Review

Review

Iredia M Olaye¹, MSc, PhD; Mia P Belovsky², BSc; Lauren Bataille³, MSc; Royce Cheng⁴, MSc, MBA; Ali Ciger⁵, MBA; Karen L Fortuna⁶, PhD, LICSW; Elena S Izmailova⁷, PhD; Debbe McCall, MBA; Christopher J Miller⁸, MStat; Willie Muehlhausen⁹, DVM; Carrie A Northcott¹⁰, PhD; Isaac R Rodriguez-Chavez¹¹, MSc, MHS, PhD; Abhishek Pratap^{12,13,14,15}, PhD; Benjamin Vandendriessche^{16,17}, PhD; Yaara Zisman-Ilani¹⁸, MA, PhD; Jessie P Bakker¹⁹, MSc, PhD

⁵Pfizer, Berlin, Germany

¹³Vector Institute, Toronto, ON, Canada

¹⁴Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, United States

¹⁶Byteflies, Antwerp, Belgium

¹⁷Department of Electrical, Computer, and Systems Engineering, Case Western Reserve University, Cleveland, OH, United States

¹⁸Department of Social and Behavioral Sciences; College of Public Health, Temple University, Philadelphia, PA, United States
¹⁹Signifier Medical Technologies, Needham, MA, United States

Corresponding Author:

Jessie P Bakker, MSc, PhD Signifier Medical Technologies 175 Highland Ave Needham, MA United States Phone: 1 6173866968 Email: jessie.b@signifiermedical.com

Abstract

Background: Suboptimal adherence to data collection procedures or a study intervention is often the cause of a failed clinical trial. Data from connected sensors, including wearables, referred to here as biometric monitoring technologies (BioMeTs), are capable of capturing adherence to both digital therapeutics and digital data collection procedures, thereby providing the opportunity to identify the determinants of adherence and thereafter, methods to maximize adherence.

Objective: We aim to describe the methods and definitions by which adherence has been captured and reported using BioMeTs in recent years. Identifying key gaps allowed us to make recommendations regarding minimum reporting requirements and consistency of definitions for BioMeT-based adherence data.

RenderX

¹Department of Medicine Division of Clinical Epidemiology and Evaluative Sciences Research, Weill Cornell Medical College Cornell University, New York, NY, United States

²Sidney Kimmel Medical College at Thomas Jefferson University, Philadelphia, PA, United States

³Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States

⁴Health Platforms, Verily Life Sciences, Cambridge, MA, United States

⁶Giesel School of Medicine at Dartmouth College, Hanover, NH, United States

⁷Koneksa Health, New York, NY, United States

⁸AstraZeneca Pharmaceuticals LP, Gaithersburg, MD, United States

⁹SAFIRA Clinical Research, Cloughjordan, Ireland

¹⁰Pfizer Inc, Cambridge, MA, United States

¹¹ICON plc, Blue Bell, PA, United States

¹²CAMH Krembil Center for Neuroinformatics, Toronto, ON, Canada

¹⁵Institute of Psychiatry, Psychology, and Neuroscience, Kings College London, London, United Kingdom

Methods: We conducted a systematic review of studies published between 2014 and 2019, which deployed a BioMeT outside the clinical or laboratory setting for which a quantitative, nonsurrogate, sensor-based measurement of adherence was reported. After systematically screening the manuscripts for eligibility, we extracted details regarding study design, participants, the BioMeT or BioMeTs used, and the definition and units of adherence. The primary definitions of adherence were categorized as a continuous variable based on duration (highest resolution), a continuous variable based on the number of measurements completed, or a categorical variable (lowest resolution).

Results: Our PubMed search terms identified 940 manuscripts; 100 (10.6%) met our eligibility criteria and contained descriptions of 110 BioMeTs. During literature screening, we found that 30% (53/177) of the studies that used a BioMeT outside of the clinical or laboratory setting failed to report a sensor-based, nonsurrogate, quantitative measurement of adherence. We identified 37 unique definitions of adherence reported for the 110 BioMeTs and observed that uniformity of adherence definitions was associated with the resolution of the data reported. When adherence was reported as a continuous time-based variable, the same definition of adherence was adopted for 92% (46/50) of the tools. However, when adherence data were simplified to a categorical variable, we observed 25 unique definitions of adherence reported for 37 tools.

Conclusions: We recommend that quantitative, nonsurrogate, sensor-based adherence data be reported for all BioMeTs when feasible; a clear description of the sensor or sensors used to capture adherence data, the algorithm or algorithms that convert sample-level measurements to a metric of adherence, and the analytic validation data demonstrating that BioMeT-generated adherence is an accurate and reliable measurement of actual use be provided when available; and primary adherence data be reported as a continuous variable followed by categorical definitions if needed, and that the categories adopted are supported by clinical validation data and/or consistent with previous reports.

(J Med Internet Res 2022;24(4):e33537) doi: 10.2196/33537

KEYWORDS

digital medicine; digital measures; adherence; compliance; mobile phone

Introduction

Background

Suboptimal adherence to clinical study procedures and/or a study intervention is often the root cause of a failed clinical trial, as it contributes to missing data; dilutes the effect of the intervention, thereby reducing statistical power; and may contribute to selection bias [1-3]. Compounding this issue is the fact that adherence is challenging to measure and account for during the analysis [4]. For example, self-reported medication adherence is straightforward to capture but is generally not a reliable measure of actual use; pill counts and pharmacy refill rates are imperfect surrogate measurements, and physical tests such as blood or urine biomarkers are expensive and difficult to administer throughout a trial [5-7]. Connected sensor technologies, such as mobile health and wearables, represent a potential solution for measuring adherence to a therapeutic intervention accurately, given that they are capable of collecting continuous, sensor-based data in real-world settings. Such technology has been increasingly used to capture trial end points [8,9]; therefore, in addition to monitoring adherence to an intervention, there exists an opportunity to monitor adherence to data collection procedures.

