Protocol

A Dual In-Person and Remote Assessment Approach to Developing Digital End Points Relevant to Autism and Co-Occurring Conditions: Protocol for a Multisite Observational Study

Isabel Yorke[1,2], PhD; Charlotte A Boatman[3], MSc; Akash Roy Choudhury[4], MSc; Bethany Oakley[3], PhD; Pauline Conde[4], BSc; Heet Sankesara[4], BSc; Yatharth Ranjan[4], MSc; Zulqarnain Rashid[4], PhD; Judith Dineley[2,4], PhD; Johnny Downs[2], MRCPsych, PhD; Christopher H Chatham[5,6], PhD; Nicholas Cummins[2,4], PhD; Amos Folarin[4,7,8,9,10], PhD; Eva Loth[3], PhD; Jan Buitelaar[11], PhD, MD; Declan Murphy[3], MD; Richard Dobson[4,7,8,9,10], PhD; Emily Simonoff[1], PhD, MD

[1]Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[2]CAMHS Digital Lab, Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[3]Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, United Kingdom

[4]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, United Kingdom

[5]Current Affiliation: Biomarker Research, Shionogi, Inc, Florham Park, NJ, United States

[6]Pharma Research & Early Development, F. Hoffmann La Roche, Ltd, Little Falls, NJ, United States

[7]Institute of Health Informatics, University College London, London, United Kingdom

[8]NIHR Biomedical Research Center, South London and Maudsley NHS Foundation Trust, London, United Kingdom

[9]Health Data Research UK, University College London, London, United Kingdom

[10]NIHR Biomedical Research Center, University College London Hospitals NHS Foundation Trust, London, United Kingdom

[11]Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands

Corresponding Author:
Isabel Yorke, PhD
Department of Child and Adolescent Psychiatry
Institute of Psychiatry, Psychology and Neuroscience
King's College London
PO85
16 De Crespigny Park
London, SE5 8AF
United Kingdom
Phone: 44 020 7848 ext 5312
Email: isabel.d.yorke@kcl.ac.uk

Abstract

Background: Research priorities for autistic people include developing effective interventions for the numerous challenges affecting their daily living, for example, mental health problems, sleep difficulties, and social well-being. However, clinical research progress is limited by a lack of validated objective measures that represent target outcomes for improvement. Digital technologies, including wearable devices and smartphone apps, provide opportunities to develop novel measures that may reflect everyday experience and complement key clinical assessments. However, little is known about the acceptability and feasibility of implementing digital data collection in this population.

Objective: The primary objective of this study is to evaluate the usability, acceptability, adherence, and feasibility of a dual in-person and remote (ie, at-home) protocol. Secondarily, we aim to explore the properties of certain resulting data with a view to developing novel digital end points for key target outcomes, including social communication, sleep, and mental health.

Methods: Eligible autistic and nonautistic in the AIMS Longitudinal European Autism Project were invited to participate in a digitally augmented in-person Autism Diagnostic Observation Schedule-2 (ADOS-2) and a 28-day remote measurement (RM) protocol involving wearing a Fitbit device, downloading a passive smartphone data collection app, and using 2 active reporting apps.

Results: The first AIMS Longitudinal European Autism Project study participants were enrolled in September 2021 (in-person component) and March 2022 (RM component). To date, 190 participants have taken part in the digitally augmented ADOS-2 component, and 86 participants have been enrolled for the RM protocol. Recruitment is now complete with some RM data collection ongoing until August 2025. Data analysis has commenced, including qualitative framework analysis of feedback interview data coproduced with autism community members, exploration of acceptability and feasibility metrics, pipeline development for ADOS-2 speech analysis, and RM sleep measures.

Conclusions: This study lays important groundwork in understanding the acceptability and feasibility of in-person and remotely implemented digital measurement procedures to capture meaningful outcomes in domains important to improving everyday life for autistic people.

International Registered Report Identifier (IRRID):   DERR1-10.2196/71145

Introduction

Background

Autism is a lifelong neurodevelopmental condition characterized by social communication differences, restricted or repetitive behaviors and interests, and differences in sensory experience [1]. Presentation across individuals is diverse and often complicated by co-occurring conditions, such as intellectual disability, attention-deficit/hyperactivity disorder (ADHD), mental health problems, epilepsy, and sleep difficulties [2]. Most autistic people face considerable challenges affecting numerous domains of everyday life, such as accessing health care, community activities, education, and employment [3]. Across studies consulting the autism community, major research priorities include enabling activities of daily living and enhancing quality of life by treating co-occurring conditions and alleviating difficulties associated with core characteristics (eg, providing training for social, vocational, and life skills and by making environments more autism friendly) [4]. However, progress in addressing these priorities is limited not only by the development of effective intervention and support strategies but also by the measurement tools available to capture desired changes in response to intervention [5-7].

Current Measurement

Self- or informant-report instruments and clinical observations or assessments form most outcome measures used in clinical autism research in domains such as social functioning [8], mental health [6], and ADHD [9]. Self-report measures are well-placed to capture real-world experiences and behavior. However, high rates of alexithymia and difficulties or differences in interpreting questionnaire items may present specific challenges for autistic people reporting on their own feelings and behaviors [10,11]. Parents, teachers, and other informants reporting on behaviors in different contexts share only modest variance, with characteristics such as parenting stress accounting for discrepancies [12,13]. When evaluating psychosocial treatment approaches, which play a key role in supporting autistic people and their families, findings may be impacted by difficulty in blinding informants involved in the treatment itself [14]. Finally, even where blinding is possible, self- or informant-report measures may show spontaneous improvement in the absence of treatment, which makes treatment effects more difficult to establish [15,16].

Assessments carried out by clinicians and researchers mitigate certain limitations; for example, in pharmacological trials targeting core autistic features (eg, social communication difficulties), placebo effects were lower for clinician versus caregiver ratings [16]. However, these are time-consuming and therefore costly to implement at scale. They may also be unrepresentative of everyday behavior or experiences because they are often conducted in unfamiliar settings that are potentially anxiety provoking and challenging for autistic people to access. Finally, whether self- or informant reported, such measures are often ordinal and limited in their effective resolution, compromising sensitivity [2] and inflating thresholds for detecting clinically meaningful results [17]. Given the profile of strengths and limitations of measurement options based on human judgment, many authors have called for the development of objective measurement tools to complement these.

Digital Measurement Opportunities

Table 1 summarizes potential areas in which digital technology could address the limitations of current measurement, as well as outlining potential challenges for measure development based on these, including those that may be particularly problematic in autism research. Smartphones and wearable devices, such as smartwatches, are widely used by the general population and are subject to commercial optimization; thus, their application in health research may benefit from already honed acceptability.

XSL•FO

RenderX

They offer an array of passively collected data types (ie, not requiring user input), including physiological (eg, heart rate), lifestyle (eg, sleep and activity levels), and nearby environmental data (eg, light levels). These may provide more objective data indexing characteristics and symptoms of interest while reducing challenges associated with subjective reporting. Furthermore, their unobtrusiveness may reduce Hawthorne effects (ie, behavioral change due to being observed) [18]. Smartphones also provide the opportunity to collect active measures efficiently and repeatedly in everyday life (eg, collecting mood data via ecological momentary assessment) [19]. Although subjective, these data have high ecological validity and minimize recall effects. They provide context for objective data and enable investigation of fluctuation in subjective symptom levels. At the same time, modern analytical methods, such as machine learning, provide tools for efficiently extracting and identifying meaningful patterns in large datasets. Suitably large datasets are readily produced by wearables, as well as through video, audio, and natural language recording, and can now be processed much more efficiently than via traditional approaches such as manual coding. Together, such technology may enable efficient measurement with relatively low burden to participants and high ecological validity. Such characteristics hold particular promises for clinical research, in which symptom changes would ideally be tracked over time, and early detection of a subtle change may be valuable.

Table 1. Summary of current measurement limitations, opportunities for digital technology to address these, and potential challenges therein.