When conducting a clinical trial, the goal is to maximize adherence to both study procedures and interventions. However, efforts focused on increasing adherence cannot be developed until the determinants of adherence are understood. In turn, the reasons for suboptimal adherence cannot be identified unless adherence is adequately measured and reported in studies that use biometric monitoring technologies (BioMeTs), defined previously as connected digital tools that process data captured by mobile sensors using algorithms to generate measures of behavioral or physiological function [10]. A growing body of literature has offered standards and guidance to improve digital medicine study design and reporting quality [8-15]; however, best practices regarding measurement and reporting of BioMeT adherence—the extent to which the tool itself or an associated intervention is used as designed—are not as clearly conceptualized [4,16].

Objectives

A research working group from the Digital Medicine Society was formed to conduct a systematic literature review of published studies reporting adherence captured by BioMeTs to (1) identify studies that have used these tools to capture adherence to data collection procedures and/or study interventions, (2) describe the various methods used to measure adherence, and (3) compare the definitions of adherence reported in the literature. We view this description of the current state of the art as a critical first step toward identifying the determinants of adherence, in order to develop adjunct interventions to maximize adherence, ultimately contributing to improving the efficiency of clinical trials using novel technology.

Methods

Literature Search

The PubMed search terms were designed in five layers as follows: (1) used a BioMeT (layer A), (2) reported adherence or compliance (layer B), (3) were clinical studies (layer C), (4) reported original data (layer D), and (5) were published between January 1, 2014, and November 19, 2019 (layer E). Layers B, C, D, and E were based on indexing data available in PubMed, such as Medical Subject Headings terms and publication types. Layer A was designed to identify studies using a BioMeT and comprised 3 Medical Subject Headings terms as well as 34

XSL•FO RenderX

keywords including *tracker*, *implantable*, *watch*, *mobile*, and *sensor* (see Multimedia Appendix 1 for complete search terms). When developing Layer A, our goal was to be sensitive rather than specific, as we anticipated variability in how BioMeTs are described in the literature.

We have adopted the term *adherence* rather than *compliance* or *concordance* throughout this manuscript, although we recognize that these terms cover a range of inconsistent definitions including patient-driven decisions and behaviors, passively conforming to medical advice, and the extent to which a research participant follows a study protocol [17]. Specifically, we included BioMeTs that measured the use of the tool itself,

such as a wrist-worn device containing a skin capacitance sensor to monitor the duration of use, in addition to BioMeTs that measured the use of a diagnostic or therapeutic tool, such as a temperature sensor to measure the use of a dental appliance.

Systematic Review

We developed a Population, Intervention, Comparison, Outcomes, and Study design (PICOS) framework [18] to formulate the eligibility criteria for prospective studies of human participants (Textbox 1). Each study deployed at least one BioMeT outside the clinical setting or a testing facility, for which a quantitative, nonsurrogate, and sensor-based measurement of adherence was reported.

Textbox 1. Eligibility criteria adopted for literature screening in Population, Intervention, Comparison, Outcomes, and Study design order.

Eligibility criteria

Population

- Identify human studies
- Identify studies capturing in vivo data

Intervention

- Identify studies that used at least one biometric monitoring technology (BioMeT):
 - The tool must be used for purposes of measurement, diagnosis, and/or treatment of a behavioral or physiological function related to a disease state or physiological condition.
 - The tool must be mobile, meaning that it is capable of collecting data in real-world settings without oversight from trained personnel or staff.
 - The tool must be connected, meaning that there is a method to move data from the tool to the clinical or laboratory for analysis.
 - The tool must capture data via sensors of a physical property.
- Identify studies that captured BioMeT data outside of the clinical or laboratory setting.

Comparison

Not applicable

Outcomes

- Identify studies that reported adherence:
 - The tool must measure adherence directly, rather than surrogate data associated with adherence.
 - The adherence data must be quantitative.
 - The adherence data must be sensor-based rather than based on self-report, observation, and/or based on manual adjustment or scoring.

Study design

• Identify studies reporting primary analyses of prospective data collection

Within the aforementioned definition of a BioMeT [10], the term *connected* was interpreted to include any wired or wireless transfer of data; thus, products that used a physical connection for data transfer were included, but devices that only displayed data on a user interface were excluded. Similarly, we interpreted the term *mobile* broadly and included wearables, proximal sensors, ingestibles, implantables, and tools that require a brief interaction such as a smartphone. When considering *sensors*, we included only those that measured physical properties such as temperature, pressure, sound, or acceleration. Therefore, we excluded tools that contained a chronometer that relied on being

turned on or off or analyses based on self-report or other subjective assessments.

We stipulated that the measurement of adherence must be quantitative, nonsurrogate, and sensor-based. We defined *nonsurrogate* as an unequivocal reflection of product use. For example, technologies such as smart pill bottles, in which a sensor records the time at which the lid is removed, were considered a surrogate measure of adherence because they do not measure the ingestion of the pills. In contrast, smart pills that combine a pharmaceutical agent with a digital radio-frequency emitter activated by chloride ions in the

digestive system were considered as nonsurrogate adherence measurements, and therefore in scope. Finally, BioMeTs were excluded if the adherence data were based on self-report or observation or if any component of the adherence data required manual adjustment or scoring.

In total, 5 independent investigators (JPB, AC, DM, WM, and IMO) applied the PICOS criteria to a subset of 42 manuscripts for training purposes. The screening results were compared, discrepancies were discussed as a group, and the PICOS criteria were refined and clarified to optimize standardization during the remaining literature screening process. The remaining 898 manuscripts were then divided across 5 trained investigators for screening (LB, AC, WM, IMO, and BV), whereas a sixth investigator (JPB) independently screened a subset of 20% (180/898) of the manuscripts for quality assessment, as described below.

Data Extraction and Analysis

Data extraction fields included the study aim, study design (observational or interventional), therapeutic area, country of data collection, participant demographics (age, sex or gender, and race or ethnicity), information related to the BioMeT (concept of interest as described previously [19], technology type, sensor or sensors, device make and model, software name and version), and adherence data (end point definition and units). End point definitions were identified as the primary metric by which the sample-level data were analyzed to describe BioMeT adherence; for example, the duration of use, the percentage of tasks completed, or the percentage of study participants achieving a specified use goal. Therapeutic areas were categorized according to the Clinical Data Interchange Standards Consortium list [20], with additional categories for healthy and overweight or obesity. The sample size extracted from each manuscript was either the number of participants contributing to the adherence data or, if not reported, the total sample size for the study.

We categorized the primary definition of adherence for each BioMeT according to decreasing levels of data resolution as follows: (1) duration of use, either as a unit of time or percentage (continuous variable); (2) the number of measurements completed or number of days containing a measurement (continuous variable); or (3) the percentage of study participants who achieved a use goal (binary variable). Each BioMeT was categorized as passive (tools designed for continuous use) or active (tools that require user engagement at defined time points). Active BioMeTs were further categorized as session-based tools, such as connected exercise equipment, or task-based tools, such as a smart scale. This distinction was made because duration-based adherence data cannot be extracted from tools that measure one-off tasks; therefore, the highest-resolution adherence data available from these tools are the number of tasks or measurements completed.

All data were presented with descriptive statistics.

Results

Literature Screening Results

The PubMed search identified 940 manuscripts, of which 100 (10.6%) were deemed eligible for inclusion in the systematic review after meeting the PICOS criteria (Figure 1; see Multimedia Appendix 2 for all 100 papers listed). Data were extracted from these 100 manuscripts as described earlier.

After removing the date constraints from the PubMed search terms, we repeated the search by year to assess the number of publications captured from 1975 to 2020 (Figure 2).



Figure 1. Literature screening results per PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Note that all papers were assessed for eligibility based on information contained in the abstract or full text. Papers were not screened based on the title only, as it was anticipated that many studies would include biometric monitoring technology (BioMeT) data as an exploratory end point and therefore, they would be not captured in the title.





Olaye et al

Figure 2. Number of publications captured by our literature search terms over time. The solid bars indicate the publications screened for inclusion in our systematic review.



Literature Screening Quality Assessment

The large number of manuscripts identified during the PubMed search (N=940) precluded the ability of more than one independent investigator to review each against our PICOS criteria. As described earlier, of the 898 manuscripts remaining after assessment of the subset of 42 manuscripts identified for training, we randomly identified 180 (20%) manuscripts to be rescreened by an independent investigator (JPB) for quality assessment. During this process, of the 180 manuscripts, there were 13 (7%) instances in which there was disagreement regarding the classification of the manuscripts. Most of these disagreements resulted from ambiguity of reporting with respect to whether the tool in question met the definition of a BioMeT (8/13, 62% of the manuscripts), whether the device measured adherence via a sensor (1/13, 8% of the manuscripts), or the setting of data collection (1/13, 8% of the manuscripts). Finally, 23% (3/13) of the disagreements were errors in which the manuscript was inadvertently marked as not including human participants or based on in vitro analyses. An additional 8 (1.11%) manuscripts that were not part of the audit subset were marked by the reviewer as ambiguous; all of these were cross-checked by another investigator to determine eligibility.

Descriptive Data

Table 1 summarizes the study design, sample size, and participant demographics of the 100 eligible studies. The sample size ranged from 10 to 128,037 participants, with an overall median of 60 participants (IQR 35-137). Most studies (92/100, 92%) used a single BioMeT; however, a second and third BioMeT was used in 6 (6%) and 2 (2%) manuscripts, respectively. Thus, 100 studies contributed data on 110 BioMeTs.

The manufacturer and/or model were reported for 90.0% (99/110) of the BioMeTs; however, only 30.9% (34/110) reported the software name and/or version. BioMeTs were categorized according to their concept of interest, with exercise or sleep (47/110, 42.7%) and sleep-disordered breathing (25/110, 22.7%) being the most common. These 2 categories also contained the highest number of BioMeT tool types; for example, exercise or sleep was captured by wearables, chest straps, smart clothing or footwear, and smartphones, whereas the sleep-disordered breathing category included positive airway pressure devices, chest straps for the treatment of positional obstructive sleep apnea, oral appliances, and implantable nerve stimulator devices. Notably, proximal sensors were considered in the scope, but none were captured in our systematic review.



Olaye et al

Table 1. Study details, demographic data, and biometric monitoring technologies (BioMeTs) by therapeutic area.