Domain	Limitation of current measurement	Opportunity for digital technology	Potential challenges
Meaningful capture of real-world outcomes	• Difficulties associated with self-reporting and recall effects • Observations and assessments are often conducted at discrete time points and in unfamiliar situations	• Reporting in the moment via smartphone may alleviate difficulty remembering thoughts, feelings, and behavior over weeks or months • Real-world, real-time data collection offers potential for high ecological validity	• Limits inclusion of certain groups, for example, those without a smartphone and those who find technology less accessible or trustworthy. Daily reporting for long periods may be burdensome. Both these points may disproportionately affect autistic individuals.
Measurement error and bias	• Susceptibility to attention or placebo effects; difficulty blinding informants in psychosocial intervention research	• Objective data capture possible via passive physiological and environmental sensors	• Subjectivity may still be present, for example, in analytic choices and model training. Models relying on neurotypical data may misclassify data from autistic individuals.
Scalability	• Observations and assessments by researchers or clinicians are expensive and time-consuming • Repeated and lengthy questionnaire administration may become burdensome	• Potentially highly scalable across individuals and over time • Passive data collection could reduce the burden of repeated or long-term data collection	• Wide variation in affordability of devices; many devices are inaccessible in lower-resource research settings. • Wearable device comfort and data privacy concerns associated with passive data collection may affect adherence and exclude certain groups, for example, those with sensory sensitivities.
Capture of differences	• Difficult for humans to detect subtle differences across or within people	• Digital sensors can capture fine-grained differences	• Processing and interpretation of large amounts of data is challenging, has an environmental impact, and often lacks validation against well-understood instruments.

For these reasons, digital measurement has drawn interest in diverse physical health applications, such as wearable sensors for monitoring recovery in oncology patients and patients who have had stroke [20,21], monitoring disease progression in Parkinson disease [22], and identifying epileptic seizures noninvasively [23]. Elsewhere, smartphone apps and wearable wristbands have been explored for use in predicting depression relapses [24,25]. In autism research specifically, several digital measurement systems have been developed. These include in-person autism diagnostics or brief conversation-based assessments using machine learning to quantify elements of social communication from videotaped interactions [26-29]. Similar techniques have also been applied remotely, including at-home attention and emotion analysis based on computer vision [30]. In the domain of sleep, objective monitoring via actigraphy has been applied relatively widely [31-33]. Integrating laboratory-based and wearable biosensors with a web-based and mobile app to collect objective and caregiver-report data, the Janssen Autism Knowledge Engine (JAKE) system [34] has been fruitful in exploring existing and novel measures relating to autism and co-occurring symptoms in the domains of sleep [35], prediction of diagnostic groups

from multimodal biosensor data [36], and daily caregiver-reported mood and behavior [37]. Published research using the JAKE system has focused on a sample of caregivers and their autistic or nonautistic children, who are predominantly children or young adults of average intellectual ability. However, to our knowledge, a comprehensive system such as the JAKE system has not been developed to capture self-report and passive smartphone data alongside wearable and in-person assessment among autistic adolescent and adult smartphone users.

Furthermore, despite growing research interest, very little has been published on the acceptability, feasibility, and implementation challenges of digital assessments in autism (although important work on evidence-based practice for developing digital supports for autistic people has been published [38]). For the JAKE system, caregiver feedback was sought and showed encouraging findings in terms of usability and device comfort; however, the experiences of the autistic participants themselves were not directly captured. Autistic people's experiences of research supported by digital technology are likely to differ in several ways from those of other populations, as is the analysis and interpretation of the data yielded. First, sensory sensitivities may affect the comfort of wearable devices. This means that important data may be missing, for example, if people find a wristband uncomfortable to wear at night in a sleep study [39]. On the other hand, wristbands may be more widely tolerated than scalp electrodes such as those used in polysomnography [40] and innovation in the characterization of sleep from data acquired via wristbands (eg, actigraphy and pulse rate monitoring) may enable the inclusion of a wider proportion of the autistic population. Second, higher rates of intellectual disability may affect the usability of smartphone apps, and differences in communication and emotional literacy may affect the way in which objective data correspond to reported subjective experience. Third, analytical tools and pipelines developed in other populations may function differently in the autistic population. Care must be taken to anticipate potential neurotypical biases in the ways that features of speech, social interaction, and facial expressions are classified in machine learning tools trained in other populations.

Study Approaches and Target Domains

This study protocol comprises a dual in-person and remote approach to digital measure development in the domains of autistic characteristics (social communication and restricted or repetitive behaviors and interests) and selected co-occurring difficulties (mood, sleep, inattention, and hyperactivity). Our primary focus is on evaluating the acceptability and feasibility of the methods developed. Co-occurring domains were selected based on being (1) associated with everyday challenges, (2) highlighted by autistic people as important targets for improvement, and (3) putatively amenable to digital measurement. As depicted in Figure 1, we aim to enable the evaluation of a range of candidate data collection methods for their individual and combined potential in indexing these target domains.

Our in-person approach explores digital augmentation to the Autism Diagnostic Observation Schedule-2 (ADOS-2) [41], a widely used in-person observational assessment, a gold-standard instrument in informing autism diagnosis. The ADOS-2 has been used in clinical studies to evaluate intervention effects on social communication [42-44]. Its scoring system is designed to allow human administrators to differentiate reliably between several levels of intensity across a range of behaviors characteristic of autism. Because fine-grained discriminations are challenging for humans, this system may not be ideal for identifying subtle changes, for example, early response to an intervention [45]. Our digital augmentations to the ADOS-2 comprise video recording from 3 angles, audio recording for speech transcription, and a wearable wristband recording heart rate using photoplethysmography, electrodermal activity (EDA), and accelerometer data. These digital augmentations are intended to enable exploration of objective data to complement subjective observer ratings of social communication and repetitive and restricted behaviors and of the underlying autonomic demands of this assessment.

Existing research has shown autonomic correlates of outcomes of interest within the autistic population, such as social function [46], anxiety [47], and depression [48]. Such outcomes are among those that are both key sources of everyday distress and difficult to verbalize for many autistic people. Thus, there is an interest in discovering objective autonomic markers that may underlie states of subjective distress and may be sensitive to treatment response. The ADOS-2 provides a standardized series of presses mimicking everyday interpersonal interactions that present various challenges to autistic people. Therefore, it is well-placed for studying autonomic responses during such interactions.

We will also use a 28-day remote measurement (RM) period by using a commercial wearable device and both passive and active smartphone-based measures. Passive smartphone data collection will use an app that records data continuously from smartphone sensors (eg, motion sensors, GPS location, and Bluetooth connectivity) and logs (eg, app use and call and SMS logs). Such apps have contributed elsewhere to assessing mental health [49,50], and diverse aspects of social interaction and environment [51-53]; however, they have not been widely applied in autism. In this study, both active and passive approaches seek to understand trade-offs between data quality, acceptability to participants, and feasibility of implementation.

For both in-person and remote assessments, we developed our approach in a stepwise manner. First, autism community representatives were consulted to discuss potential opportunities and pitfalls for using candidate technologies in autistic populations, to select specific devices, to maximize procedural acceptability, and to identify potential facilitators and barriers for participation. Protocols were then piloted before implementation as part of the third time point of the AIMS Longitudinal European Autism Project (LEAP) [54,55]. As a large, well-characterized cohort comprising both autistic individuals and comparison individuals, data collection in LEAP enables a detailed understanding of participant characteristics that may affect acceptability and feasibility and how our measures relate to existing gold-standard and widely used conventional measures.