	Therapeutic area of focus								
	All (N=100)	Healthy (n=11)	Cardiovas- cular (n=17)	Endocrine (n=13)	Neural (n=10)	Over- weight or obesity (n=6)	Respiratory (n=29)	Pain treat- ments (n=5)	Other ^a (n=9)
Study design, n (%)			·						
Observational studies	26 (26)	3 (27)	1 (6)	4 (31)	3 (30)	1 (17)	9 (31)	1 (20)	4 (44)
Interventional studies	74 (74)	8 (73)	16 (94)	9 (69)	7 (70)	5 (83)	20 (69)	4 (80)	5 (56)
Sample size (participants), median; range	60; 10- 128,037	179; 42- 1381	84; 40- 1732	46; 10-234	22; 10-780	86; 11-174	70; 10- 128,037	35; 10-68	56; 20- 281
Participant characteristics,	n (%)								
Sex or gender									
Females or women only ^b	9 (9)	3 (27)	1 (6)	0 (0)	0 (0)	3 (50)	0 (0)	0 (0)	2 (22)
Both sexes or gen- ders	84 (84)	7 (64)	16 (94)	12 (92)	8 (80)	3 (50)	26 (90)	5 (100)	7 (78)
Not reported	7 (7)	1 (9)	0 (0)	1 (8)	2 (20)	0 (0)	3 (10)	0 (0)	0 (0)
Age (years), n (%)									
≥60	24 (24)	4 (36)	4 (24)	5 (38)	2 (20)	0 (0)	5 (17)	1 (20)	3 (33)
>21 to <60	57 (57)	5 (45)	9 (53)	4 (31)	5 (50)	3 (50)	22 (76) ^c	3 (60)	6 (67)
≤21	19 (19)	2 (18)	4 (24)	4 (31) ^c	3 (30)	3 (50)	2 (7)	1 (20)	0 (0)
Race or ethnicity, n (%)									
Reported	39 (39)	7 (64)	11 (65)	6(46)	2 (20)	4 (67)	3 (10)	2 (40)	4 (44)
Not reported	61 (61)	4 (36)	6 (35)	7 (54)	8 (80)	2 (33)	26 (90)	3 (60)	5 (56)
BioMeT tool type, n, (%)									
Wearable	46 (42)	10 (91)	9 (41)	4 (27)	6 (60)	5 (83)	2 (7)	5 (83)	5 (50)
Positive airway pressure device	18 (16)	d	_	—	_	_	18 (60)	_	—
Smart clothing	8 (7)	—	3 (14)	3 (20)	1 (10)	—	—	—	1 (10)
Blood pressure monitor	6 (5)	_	5 (23)	1 (7)	_	_	_	_	_
Chest strap	5 (5)	_	2 (9)	_	—	_	3 (10)	_	_
Smartphone	5 (5)	—	1 (5)		2 (20)	1 (17)	_	—	1 (10)
Oral appliance	5 (5)	_	—	_	—	_	2;7%	_	3 (30)
Glucometer; continuous	3 (3)	—	—	3 (20)	_	_	_	_	—
Glucometer; noncontinu- ous	3 (3)	_	_	3 (20)	_	_	_	_	—
Ingestible	2 (2)	—	1 (5)	_	—	_	—	1 (17)	—
Implantable	2 (2)	—	—	—	—	_	2 (7)	—	—
Smart scale	2 (2)	—	1 (5)	1 (7)	—	—	—	—	—
Adhesive patch	1 (1)	1 (9)	—		—	—	—	—	—
Exercise equipment	1 (1)	—	—	_	—	—	1 (3)	—	—
Muscle trainer	1 (1)	—	—	—	—		1 (3)	—	—
Hearing aid	1 (1)	—	—	—	1 (10)		—	—	—
Home oxygen	1 (1)	—	—	—	—	—	1 (3)	—	—
Total	110	11	22	15	10	6	30	6	10



XSL•FO RenderX

^aOther category included oncology, gastrointestinal, bone structure, anatomy, or orthodontics, pregnancy, and vocal cord dysfunction. ^bNo studies included only males or men.

^cEach of these categories contained 1 study that reported age only qualitatively or by providing a range; all other studies reported an average age. d No studies falling into that category.

Adherence Data

Overall, we identified 37 unique definitions for the 110 BioMeTs. The most commonly reported definition (duration of use) was reported for 41.8% (46/110) of the tools; however, the next most common definitions (number or percentage of tasks completed and number or percentage of days with data) were reported for only 8.2% (9/110) and 6.4% (7/110) of the BioMeTs, respectively.

As shown in Table 2, each BioMeT was categorized as passive (69/110, 62.7%), session-based (24/110, 21.8%), or task-based (17/110, 15.5%). The duration of use was reported for 46% (32/69) of the passive BioMeTs and 75% (18/24) of the session-based BioMeTs. Of the task-based BioMeTs for which the duration of use could not be meaningfully reported, the highest resolution of adherence data (the number of measurements or days) was reported for 41% (7/17) of the BioMeTs. The lowest resolution of adherence data (achievement of a goal as a binary or categorical variable) was reported for 33.6% (37/110) of all BioMeTs.

Table 2.	Adherence data res	olution and def	inition captured	by passive and	active biometric	monitoring tech	nologies (BioMeTs)
----------	--------------------	-----------------	------------------	----------------	------------------	-----------------	------------	----------

Parameters Highest resolution adherence data					Lowest resolution adherence data		
	Duration of use (variable)	based on a continuous	Number of mease (based on a cont	urements or days used inuous variable)	Achievement of a goal (based on a bi- nary variable)		
	Number of BioMeTs, n (%)	Number of unique adherence defini- tions	Number of BioMeTs, n (%)	Number of unique adherence defini- tions	Number of BioMeTs, n (%)	Number of unique adherence defini- tions	
Monitoring type		·	•			·	
Passive	32 (64)	4	16 (70)	7	21 (57)	19	
Active; session-based	18 (36)	1	0 (0)	0	6 (16)	2	
Active; task-based	N/A ^a	N/A	7 (30)	4	10 (27)	5	
All BioMeTs	50 (100)	4 ^b	23 (100)	8 ^b	37 (100)	25 ^b	
All BioMeTs apart from sleep-disordered breathing	30 (60)	4	23 (100)	8	32 (86)	24	

^aN/A: not applicable.

^bThese data are not simply the sum of the rows above, as there were instances where the same adherence definition was adopted for different tool types.

As shown in Figure 3, the number of unique definitions of adherence increased as the resolution of the reported data decreased. For example, adherence was reported as the duration of use (highest resolution) for 50 BioMeTs, 46 (92%) of which reported an actual unit of time, along with 3 other unique definitions of adherence, such as the duration of use in which a heart rate goal was achieved. In contrast, among the 37

BioMeTs for which adherence was reported as a categorical variable (lowest resolution), 25 unique definitions of adherence were identified. Sleep-disordered breathing was the BioMeT category with the most consistent definitions of adherence (2 definitions reported for 25 BioMeTs); however, the pattern of decreasing uniformity score alongside decreasing data resolution persisted after removal of these BioMeTs (Table 2).



Figure 3. Uniformity of adherence definitions according to whether the biometric monitoring technology (BioMeT) was a passive, session-based, or task-based tool. Passive BioMeTs are those designed for continuous use. Active BioMeTs are those that require user engagement at defined time points, further categorized as session-based (for which duration of use is meaningful) versus task-based (for which the duration of use is not meaningful). The colored bands represent unique definitions of adherence within each bar. The colors are comparable across the bars within each category of adherence definition (duration of use, number of measurements, and categorical variables).