Figure 1. Conceptual depiction of study approaches to exploring individual and combined potential for candidate data collection modes to index variance in domains of interest. Arrows represent potential indexing of domains of interest by data collection modes. For example, sleep disruption (blue arrows) may be indicated by a self-report sleep quality rating and also by several passive data sources. ADHD: attention-deficit/hyperactivity disorder; ADOS-2: Autism Diagnostic Observation Schedule-2; EDA: electrodermal activity; RRBI: restricted or repetitive behaviors and interests.
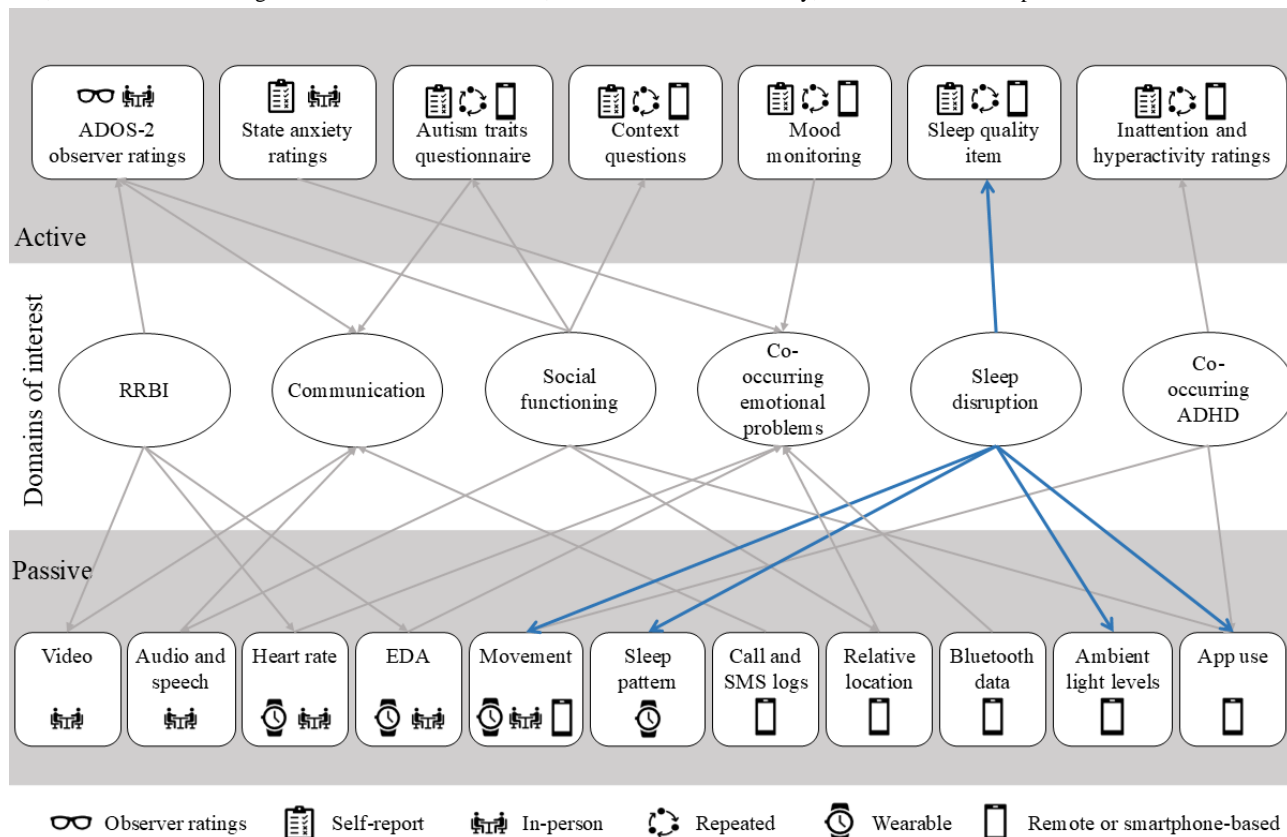
Study Objectives

The primary aim of this study is to evaluate, via quantitative and qualitative methods, the acceptability and feasibility of the measurement systems developed. Primary acceptability outcomes are uptake, adherence, completion rates for each component, and qualitative feedback on participant experience for RM. Primary feasibility outcomes are data availability and proportion of usable data relevant to each target domain. Secondarily, we aim to develop measures from the resulting data relevant to core autism characteristics (eg, social communication) and co-occurring conditions (eg, mental health and sleep problems). Finally, we aim to explore the resulting measures in terms of the association between subjective and objective measures and psychometric properties. In so doing, we hope to identify key candidate measures that show promising characteristics for further validation. This paper will describe the overall design, protocols for in-person and remote assessment, and key target analyses before discussing the potential impact, strengths, and limitations of our approaches.

Methods

Overall Design

Autism Community Involvement

In developing our protocol, we involved community representatives in a series of online focus groups aiming to explore community views on practical and ethical concerns regarding the use of digital technology in outcome measure development. Community views informed key elements of the protocol, such as study information processes, choice of wearable devices, and the interview schedule for qualitative feedback.

Participants

Data collection has been implemented primarily as part of the third (and, at the time of writing, most recent) time point of the LEAP study [54,55]. LEAP constitutes a large, well-characterized cohort recruited from centers across Europe: King's College London (KCL), University of Cambridge (UCAM), Radboud University Medical Centre, Nijmegen (RUMC), University Medical Centre, Utrecht, and Central Institute of Mental Health, Mannheim. At inception, LEAP included 437 autistic participants and 300 nonautistic participants aged between 6 and 30 years who have undergone repeated comprehensive phenotyping at 3 time points (T1: 2014-2017, T2: 2015-2017, and T3: 2020-2024). This includes gold-standard autism clinical profiling (ie, ADOS-2, Autism Diagnostic Interview-Revised, and Vineland Adaptive Behavior Scales); neurocognitive assessments; co-occurring symptom assessments; brain imaging; biochemical markers; and genomics. Participants with an IQ in the typical range were recruited in a 3:1 male-to-female ratio within age bands defined as (A) adult (aged between 18 and 30 y), (B) adolescent (aged between 12 and 17 y), and (C) child (aged between 6 and 11 y). Participants aged between 12 and 30 years with mild intellectual disability

(IQ 50-74) were additionally recruited to a fourth schedule (D). The eligibility criteria for autistic participants were an existing clinical diagnosis according to the *DSM-IV* (*Diagnostic and Statistical Manual of Mental Disorders* [Fourth Edition]); *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]); or *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) criteria.

Preliminary recruitment figures for LEAP T3 indicate that 225 autistic participants and 132 nonautistic participants took part across sites. A sample of 36 autistic participants and 44 nonautistic participants but with epilepsy (referred to as the LEAP epilepsy sample) was additionally recruited at KCL to complete the LEAP T3 protocol, including digital measures. All autistic participants in schedules A and B (ie, those aged ≥12 years) who were contactable at T3 at the KCL, UCAM, and RUMC sites (including University Medical Centre, Utrecht participants, who were seen at RUMC) were eligible to take part in our digitally augmented ADOS-2. In total, 30 nonautistic participants from across sites and schedules were also to be recruited. For RM, additional eligibility criteria were enrollment at KCL or UCAM (non-UK sites were not included because app translation was beyond our scope), independent use of an up-to-date Android (Google LLC) smartphone (version 8 or later), and willingness to try wearing the Fitbit Inspire 2 device as much as possible (day and night) for 28 days. There were no additional exclusion criteria to not meeting the inclusion criteria. We planned to only include Android users because our passive smartphone data collection app is only compatible with Android smartphones. Although we considered providing Android study phones to participants, we reasoned that these additional smartphones were unlikely to be used in a habitual manner and would thereby compromise ecological validity. Participants indicating interest but also a concern that they would find the Fitbit uncomfortable were invited to try to participate. It was emphasized that they would be free to discontinue wearing the Fitbit at any time or work with researchers to find a more comfortable strap. Finally, participants in schedule D were also eligible for each component, according to the same criteria, where it was felt that the protocol would be manageable within the context of overall demands presented by the wider LEAP T3 protocol and adequately understood for the purpose of providing informed consent. Participants in LEAP epilepsy were also invited to take part in both components according to the same criteria.

In addition to data collection in the LEAP sample, we recruited a volunteer sample of 25 nonautistic individuals or without known neurodevelopmental, neurological, mental health, or sleep conditions (RM comparison sample) to participate in an extended version of the RM component only. While data collection in LEAP and LEAP epilepsy is complete, data collection in our RM comparison sample is ongoing and will be completed in August 2025.