Discussion

Principal Findings

The purpose of this review was to describe the various approaches taken to evaluate adherence to study procedures and/or interventions using BioMeTs in recent clinical studies and discuss best practices that can improve the reliability and comparability of adherence measurements to support further BioMeT evaluation and decision-making in both research and clinical care settings. Notably, we found that 29.9% (53/177) of the studies that used a BioMeT outside the clinical or laboratory setting failed to report a sensor-based, nonsurrogate, quantitative measurement of adherence, thus impeding a complete understanding of the study data. Among the 100 studies that reported sensor-based adherence data, we found substantial variability in terms of the definitions of adherence adopted, and that the degree of variability was associated with the resolution of the data reported. For example, when adherence was reported as a continuous time variable, the same definition of adherence was adopted for 92% (46/50) of the tools. However, when the adherence data were simplified to a categorical variable, we observed 25 unique definitions of adherence reported for 37 tools, and the most common definition was adopted for only 19% (7/37) of the BioMeTs. Examples of adherence definitions that were reported only once each within our data set include the percentage of participants with use <85% of the total time (passive BioMeT), the percentage of participants with use of ≥ 4 hours on ≥ 70 days (active, session-based BioMeT), and the percentage of participants completing readings on 100% of days (active, task-based BioMeT). All 3 of these adherence definitions were relevant and useful for the study in question; however, by adopting a specific threshold and reporting adherence as the percentage of the sample that achieved the goal, the adherence data were not readily interpretable against other studies. If adherence data were provided as higher-resolution variables, such as duration of use or number of readings, readers would be better positioned

```
https://www.jmir.org/2022/4/e33537
```

RenderX

to make comparisons. Considering that adherence to any given procedure or intervention is a critical driver of desired behavior change and improved health outcomes in research and real-world settings [21], greater consistency in defining adherence may help more clearly associate BioMeT adherence to study outcomes and may offer a critical lens in the design and implementation of customized BioMeTs that are fit-for-purpose within their context of use.

In addition to consistent reporting of adherence, it is critical to understand exactly what digital medicine tools are used in a given study, consistent with the 2021 EVIDENCE (Evaluating Connected Sensor Technologies) Publication Checklist [11]. A complete description of the tool used ensures reproducibility, allows for meaningful comparisons across studies, and opens up the possibility of merging data across cohorts. Although the manufacturer or model (or both) were reported for 90.0% (99/110) of the tools captured in our review, this still leaves 10.0% (11/110) for which the tools were described in generic terms. Moreover, we found that the software name or version (or both) used for data processing was reported for only 30.9% (34/110) of the BioMeTs, indicating that the data cannot be reproduced even if the hardware details are known. We also noted key gaps in the descriptive data; most notably, only 41% (41/100) reported the ethnicity or race of participants, which is a persistent problem in clinical research [22,23]. Even among the 46 studies performed in the United States, for which there are clear guidelines for collecting and reporting race and ethnicity [24], 30% (14/46) of the studies did not provide these data. Age, sex or gender, and race or ethnicity are essential for understanding the representativeness of study samples and generalizability of findings and reflect only a subset of a broader set of characteristics such as socioeconomic information that must be captured and reported to understand how BioMeT adherence relates to issues of access, uptake, equity, and equality [25,26].

Differences across studies using BioMeTs are inevitable; however, ideally, there should be standardization and harmonization of collecting and reporting adherence to allow for the evaluation, interpretation, and statistical comparison of outcomes. Thus, although this study was not designed to develop standards, the aforementioned gaps and shortcomings have led us to recommend minimum reporting requirements and that consistent definitions or units for adherence should be adopted. Specifically, we recommend that (1) quantitative, nonsurrogate, sensor-based adherence data be reported for all BioMeTs when feasible; (2) a clear description of the sensor or sensors used to capture adherence data, the algorithm or algorithms that convert sample-level measurements to a metric of adherence, and the analytic validation data demonstrating that BioMeT-generated adherence is an accurate and reliable measurement of actual use be provided when available; and (3) primary adherence data be reported as a continuous variable followed by categorical definitions if needed, and that the categories adopted are supported by clinical validation data or consistent with previous reports (or both). These recommendations are in addition to the minimum requirements recommended elsewhere, such as providing a description of verification, validation, and usability data explaining the fit-for-purpose characteristics of the BioMeT

technology used within a specific context as well as detailed demographic and descriptive data for the study sample [10,11,15]. More detailed descriptions of our recommendations for reporting BioMeT adherence are provided in Table 3, which includes a case study that we identified as an exemplar that followed all included recommendations [27].

On a positive note, it is clear that BioMeTs have been increasingly deployed in clinical research studies. Repeating our PubMed search over successive years revealed that the number of publications captured in our search terms increased steadily between 1975 and 2005 and became increasingly prevalent until 2015. The reduced number of papers captured in 2020 may reflect a delay in PubMed indexing and possibly a reduced submission and/or acceptance rate during the initial stages of the COVID-19 pandemic. It is encouraging to observe adherence data reported from a wide range of monitoring, diagnostic, and therapeutic tools from studies conducted in 22 different countries. Furthermore, although accelerometry-based tools for estimating sleep and activity have been in use for several decades [28], our literature search captured adherence data for more recently developed tools, such as automated speech assessments [29] and upper-limb training systems for motor disorders [30].



Olaye et al

Table 3. Recommendations for capturing and reporting adherence measured by biometric monitoring technologies (BioMeTs).

Identified gaps and recommendations	Case study [27]			
Gap 1: Quantitative, nonsurrogate, sensor-based adherence data were no data outside the clinical or laboratory setting.	t reported in 29.9% of screened manuscripts that captured BioMeT			
Recommendation 1: Investigators are encouraged to develop and/or use BioMeT sensors to capture sensor-based adherence data in addition to their primary purpose.	This study aimed to evaluate adherence to a physical activity among students recruited from 20 schools. Quantitative adherence data were derived from wrist-worn accelerometers, considered a direct reflection of wear-time.			
Recommendation 2: Where feasible, we encourage investigators to collect and report adherence data that are a direct reflection of actual use, rather than a surrogate.	N/A ^a			
Gap 2: BioMeT manufacturer or model and software information was n	nissing for 10% and 68% of included tools, respectively.			
Recommendation 3: In addition to reporting the BioMeT manufacturer or model and software used for generating adherence data (where appli- cable), we recommend that investigators provide a clear description of the sensor or sensors capturing adherence data.	BioMeT model: GENEActiv wrist-worn device (ActivInsights Ltd). Sensor description: 3-axis accelerometer. Software: GENEActiv PC software (version 2.9), with subsequent signal processing performed in R-package (GGIR; version 1.2-2).			
Recommendation 4: We recommend that investigators describe the algo- rithm or algorithms that convert sample-level measurements into a measurement of adherence. If a description is not available from the manufacturer, this should be stated.	The paper included the data sampling frequency (100 Hz); a description of the signal processing steps including calibration; the epoch length (5 seconds) over which the sample-level data were averaged; and the units (milligravitational units; m g). A description of the nonwear detection algorithm was summarized as, "Non-wear is estimated on the basis of the SD and value range of each axis, calculated for 60-min windows with 15-min sliding window. The window is classified as non-wear if, for at least two of the three axes, the SD is less than 13 mg or the value range is less than 50 mg."			
Recommendation 5: We recommend that investigators describe the ana- lytic validation data supporting the adherence algorithm; that is, the data indicating that adherence per the BioMeT is an accurate estimate of ac- tual use. If analytic validation data is not available, this should be stated.	A reference to previous verification and analytic validation work was included.			
Gap 3: Heterogeneity of adherence definitions increased alongside decre	asing resolution of adherence data reported.			
Recommendation 6: We recommend that investigators using BioMeTs that are either passive (designed to capture data passively over long pe-	The BioMeT was categorized as passive, as the wrist-worn accelerom- eter was designed to capture data continuously over 3 separate periods			

riods) or session-based (designed for user engagement at certain time points, for which the duration of use is meaningful) report primary adherence as a continuous variable of time; that is, total minutes or hours or days, or average hours per day, days per week, and so on. Example of a passive BioMeT: smart clothing. Example of a session-based BioMeT: connected exercise equipment.

Recommendation 7: We recommend that investigators using BioMeTs that are task based (designed for user engagement at certain time points, for which the duration of use is not meaningful) report primary adherence as a continuous variable; that is, the number of tasks or days completed. Example of a task-based BioMeT: connected scale.

Recommendation 8: We recommend that categorical adherence data are reported only in addition to continuous adherence data; for example, the percentage of participants with use >x hours per day or percentage of participants completing >y tasks.

Recommendation 9: We recommend that categorical definitions of adherence be based on clinical validation data indicating the level of adherence associated with a clinically meaningful change in the outcome of interest, when available. If clinical validation data are not available, this should be stated.

^aN/A: not applicable.

Strengths and Limitations

To our knowledge, this is the first systematic review to focus specifically on BioMeT adherence. Further strengths of our study include the large number of manuscripts screened for potential inclusion, and the quality control processes that we

```
https://www.jmir.org/2022/4/e33537
```

RenderX

of 7 days. Adherence was reported as the total hours of wear-time, and hours per day of wear-time.

N/A, as the BioMeT was categorized as passive rather than task based.

Categorical adherence data included the number of participants with \geq 16 hours of wear-time per day.

The investigators include a reference to previous work that adopted the threshold of ≥16 hours of wear-time per day and describe another study that compared thresholds of 8 hours, 16 hours, and 24 hours of weartime.

implemented, which resulted in few disagreements among the reviewers. The sensitive, rather than specific, search terms we adopted increased our confidence that we were able to capture the relevant set of literature, given our initial concern that many BioMeTs were used for exploratory analyses and therefore not referred to in study titles, abstracts, or keywords. The quality

control process was particularly important, given that 76.4% (718/940) of the manuscripts were screened by a single investigator. Alongside these strengths, this review has several limitations that should be noted. Owing to the inconsistencies in the study outcome measures and variability in the definition of adherence adopted across studies, we did not undertake a methodological assessment and could not determine statistical inference. We also did not extract every element of the study design, such as the duration of the study itself or the length of time the BioMeTs were used per protocol. We included only peer reviewed publications indexed in PubMed; therefore, our findings may not be representative of all studies capturing BioMeT data in related fields, such as engineering. Finally, due to the vast number of manuscripts captured in our PubMed search terms, we limited the time frame to a 5-year period, and we limited our review to studies reporting nonsurrogate measurements of adherence, thereby excluding technology such as smart pill bottles which may offer valuable adherence data where a nonsurrogate measurement is not feasible.

Conclusions

This review provides a description of the numerous methods that have been used in recent years to measure BioMeT adherence, allowing us to identify gaps and make specific reporting recommendations. Several important questions remain, which we hope will be addressed in future studies. For example, it will be interesting to compare our findings to a similar review covering the subsequent 5-year period (2020-2025), as the abrupt acceleration of digital monitoring and interventions including telemedicine during the current COVID-19 era [31,32] will likely, in hindsight, be considered a paradigm shift in both research and health care delivery. We hope that with increased consistency and reporting of data elements, it will become possible to meta-analyze adherence data to identify the possible determinants of BioMeT use patterns. Only when adherence data are adequately reported will the field of digital medicine be able to advance our understanding of the reasons underlying acceptance and adherence, which will ultimately allow investigators to optimize the design of tools, studies, implementation methods, and user engagement strategies to realize the full potential of BioMeTs as digital monitoring, diagnostic, and therapeutic tools. To support these actions, we recommend that (1) quantitative, nonsurrogate, sensor-based adherence data be reported for all BioMeTs when feasible; (2) a clear description of the sensor or sensors used to capture adherence data, the algorithm or algorithms that convert sample-level measurements to a metric of adherence, and the analytic validation data demonstrating that BioMeT-generated adherence is an accurate and reliable measurement of actual use be provided when available; and (3) primary adherence data be reported as a continuous variable followed by categorical definitions if needed, and that the categories adopted are supported by clinical validation data and/or consistent with previous reports.