RADAR-Base Data Collection Platform

Both in-person and remote assessments use the RADAR-base platform for digital data collection and management [56]. The platform comprises the collection of multiple data sources, for example, passive and active data collection smartphone apps, as well as application programming interface integration for ingesting third-party (eg, Fitbit) data, secure data storage systems, and a management portal for linking data sources to participant identifying codes. The active RM technology (aRMT) app is a questionnaire administration app that allows a customizable schedule of questionnaires to be administered. The use of RADAR-base enables the secure collection and management of automatically time-stamped active and passive data for downstream analysis. The passive RM technology (pRMT) app [23,50] collects smartphone sensor data (eg, movement, relative location, light levels, and app use) as well as data from integrated wearable devices (eg, the Empatica E4 device used in this study).

In-Person Assessment

Procedure

The digitally augmented ADOS-2 is implemented in LEAP T3 as part of a 2-day in-person assessment visit. Cameras and a lapel microphone are set up before the participant enters the room (as depicted in Figure S2 in Multimedia Appendix 1 [27,28,49,57-75]). At the beginning of the assessment, participants are introduced to the Empatica E4 wristband and shown how to put it on their nondominant wrist. Video recording is then started on all cameras, and ceiling lights are switched off and on again as a synchronizing marker. The E4 is then turned on, and connection with the RADAR passive app is automatically established. Before starting the ADOS-2 assessment, participants complete a pencil-and-paper measure of state anxiety (detailed in subsequent sections). The ADOS-2 assessment is then administered according to standard protocol. Data collection started in September 2021, while certain social distancing and enhanced hygiene policies were still in place. We accommodated these via several adaptations to usual ADOS-2 procedures (summarized in Table S1 in Multimedia Appendix 1). The exact implementation for each "ADOS-2 informed assessment" was allowed to differ according to current site-specific policies and guidelines and was documented in case report files. For example, participants and administrators sat at a distance (eg, 2 m), clear masks were used, and certain ADOS-2 materials were adapted or removed. Between each ADOS-2 activity, the participants are asked to wave the wrist on which they were wearing the E4 to create an event marker. At the end of the assessment, recording is stopped, the E4 wristband is removed, and participants fill in the state anxiety measure again. Participants are also invited to wear the E4 wristband during a separate electroencephalography and eye-tracking task battery that forms part of the LEAP T3 visit. This acts as a comparison condition with lower social interaction demands.

Measures

Empatica E4

The Empatica E4 is a wrist-worn device that captures motion via accelerometer, heart rate via photoplethysmography, and EDA activity via electrodes on the device or via finger attachments (Figure S1 in Multimedia Appendix 1). Previous research has shown poor accuracy for wrist-worn EDA capture [57], and our own piloting demonstrates better capture using

finger electrodes. Therefore, we use these unless participants find them uncomfortable, in which case we use the on-device electrodes. This option is provided primarily to maximize collection of photoplethysmography and motion data from the E4 and to explore acceptability of EDA electrode options in this context. This is because EDA readings differ across sites [76], and our protocol is not optimized for wrist placement (to date, all participants willing to wear the E4 were happy with finger placement). E4 data are ingested via the RADAR pRMT app installed on a study-owned compatible Android smartphone. The RADAR pRMT app is configured to pair automatically with the device via Bluetooth to ingest data. The Empatica E4 has been shown elsewhere to be well tolerated in naturalistic observation sessions involving autistic youth and to produce usable data for prediction of overt behavior from physiological data [77].

Video and Audio Recording

Video data are captured using commercially available high-definition video cameras, exporting at a minimum of 1920×1080-pixel resolution and 24 frames/second sampling rate, from three angles: (1) recording participant face from researcher perspective, (2) recording researcher face from participant perspective, and (3) recording the scene from observer perspective (camera setup is shown in Figure S2 in Multimedia Appendix 1). Cameras are mounted on tripods to minimize movement. The exact room setup varies across sites and sometimes within sites due to the use of different rooms and the constraints imposed by furniture size and layout. It is also not possible to standardize room acoustics and ambient lighting due to windows and differences in electrical lighting across rooms. In addition, the researcher wears a mobile lapel microphone to aid in distinguishing speakers (ie, researcher speech from participant speech).

Researcher Observational Ratings

The ADOS-2 [41] is a widely used gold-standard diagnostic instrument for autism in which a trained administrator leads a series of activities designed to elicit social interaction. The administrator assigns scores on numerous ordinal scales within domains, such as language and communication, reciprocal social interaction, and stereotyped behaviors and restricted interests. In this study, as per the LEAP protocol, we used modules 3 and 4 activity schedules, which were designed for individuals who are verbally fluent.

State Anxiety

The 6-item short form of the Spielberger State-Trait Anxiety Inventory state scale [78] comprises the 6 items exhibiting the highest item-remainder correlations in the full scale. The short form displays excellent correlation with full-scale scores ($r \geq 0.90$), and acceptable reliability and validity. Participants rate items (eg, I feel calm) on a 4-point scale between 1 (not at all) and 4 (very much) according to how they feel "right now."

Remote Measurement

The RM System

Overview

The RM system comprises 4 components encompassing both active and passive data collection for a period of 28 days. This period was chosen pragmatically, considering the trade-off between gathering sufficient data to enable target analyses and acceptability to participants. Active data collection is via the RADAR aRMT app and MyJournE (Digital Marmalade) mood monitoring app. Passive data collection is via the RADAR pRMT app, Fitbit Inspire 2 wristband, and Fitbit app. Both active and passive data acquisition components are described in detail in subsequent sections. Figure 1 shows how multimodal data from the RM system map onto the domains of interest for measure development.

Active Components

We will use the RADAR aRMT primarily to administer our questionnaire schedule with prompts sent by push notification. The schedule includes a daily sleep quality item, biweekly attention items, and fortnightly social communication items (instruments detailed in subsequent sections). We will also use the aRMT to provide information about weekly progress, informing participants when each of the 4 weeks has been completed. Finally, at the end of the 28 days, system usability and user experience items will be administered to quantitatively capture participant satisfaction with the system. The schedule for the active components is summarized in Table S2 in Multimedia Appendix 1, and the instruments used are detailed in subsequent sections.

MyJournE is a mood monitoring app codesigned by young people [79] and researchers at KCL and developed by a UK-based software development company Digital Marmalade [80]. The app sends reminders via push notification twice per day for participants to complete the mood logs. These include a series of questions to efficiently capture context (eg, who a participant is with, where they are, and what they are doing), followed by a general happiness rating on a 5-point scale from very unhappy to very happy. This is followed by an implementation of the Positive and Negative Affect Scale (PANAS; refer to subsequent sections) in which each item is accompanied by a visualization of the emotion word involving a purple bean character (Figure S3 in Multimedia Appendix 1). Participants respond using a movable visual slider. Visual supports are commonly used by autistic people to aid in processing abstract concepts [81] and thus may contribute to acceptability, especially in the context of high rates of alexithymia among autistic individuals.

Passive Components

Fitbit Inspire 2 is a commercial fitness tracking device that records heart rate via photoplethysmography and motion via an accelerometer. It is lightweight, unobtrusive, and allows for a choice of strap material; thus, it is well-placed to cater for some degree of sensory sensitivity. Using proprietary algorithms, Fitbit provides heart rate, step count, and sleep staging data. Fitbit models manufactured after 2017 estimate time spent in 4 sleep stages (awake, light sleep, deep sleep, and rapid eye

movement) using machine learning methods applied to motion and heart rate [82]. From this, sleep parameters such as total sleep time and wake after sleep onset may be derived [49].

The RADAR pRMT collects a variety of Android smartphone sensor data without the need for regular participant input. Specifically, these data include physical motion (eg, from the onboard accelerometer and gyroscope), ambient light levels, GPS location, number of Bluetooth-enabled devices in range, phone and app use data, and battery levels. GPS coordinates are offset by a random number for each participant to obscure the absolute location while preserving the relative location (ie, information about distance, duration, and direction of travel). In addition, for this project, we used a custom build of the app [83] that included functionality to access call and SMS logs including data on incoming and outgoing call duration, SMS length, and whether they were with a saved contact. These data are potentially relevant in capturing aspects of social functioning and, due to Google Play Store restrictions, are not included as standard in the app. No call or SMS content is recorded, and participants are able to freely grant or remove permissions for different data sources in the app settings.

Procedure

Given the novelty, complexity, and privacy considerations of RM data collection, our community representatives stressed the importance of a comprehensive information process with a clear description of which information is collected, how it is managed, and how it will be used. With this in mind, we developed an extended information and consent process. Eligible participants will be sent an email with attachments comprising the participant information sheet and a participant handbook (Multimedia Appendix 2). The participant handbook contains additional information on the types of data collected, particularly those that may feel more personal (eg, location and app use data), and how the data are pseudonymized, stored, and processed. Participants are also invited to a study information video call, which is typically scheduled at least 1 week after the email to allow time for reading the information documents. The call takes approximately 30 minutes, aims to run through the study information at a suitable pace, and allows time for questions. At the end of the call, participants who are happy to proceed are provided with a link to a web-based consent form, which they may complete as part of the call or in their own time. Consenting participants are invited to an enrollment video call approximately 1 week after the information call to allow for a Fitbit to be posted to their home.

The enrollment call takes approximately 1 hour and comprises app installation, setup, and Fitbit connection. The MyJournE app, RADAR active app, and the Fitbit app are installed from the Google Play Store whereas the custom RADAR passive app is installed from GitHub to the Android Application Package. Participants are given log-in details for each app and given a chance to explore the active data collection apps. The RADAR passive app requests access to several smartphone data sources. Participants are taken through these and reminded that they can control the data the app accesses by toggling permissions and are free to disallow access to any data source at any time if they

feel uncomfortable. Finally, participants are taken through setting up their Fitbit device and pairing it with the app.

The 28-day data collection period starts the day after the enrollment call. Participants receive reminders to respond to the active components. A day 3 check-in is administered via video or phone call or email to check if the participant is happy with how the month is proceeding and to troubleshoot any issues. Researchers complete regular checks by accessing the RADAR-base platform to ensure data are coming in for each component (ie, Fitbit, pRMT app, aRMT app, and MyJournE). We contact the participant or relevant technical support teams for troubleshooting if any data component is not received or largely missing. For example, if data checks show that Fitbit data are missing, the participant is contacted to establish if the device is being worn or if there are syncing or charging issues. If the participants report that they are finding the device uncomfortable, an alternative strap is offered if it is felt that this may help. If data are missing from the smartphone apps, similar troubleshooting occurs, for example, to establish if notifications are being received or if the participant is simply finding it difficult to respond.

At the end of the 28 days, participants are notified to complete the user experience questionnaires via the RADAR aRMT. These questionnaires are available for 1 week, after which data are no longer collected via the aRMT app. Participants are also invited to complete a feedback interview during the week following completion. The feedback interview is conducted via a video call and lasts around 1 hour. It also includes questions eliciting participant opinions on the study information and enrollment processes, general experience and impact on daily life (including discussion of potential Hawthorne effects), Fitbit usability and comfort, active app scheduling, and passive data collection acceptability. Participants keep the Fitbit device at the end of the month and receive £20 (US $26.81) in Amazon vouchers for their time spent taking part in the feedback interview.

Measures

Autism Ratings

The Social Responsiveness Scale Short Form [84] is a questionnaire measure of autistic characteristics, particularly, in the social communication domain. It comprises 16 items derived from the full parent report version (65 items). The short form showed excellent reliability (α=.96) and substantial associations with social features of autism measured by gold-standard instruments (eg, Autism Diagnostic Interview-Revised Social subscale, r=0.48, and ADOS-2 module 4 Social Affect, r=0.49) [84]. In this study, it is administered as a self-report questionnaire on days 8 and 15 of the 28-day RM period.

Inattention and Hyperactivity Ratings

On the basis of a previous experience sampling study [85], we used a 4-item scale comprising 3 inattention items (eg, "I was easily distracted and could not concentrate over a longer period of time") and 1 hyperactivity item ("I felt restless or fidgety"), based on *DSM-5* ADHD criteria. In a community sample of adolescents aged between 14 and 17 years, internal consistency

for the inattention items was satisfactory (α=.67-.79) [85]. In this study, items are rated every 4 days on a scale from 0 (not at all) to 100 (all the time) using a movable slider.

Sleep Quality Ratings

The Single-item Sleep Quality Scale [86] is a simple measure of overall sleep quality rated on a ten-point scale. It shows a strong association with widely used questionnaire measures and was able to differentiate normal, borderline, and problem sleepers in a study of patients with depression [86]. In this study, participants are asked to rate their sleep quality for the night before each morning on a scale from 0 (terrible) to 100 (Excellent) using a moveable slider. The 0-100 scale was used for consistency with the inattention and hyperactivity items.

Mood Ratings

The PANAS is a widely used 20-item measure of positive and negative affect in which participants rate the intensity with which they are feeling each of 10 positive and 10 negative mood descriptors (eg, "enthusiastic" and "distressed"). Internal consistency for both scales was shown to be excellent (α=.84-.90) [87]. In the MyJournE app, items are rated according to the question "How appropriate is this word today?" using a 5-point movable slider scale ranging from "very slightly/not at all" to "extremely."

User Experience Ratings

The System Usability Scale [88] is a 10-item usability questionnaire in which statements (eg, "I thought the system was easy to use") are rated on a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree"). The scale has been demonstrated to have excellent internal consistency (α=.91) [89]. In this study, participants are notified via the RADAR aRMT app to complete the System Usability Scale based on their experience of the RM system *overall* (Fitbit and smartphone apps) on day 29 after enrollment. At the same time, participants were also notified to complete a separate user experience questionnaire with 7 items specific to the Fitbit and 10 items specific to the active questionnaire apps [90].

Ethical Considerations

Ethics approval was obtained for LEAP and LEAP epilepsy groups at KCL and UCAM from the London-Central and Queen Square Health Research Authority Research Ethics Committee (13/LO/1156) and at RUMC from the RUMC's Medisch-Ethische Toetsingscommissie (2019-5942; NL72033.091.19). For the RM comparison group, KCL Research Ethics was obtained (LRS/DP-23/24-42359). Compliance with European and UK data protection policies was ensured by performing a data protection impact assessment in consultation with the KCL information compliance department and compliance teams at local sites. RM data are pseudonymized via a study identification code, and dummy Fitbit accounts are set up for participants, which are not linked to their personal information. Data acquired from the Fitbit and smartphone apps are automatically encrypted and uploaded to secure servers as part of the RADAR platform. Video and other identifiable data are kept separately on secure servers and shared across study sites for analysis purposes with explicit consent.

Planned Analyses

Overview

This protocol is primarily intended to support evaluation of the acceptability and feasibility of the range of candidate technologies and methods that it uses. Secondarily, it may provide rich, multimodal data supporting the exploration of a range of potentially meaningful outcome measures. The potential avenues for such exploration are numerous and dependent not only on the quality of the resulting data but also on available data processing tools and analytic approaches, which are expected to evolve rapidly. Therefore, in this paper we describe our primary analyses of acceptability and feasibility in detail and give only several examples of candidate outcome measure exploration planned in this study. The specifics of these secondary analyses are beyond the scope of this paper but will be presented in individual outputs. Our featured initial domains for outcome measure development are based on community priorities (eg, sleep and mental health) [91,92], tractability based on currently available analytic tools, and innovation in terms of building on existing work. Initial data processing plans are provided in Table S3 of Multimedia Appendix 1.