Acknowledgments

This publication is a result of collaborative research performed under the auspices of the Digital Medicine Society (DiMe). DiMe is a 510(c)(3) nonprofit professional society for the digital medicine community and is not considered a sponsor of this work. All authors are members of DiMe, who volunteered to participate in this systematic review. All DiMe research activities are overseen by a research committee, the members of which were invited to comment on the manuscript before submission.

Conflicts of Interest

The authors have reported the following conflicts of interest through employment or stock ownership: JPB (Signifier Medical Technologies and Philips), LB (Novartis), RC (Verily Life Sciences), AC (Pfizer and Ali Ciger Ventures UG [haftungsbeschränkt]), KLF (K Health, Trusst Health Inc, InquistHealth, and Social Wellness), ESI (Koneksa Health), CJM (Astra Zeneca and Abbvie), CAN (Pfizer), IRRC (ICON plc), and BV (Byteflies).

Multimedia Appendix 1

PubMed search terms. [DOCX File, 100 KB-Multimedia Appendix 1]

Multimedia Appendix 2

All manuscripts identified for data extraction. [DOCX File , 124 KB-Multimedia Appendix 2]

References

- 1. Breckenridge A, Aronson JK, Blaschke TF, Hartman D, Peck CC, Vrijens B. Poor medication adherence in clinical trials: consequences and solutions. Nat Rev Drug Discov 2017;16(3):149-150. [doi: <u>10.1038/nrd.2017.1</u>] [Medline: <u>28154411</u>]
- 2. Shiovitz TM, Bain EE, McCann DJ, Skolnick P, Laughren T, Hanina A, et al. Mitigating the effects of nonadherence in clinical trials. J Clin Pharmacol 2016;56(9):1151-1164 [FREE Full text] [doi: 10.1002/jcph.689] [Medline: 26634893]
- 3. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. Nephron Clin Pract 2010;115(2):c94-c99 [FREE Full text] [doi: 10.1159/000312871] [Medline: 20407272]

- 4. Roe D, Jones N, Hasson-Ohayon I, Zisman-Ilani Y. Conceptualization and study of antipsychotic medication use: from adherence to patterns of use. Psychiatr Serv 2021;72(12):1464-1466. [doi: 10.1176/appi.ps.202100006] [Medline: 34126781]
- Valdés Y Llorca C, Cortés-Castell E, Ribera-Casado JM, de Lucas-Ramos P, de Palacio-Guerrero LM, Mugarza-Borqué F, et al. Validation of self-reported adherence in chronic patients visiting pharmacies and factors associated with the overestimation and underestimation of good adherence. Eur J Clin Pharmacol 2020;76(11):1607-1614. [doi: 10.1007/s00228-020-02950-9] [Medline: 32613537]
- 6. Anghel LA, Farcas AM, Oprean RN. An overview of the common methods used to measure treatment adherence. Med Pharm Rep 2019;92(2):117-122 [FREE Full text] [doi: 10.15386/mpr-1201] [Medline: 31086837]
- Stirratt MJ, Dunbar-Jacob J, Crane HM, Simoni JM, Czajkowski S, Hilliard ME, et al. Self-report measures of medication adherence behavior: recommendations on optimal use. Transl Behav Med 2015;5(4):470-482 [FREE Full text] [doi: 10.1007/s13142-015-0315-2] [Medline: 26622919]
- Perry B, Herrington W, Goldsack JC, Grandinetti CA, Vasisht KP, Landray MJ, et al. Use of mobile devices to measure outcomes in clinical research, 2010-2016: a systematic literature review. Digit Biomark 2018;2(1):11-30 [FREE Full text] [doi: 10.1159/000486347] [Medline: 29938250]
- 9. Walton MK, Cappelleri JC, Byrom B, Goldsack JC, Eremenco S, Harris D, et al. Considerations for development of an evidence dossier to support the use of mobile sensor technology for clinical outcome assessments in clinical trials. Contemp Clin Trials 2020;91:105962 [FREE Full text] [doi: 10.1016/j.cct.2020.105962] [Medline: 32087341]
- Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digit Med 2020;3:55 [FREE Full text] [doi: 10.1038/s41746-020-0260-4] [Medline: 32337371]
- Manta C, Mahadevan N, Bakker J, Ozen Irmak S, Izmailova E, Park S, et al. EVIDENCE publication checklist for studies evaluating connected sensor technologies: explanation and elaboration. Digit Biomark 2021;5(2):127-147 [FREE Full text] [doi: 10.1159/000515835] [Medline: 34179682]
- Godfrey A, Goldsack JC, Tenaerts P, Coravos A, Aranda C, Hussain A, et al. BioMeT and algorithm challenges: a proposed digital standardized evaluation framework. IEEE J Transl Eng Health Med 2020;8:0700108 [FREE Full text] [doi: 10.1109/JTEHM.2020.2996761] [Medline: 32542118]
- Coravos A, Doerr M, Goldsack J, Manta C, Shervey M, Woods B, et al. Modernizing and designing evaluation frameworks for connected sensor technologies in medicine. NPJ Digit Med 2020;3:37 [FREE Full text] [doi: 10.1038/s41746-020-0237-3] [Medline: 32195372]
- Goldsack JC, Dowling AV, Samuelson D, Patrick-Lake B, Clay I. Evaluation, acceptance, and qualification of digital measures: from proof of concept to endpoint. Digit Biomark 2021;5(1):53-64 [FREE Full text] [doi: 10.1159/000514730] [Medline: <u>33977218</u>]
- Coran P, Goldsack JC, Grandinetti CA, Bakker JP, Bolognese M, Dorsey ER, et al. Advancing the use of mobile technologies in clinical trials: recommendations from the clinical trials transformation initiative. Digit Biomark 2019;3(3):145-154 [FREE Full text] [doi: 10.1159/000503957] [Medline: 32095773]
- 16. Klonoff DC. Behavioral theory: the missing ingredient for digital health tools to change behavior and increase adherence. J Diabetes Sci Technol 2019;13(2):276-281 [FREE Full text] [doi: 10.1177/1932296818820303] [Medline: 30678472]
- 17. Chakrabarti S. What's in a name? Compliance, adherence and concordance in chronic psychiatric disorders. World J Psychiatry 2014;4(2):30-36 [FREE Full text] [doi: 10.5498/wjp.v4.i2.30] [Medline: 25019054]
- Tacconelli E. Systematic reviews: CRD's guidance for undertaking reviews in health care. Lancet Infect Dis 2010;10(4):226. [doi: <u>10.1016/s1473-3099(10)70065-7</u>]
- Bakker JP, Goldsack JC, Clarke M, Coravos A, Geoghegan C, Godfrey A, et al. A systematic review of feasibility studies promoting the use of mobile technologies in clinical research. NPJ Digit Med 2019;2:47 [FREE Full text] [doi: 10.1038/s41746-019-0125-x] [Medline: 31304393]
- 20. Therapeutic areas. Clinical Data Interchange Standards Consortium. URL: <u>https://www.cdisc.org/standards/therapeutic-areas</u> [accessed 2021-08-11]
- 21. Brown MT, Bussell JK. Medication adherence: WHO cares? Mayo Clin Proc 2011;86(4):304-314 [FREE Full text] [doi: 10.4065/mcp.2010.0575] [Medline: 21389250]
- Rochon PA, Mashari A, Cohen A, Misra A, Laxer D, Streiner DL, et al. The inclusion of minority groups in clinical trials: problems of under representation and under reporting of data. Account Res 2004;11(3-4):215-223. [doi: 10.1080/08989620490891412] [Medline: 15812967]
- Medina LD, Torres S, Gioia A, Ochoa Lopez A, Wang J, Cirino PT. Reporting of demographic variables in neuropsychological research: an update of O'Bryant et al.'s trends in the current literature. J Int Neuropsychol Soc 2021;27(5):497-507. [doi: 10.1017/S1355617720001083] [Medline: <u>33176898</u>]
- 24. Evaluation and reporting of age-, race-, and ethnicity-specific data in medical device clinical studies: guidance for industry and food and drug administration staff. U.S. Food & Drug Administration. 2017. URL: <u>https://www.fda.gov/media/98686/</u> <u>download</u> [accessed 2022-03-15]

```
https://www.jmir.org/2022/4/e33537
```