Acceptability and Feasibility

Overview

Acceptability and feasibility will be evaluated comprehensively via quantitative and qualitative analyses (Table 2). Uptake and completion rates will be calculated at the level of each system (ie, augmented ADOS-2 and RM protocols), whereas quantification of adherence, tolerance, and participant experience will differ according to subcomponents; for example, adherence to wearing the Fitbit is examined separately to adherence to questionnaire completion. Qualitative data, for example, reasons for declining to participate and for withdrawal will contribute to our understanding of quantitative information and, in the case of our RM protocol, will support a larger qualitative framework analysis of feedback interviews. Similarly, feasibility will be evaluated with reference to participant eligibility, data availability, and data quality and may be specific to each analytic approach. Limiting factors for the proportion of usable data will be identified; for example, establishing whether active RM data are missing due to technical failure (eg, notifications not sending) or nonresponse.

Table 2. Planned metrics for acceptability and feasibility.

	In-person components (E4 and video recording)	Remote measurement components (Fitbit, pRMT[a], aRMT[b], and MJE[c])
Acceptability		
Uptake	Proportion consenting and reasons for decline	Proportion consenting and reasons for decline
Adherence and tolerance	E4 wear time and reasons for removal and requests for discontinuation of filming	Fitbit wear time and reasons for removal, pRMT app permissions granted, and aRMT and MJE response rates
Experience	STAI State-6[d] (before and after assessment)	Usability ratings (SUS[e] and UEQ[f]) and feedback interview (qualitative)
Completion	Proportion completing protocol and reasons for withdrawal	Proportion completing protocol and reasons for withdrawal
Feasibility		
Eligibility	Proportion eligible and reasons for ineligibility	Proportion eligible and reasons for ineligibility
Data availability	Missing data and reasons for missingness (eg, equipment malfunction)	Missing data; reasons for missingness (eg, hardware and software malfunction versus acceptability issue)
Data quality	Proportion of usable data for each analysis and reasons for low-quality data	Proportion of usable data for each analysis and reasons for low-quality data

[a]pRMT: passive remote measurement technology.

[b]aRMT active remote measurement technology.

[c]MJE: MyJournE app.

[d]STAI State-6: 6-item short form of the Spielberger State-Trait Anxiety Inventory.

[e]SUS: System Usability Scale.

[f]UEQ: User Experience Questionnaire.

Predictors of Acceptability and Feasibility

Predictors of certain acceptability and feasibility metrics will be explored. Primarily, differences between autistic and nonautistic groups will be examined for each metric with formal parametric and nonparametric statistical methods applied where possible. Age, cognitive ability, and adaptive ability will be assessed as predictors of uptake, adherence (wear time and active app response rates), usability ratings, and completion. We will also explore the association between Fitbit wear time and the tactile sensitivity subscale of the Short Sensory Profile [93] completed by participants as part of the main LEAP study.

Planned Exploration of Candidate Outcome Measures

Digitally Augmented ADOS-2 Data

Audiovisual data from ADOS-2 will be processed using machine learning–based tools, assisted by researcher annotation of key behaviors. We will explore associations between resulting features extracted from the audiovisual data and observer-assigned scores in key domains. In so doing, we will explore the performance of existing speech and video processing tools and the suitability of extracted features for quantifying clinically relevant characteristics via more objective, precise, automatically generated scores. Note that it is not our intention to use these data to explore diagnostic classification approaches as has been done elsewhere. Our focus is on developing tools to quantify characteristics that may be used to understand heterogeneity within clinical classification and to evaluate interventions. For example, drawing on previously published

pipelines [58], our speech data will be processed using open-source, automated diarization, transcription, and conversation analysis tools such as pyannoteAI Audio [59], OpenAI Whisper [60], and ConvoKit [61], respectively (Table S3 in Multimedia Appendix 1). Resulting features (eg, number of conversational turns and the length of participant and researcher utterances) will be tested for association with ADOS-2 language and communication codes indicating the amount and nature of social speech and reciprocal social interaction codes for the frequency and quality of social initiation and response. Such associations will be tested using conventional regression-based statistical analyses.

In addition, we will explore the ability of machine learning and computer vision tools to identify instances of verbal and nonverbal behaviors relevant to ADOS-2 scoring algorithms. For example, drawing on existing work identifying and classifying stereotyped movements from ADOS-2 data in children [62], we will explore the feasibility of applying such methods to our adolescent and adult data, as well as assessing the utility of integrating E4 motion and psychophysiological data to refine classification.

Using Empatica E4 photoplethysmography and EDA data, we plan to develop pipelines for investigating autonomic arousal in the context of the ADOS-2 assessment. Data will undergo quality control procedures (eg, [63,94]) and processing (Table S3 in Multimedia Appendix 1). Tasks will be demarcated using event markers generated via hand waves. Key tasks of interest within the ADOS-2 will be identified, that is, those hypothesized to have high (presenting a story from a cartoon and discussing

emotions) versus low (completing a puzzle and describing a picture) psychosocial demand for autistic people. Within each relevant task, heart rate, heart rate variability, and EDA features will be calculated, applying and testing existing pipelines [95,96]. Derived indices of autonomic arousal will then be compared across conditions within participants. If data are of sufficient quality and power, linear mixed effects models (LMEMs) will investigate group differences in autonomic response to psychosocial demand [47]. Association of autonomic responses with subjective state anxiety scores from the beginning and end of the assessment will also be explored.

RM Data

Overview

Using RM data, we will focus on exploring objective, multimodal indices from passive data in several key domains (sleep, mental health, and social functioning) and investigate their associations with corresponding self-report data. Due to the relatively short observation period and lack of planned intervention during that time, this exploration will focus on between-subject analyses as opposed to within-subject change over time. Given the hierarchical data structures produced by repeated measurement, associations will be investigated using LMEMs. These account for dependent data points flexibly, with minimal loss of power [97].

Missing Data

Given the challenges associated with collecting data remotely, we anticipate missing data due to varying participant engagement levels, device issues, or network connectivity. We will draw on published research establishing feature-specific guidance on acceptable levels of missing data to determine data eligibility for further analysis [98]. Where possible, we will use multiple imputation to account for missing values (eg, where certain modalities have some missing values for a participant in a multimodal analysis). Coupled with this, LMEMs allow for retention of cases with missing observations and are compatible with certain maximum likelihood–based estimation procedures [99]. Multiple imputation– and maximum likelihood–based approaches are considered unbiased where data are missing at random [100]. The plausibility of this assumption will be assessed on a case-by-case basis and be addressed where necessary, for example, using sensitivity analyses.

Sleep

As well as drawing on Fitbit-provided sleep staging data, we will incorporate passive smartphone data, such as light levels, phone motion, and app use, to derive objective indices of sleep architecture. This novel approach is intended to improve accuracy in differentiating periods of sleep versus wakefulness. In particular, we will target sleep parameters identified commonly as disrupted in autistic youth and adults, for example, sleep onset latency, wake after sleep onset, sleep efficiency, and percentage of rapid eye movement sleep [33,101]. We will test the associations between these objective parameters and daily subjective sleep quality ratings using LMEMs. In addition, we will use cluster analysis to identify more homogeneous, objectively recorded sleep patterns.

Mental Health

Drawing on existing research into smartphone sensor-based indicators of mental health, for example, from location and activity data, phone and app use patterns, and call and SMS logs [102], we plan to test currently available tools and procedures for feature extraction and assessment of correspondence between passive data and subjective positive and negative affect as reported in the PANAS, administered via the MyJournE app (eg, [103]). We will also explore associations between extracted features and mental health measures collected as part of the LEAP questionnaire battery (ie, the Beck Depression Inventory [104] and Beck Anxiety Inventory [105]) and reported diagnoses.

Social Functioning

Finally, we plan to use similar methods to explore aspects of everyday social functioning and environments that may be indexed in our passive RM data. Our data offer an opportunity to explore objective indices of social functioning and environment in everyday life in relation to subjective reports of social communication, social activities, and mood. We will adapt pipelines developed in other populations to index various aspects of social functioning, including social activity [106], "sociability" (comprising communication and social exploration domains) [51], and loneliness [52], from multimodal data (including duration, frequency, and diversity of telephone calls, Bluetooth connectivity, GPS location, and smartphone screen interactions). We aim to develop similar multimodal indices that may be referenced against subjective measures, such as the Vineland Adaptive Behavior Scales, Social Responsiveness Scale Short Form, and the MyJournE context questions (eg, "Who are you with?").