RenderX

- 25. Kontos E, Blake KD, Chou WY, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
- 26. Azzopardi-Muscat N, Sørensen K. Towards an equitable digital public health era: promoting equity through a health literacy perspective. Eur J Public Health 2019;29(Supplement_3):13-17 [FREE Full text] [doi: 10.1093/eurpub/ckz166] [Medline: 31738443]
- Rowlands AV, Harrington DM, Bodicoat DH, Davies MJ, Sherar LB, Gorely T, et al. Compliance of adolescent girls to repeated deployments of wrist-worn accelerometers. Med Sci Sports Exerc 2018;50(7):1508-1517. [doi: 10.1249/MSS.000000000001588] [Medline: 29474208]
- 28. Martin JL, Hakim AD. Wrist actigraphy. Chest 2011;139(6):1514-1527 [FREE Full text] [doi: 10.1378/chest.10-1872] [Medline: 21652563]
- 29. Whitling S, Lyberg-Åhlander V, Rydell R. Absolute or relative voice rest after phonosurgery: a blind randomized prospective clinical trial. Logoped Phoniatr Vocol 2018;43(4):143-154. [doi: 10.1080/14015439.2018.1504985] [Medline: 30183437]
- 30. Gerber CN, Kunz B, van Hedel HJ. Preparing a neuropediatric upper limb exergame rehabilitation system for home-use: a feasibility study. J Neuroeng Rehabil 2016;13:33 [FREE Full text] [doi: 10.1186/s12984-016-0141-x] [Medline: 27008504]
- Seshadri DR, Davies EV, Harlow ER, Hsu JJ, Knighton SC, Walker TA, et al. Wearable sensors for COVID-19: a call to action to harness our digital infrastructure for remote patient monitoring and virtual assessments. Front Digit Health 2020;2:8 [FREE Full text] [doi: 10.3389/fdgth.2020.00008] [Medline: 34713021]
- 32. Goldsack JC, Izmailova ES, Menetski JP, Hoffmann SC, Groenen PM, Wagner JA. Remote digital monitoring in clinical trials in the time of COVID-19. Nat Rev Drug Discov 2020;19(6):378-379. [doi: <u>10.1038/d41573-020-00094-0</u>] [Medline: <u>32409759</u>]

Abbreviations

BioMeT: biometric monitoring technologyDiMe: Digital Medicine SocietyPICOS: Population, Intervention, Comparison, Outcomes, and Study design

Edited by A Mavragani; submitted 11.09.21; peer-reviewed by B Peterson, C Drummond, M Nissen; comments to author 20.10.21; revised version received 03.11.21; accepted 14.01.22; published 14.04.22 <u>Please cite as:</u> Olaye IM, Belovsky MP, Bataille L, Cheng R, Ciger A, Fortuna KL, Izmailova ES, McCall D, Miller CJ, Muehlhausen W, Northcott CA, Rodriguez-Chavez IR, Pratap A, Vandendriessche B, Zisman-Ilani Y, Bakker JP Recommendations for Defining and Reporting Adherence Measured by Biometric Monitoring Technologies: Systematic Review

J Med Internet Res 2022;24(4):e33537 URL: <u>https://www.jmir.org/2022/4/e33537</u> doi: <u>10.2196/33537</u> PMID:

©Iredia M Olaye, Mia P Belovsky, Lauren Bataille, Royce Cheng, Ali Ciger, Karen L Fortuna, Elena S Izmailova, Debbe McCall, Christopher J Miller, Willie Muehlhausen, Carrie A Northcott, Isaac R Rodriguez-Chavez, Abhishek Pratap, Benjamin Vandendriessche, Yaara Zisman-Ilani, Jessie P Bakker. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 14.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.