Results

Following ethics approvals documented earlier, the first participant was enrolled for the digitally augmented ADOS-2 component on September 11, 2021, and for the RM component on March 3, 2022. Data collection within the LEAP and LEAP epilepsy samples is complete as of October 2024. In the digitally augmented ADOS-2 component, 190 participants took part (LEAP: n=167, 87.9% and LEAP epilepsy: n=23, 12.1%). In the RM component, 86 participants have been enrolled (LEAP: n=47, 55%, LEAP epilepsy: n=14, 16%, and RM comparison: n=25, 29%). In the RM comparison group, data will be collected in an ongoing manner until August 2025.

Several planned analyses are underway. In our primary outcome domains of acceptability and feasibility, qualitative framework analysis coproduced with community members is in its final stages and will contribute to a key planned mixed methods publication on the acceptability and feasibility of RM. This publication will also outline work with autism community members during initial stages and the impact on the protocol. Quantitative acceptability and feasibility analyses contributing to this publication are also underway; for example, we have explored adherence to Fitbit wearing and the active questionnaire schedule. Elsewhere, we have developed working pipelines for analysis of ADOS-2 speech data and RM sleep data.

Discussion

Study Contributions

Autistic people face diverse and complex challenges in daily lives that are difficult to quantify using conventional techniques. Measuring treatment response in a scalable, precise, and unbiased manner is particularly problematic, and a lack of tools available to do this has impeded progress in evaluating autism-specific interventions for such challenges. Development of better measurement tools for quality of life, mental health, social well-being, and sleep was among the top 5 priorities identified by autistic adults and other stakeholders in the study by Benevides et al [91]. Digital technology offers an array of compelling opportunities to measure more objectively, with lower burden and greater ecological validity. This study protocol was designed with input from autism community representatives to develop and evaluate acceptable digital measures relevant to priority outcomes. Our dual in-person and remote approach enables us to develop an understanding of how such measures may relate to and potentially complement each other in the context of a range of study designs. If demonstrated to be adequately acceptable and feasible, the measures explored in this study could be used to enhance understanding of everyday challenges for autistic people. They could support clinical research by providing outcome measures that are specific and responsive to small changes in meaningful outcomes.

Most existing studies applying digital technology to challenges in the diagnosis and management of autism are aimed at proof-of-concept and have relatively small sample sizes [107]. Understanding generated from this study of digital acceptability and feasibility could facilitate the application of digital measures at scale and assist implementation of digital supports for autistic people, for example, app-based mental health interventions [108] and wearable devices that may help autistic people to understand and communicate their emotions [109] or predict maladaptive behaviors [77].

Potential Strengths and Limitations

Data collection in the LEAP study is well positioned for achieving these aims. The sample is already well characterized, meaning that factors potentially associated with acceptability and feasibility may be thoroughly investigated. LEAP includes high-quality measures of autism clinical characteristics, intellectual ability, and co-occurring symptoms, factors that may influence participant uptake and experience of digital measures. Certain candidate measures may also be validated against robust conventional measures in domains of interest, for example, gold-standard autism assessments and well-validated questionnaire instruments. For other novel measures, we do not incorporate a reference standard. In the domain of sleep, polysomnography is not included because its feasibility for implementation outside of clinical settings is problematic and its representativeness in capturing everyday sleep is unclear. However, we do incorporate a daily subjective sleep quality measure, which is an important clinical end point.

For the RM component, instruments were selected to build a system that maximizes acceptability and accessibility (within the context of independent smartphone users). Accordingly, we selected very brief active measures to minimize participant daily time burden and interference with everyday life. We also selected Fitbit Inspire 2 as our wearable device of choice, given its widespread use in the broader population, unobtrusive design, and affordability. However, this approach comes with challenges in terms of data quality and analytic potential. Both wearable devices use photoplethysmography to estimate heart rate, rather than obtaining direct measures via electrocardiography. Use of Fitbit also necessitates relying on proprietary algorithms for heart rate, step count, and sleep pattern, which limits flexibility in analysis. On the other hand, Fitbit provides ready-to-use processed data with established performance against gold-standard measures [82,110,111]. Across studies, recent-generation Fitbit models have been shown to have high sensitivity (0.95-0.96) in identifying sleep epochs, lower specificity (0.58-0.69), and a tendency to underestimate sleep onset latency [82]. Despite efforts to maximize the acceptability of remote data collection, some level of noncompliance is expected, and understanding factors associated with these forms a primary area of investigation for this study.

Another limitation is that the focus on independent smartphone users is likely to exclude many participants with intellectual disabilities, a group forming a considerable proportion of autistic people [112]. Our hope is that the current research may identify promising candidate measures that could be adapted for populations with ID. Indeed, enhancing our understanding of how objective, passive measures relate to subjective experience may pave the way for the development of novel measures of internal states, such as anxiety, which do not require explicit self-report [113]. Similarly, our inclusion criterion of "willingness to try to wear the Fitbit Inspire 2 as much as possible for 28 days" may exclude participants with particular sensitivity to touch. However, our selection of the Fitbit Inspire 2 was largely based on discussion with autistic community members who advised us that selecting a lightweight device with a choice of strap material may minimize exclusion of those with sensory sensitivities and support adherence. Furthermore, our explicit inclusion of questions on device comfort in the feedback interview is intended to advance understanding of the factors that may influence adherence to wearing such a device for autistic participants in research studies.

More generally, our sample is primarily derived from LEAP, a volunteer sample that has been further selected based on eligibility criteria mentioned earlier. This limits the generalizability of the current findings to the wider, highly heterogeneous autistic population. However, it is hoped that responding to acceptability and feasibility challenges highlighted by this study will inform future research that may be more widely applicable. In turn, highly precise, objective, digital end points may ultimately contribute to identifying subgroups of the autistic population that may reflect different mechanisms underpinning challenges and different responses to intervention.

Data privacy is another important issue raised in our community involvement activities and those of others [114]. Digital technology is powerful in the sense that large quantities of data capturing potentially sensitive features are generated with low effort and awareness. Research has shown that most owners of smart devices are willing or would consider sharing their digital

data for health research purposes [115]. However, for certain others, this constitutes an uncomfortable invasion of privacy. In the general population, acceptability may depend on how data are used (eg, for the good of the public), managed (eg, using anonymization procedures), and with whom they are shared [116]. This study has the potential to provide insight into the circumstances under which such data collection is acceptable for autistic people and how to communicate effectively about data use and sharing.

Finally, although this is an observational study, our RM procedures may cause people to change aspects of daily living, introducing Hawthorne effects. For example, use of the Fitbit for its intended purpose as a fitness tracker may lead to changes in activity levels and sleep habits, and knowledge that app use is recorded may cause a participant to change app use habits. Because a putative advantage of digital measurement (particularly, passive and objective measures) is a reduction in potential observation or attention effects in the context of clinical trials, it is important to investigate these during measure development. Therefore, we ask specifically about perceived behavioral changes in our qualitative interview. However, quantitative investigation of within-subject variance over time is out of scope in our study due to the relatively short observation period. Future studies may wish to draw on platforms providing the capability to blind participants from on-device information and app-based dashboards showing summarized data (eg, Oura Ring [117]).

Conclusions

Despite the range of potential challenges in digital measure development, particularly in the field of autism research, identification and understanding of these are incorporated as key aims of this study. Our collection of both quantitative and qualitative data on acceptability will allow a more thorough understanding of factors influencing uptake, adherence, completion, and, importantly, experience of the protocol. Careful consideration of the acceptability and feasibility of candidate measurement techniques, alongside extracting and validating meaningful indices of outcomes important to community stakeholders, is essential in developing measures that live up to the potential afforded by digital innovation.

Authors' Contributions

IY, CC, EL, JB, DM, RD, and ES were responsible for the conceptualization. IY, CB, ARC, PC, HS, YR, and ZR were responsible for data curation. ARC, NC, AF, and RD were responsible for the formal analysis. EL, JB, DM, RD, and ES were responsible for the funding acquisition. IY, CB, and ARC were responsible for the investigation. IY, ARC, BO, PC, HS, YR, ZR, J Dineley, J Downs, CC, NC, AF, RD, and ES were responsible for the methodology. J Downs, EL, JB, DM, RD, and ES were responsible for resources. IY, CB, ARC, BO, AF, EL, JB, DM, RD, and ES were responsible for the project administration. ARC, PC, HS, YR, ZR, J Downs, AF, and RD were responsible for software. J Downs, NC, AF, RD, and ES were responsible for supervision. IY was responsible for writing the original draft. All the authors were responsible for writing—reviewing and editing.

Conflicts of Interest

RD is the director of CogStack Ltd and Onsentia Ltd. CC is a full-time employee of Shionogi, Inc. AF has shares in Google, which acquired Fitbit. All other authors declare no conflicts of interest.

Multimedia Appendix 1

Additional tables and figures comprising Autism Diagnostic Observation Schedule-2 (ADOS-2) pandemic–related adaptations, schedule of events for the 28-day Mobile Measures Month, Empatica E4 device set-up, videocamera and lapel microphone set-up for the ADOS-2, screenshots from the MyJournE app, and additional detail on proposed data processing for passive data collection modalities.

[DOCX File , 789 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Participant handbook for the remote measurement component.

[PDF File (Adobe PDF File), 1368 KB-Multimedia Appendix 2]

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). 5th edition. Arlington, VA. American Psychiatric Publishing; 2013.
2. Lord C, Charman T, Havdahl A, Carbone P, Anagnostou E, Boyd B, et al. The Lancet Commission on the future of care and clinical research in autism. Lancet. Jan 15, 2022;399(10321):271-334. [doi: 10.1016/S0140-6736(21)01541-5] [Medline: 34883054]
3. Howlin P, Magiati I. Autism spectrum disorder: outcomes in adulthood. Curr Opin Psychiatry. Mar 2017;30(2):69-76. [doi: 10.1097/YCO.0000000000000308] [Medline: 28067726]
4. Roche L, Adams D, Clark M. Research priorities of the autism community: a systematic review of key stakeholder perspectives. Autism. Feb 2021;25(2):336-348. [doi: 10.1177/1362361320967790] [Medline: 33143455]
5. McConachie H, Parr JR, Glod M, Hanratty J, Livingstone N, Oono IP, et al. Systematic review of tools to measure outcomes for young children with autism spectrum disorder. Health Technol Assess. Jun 2015;19(41):1-506. [FREE Full text] [doi: 10.3310/hta19410] [Medline: 26065374]
6. Linden A, Best L, Elise F, Roberts D, Branagan A, Tay YB, et al. Benefits and harms of interventions to improve anxiety, depression, and other mental health outcomes for autistic people: A systematic review and network meta-analysis of randomised controlled trials. Autism. Jan 2023;27(1):7-30. [FREE Full text] [doi: 10.1177/13623613221117931] [Medline: 35957523]
7. McCracken JT, Anagnostou E, Arango C, Dawson G, Farchione T, Mantua V, et al. ISCTM/ECNP ASD Working Group. Drug development for Autism Spectrum Disorder (ASD): progress, challenges, and future directions. Eur Neuropsychopharmacol. Jul 2021;48:3-31. [FREE Full text] [doi: 10.1016/j.euroneuro.2021.05.010] [Medline: 34158222]
8. Pallathra AA, Cordero L, Wong K, Brodkin ES. Psychosocial interventions targeting social functioning in adults on the autism spectrum: a literature review. Curr Psychiatry Rep. Jan 31, 2019;21(1):5. [FREE Full text] [doi: 10.1007/s11920-019-0989-0] [Medline: 30706149]
9. Martins PL, Torquato GC, Dias GA, Leite IB, Gaspar TM, Pinto JP, et al. Effectiveness of pharmacological interventions for managing ADHD symptoms in individuals with autism spectrum disorder: a systematic review and meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry. Aug 30, 2024;134:111089. [doi: 10.1016/j.pnpbp.2024.111089] [Medline: 39004333]
10. Mazefsky CA, Kao J, Oswald DP. Preliminary evidence suggesting caution in the use of psychiatric self-report measures with adolescents with high-functioning autism spectrum disorders. Res Autism Spectr Disord. Jan 2011;5(1):164-174. [FREE Full text] [doi: 10.1016/j.rasd.2010.03.006] [Medline: 24013401]
11. Cassidy SA, Bradley L, Bowen E, Wigham S, Rodgers J. Measurement properties of tools used to assess depression in adults with and without autism spectrum conditions: a systematic review. Autism Res. May 2018;11(5):738-754. [FREE Full text] [doi: 10.1002/aur.1922] [Medline: 29360200]
12. Schwartzman JM, Hardan AY, Gengoux GW. Parenting stress in autism spectrum disorder may account for discrepancies in parent and clinician ratings of child functioning. Autism. Aug 2021;25(6):1601-1614. [doi: 10.1177/1362361321998560] [Medline: 33691519]
13. Palmer M, Tarver J, Carter Leno V, Paris Perez J, Frayne M, Slonims V, et al. Parent, teacher and observational reports of emotional and behavioral problems in young autistic children. J Autism Dev Disord. Jan 2023;53(1):296-309. [FREE Full text] [doi: 10.1007/s10803-021-05421-x] [Medline: 35028809]
14. Sonuga-Barke EJ, Brandeis D, Cortese S, Daley D, Ferrin M, Holtmann M, et al. European ADHD Guidelines Group. Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. Am J Psychiatry. Mar 2013;170(3):275-289. [doi: 10.1176/appi.ajp.2012.12070991] [Medline: 23360949]
15. Jones RM, Carberry C, Hamo A, Lord C. Placebo-like response in absence of treatment in children with Autism. Autism Res. Sep 2017;10(9):1567-1572. [doi: 10.1002/aur.1798] [Medline: 28401674]

16. Siafis S, Çıray O, Schneider-Thoma J, Bighelli I, Krause M, Rodolico A, et al. Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): systematic review and meta-regression analysis. Mol Autism. Aug 26, 2020;11(1):66. [FREE Full text] [doi: 10.1186/s13229-020-00372-z] [Medline: 32847616]

17. Chatham CH, Taylor KI, Charman T, Liogier D'ardhuy X, Eule E, Fedele A, et al. Adaptive behavior in autism: minimal clinically important differences on the Vineland-II. Autism Res. Feb 2018;11(2):270-283. [FREE Full text] [doi: 10.1002/aur.1874] [Medline: 28941213]

18. Simons AM, Beltramo T, Blalock G, Levine DI. Using unobtrusive sensors to measure and minimize Hawthorne effects: evidence from cookstoves. J Environ Econ Manage. Nov 2017;86:68-80. [doi: 10.1016/j.jeem.2017.05.007]

19. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol. 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]

20. van der Stam JA, Mestrom EH, Nienhuijs SW, de Hingh IH, Boer A, van Riel NA, et al. A wearable patch based remote early warning score (REWS) in major abdominal cancer surgery patients. Eur J Surg Oncol. Jan 2023;49(1):278-284. [FREE Full text] [doi: 10.1016/j.ejso.2022.08.034] [Medline: 36085116]

21. Dobkin BH, Carmichael ST. The specific requirements of neural repair trials for stroke. Neurorehabil Neural Repair. Jun 2016;30(5):470-478. [FREE Full text] [doi: 10.1177/1545968315604400] [Medline: 26359342]

22. Mirelman A, Volkov J, Salomon A, Gazit E, Nieuwboer A, Rochester L, et al. Digital mobility measures: a window into real-world severity and progression of Parkinson's disease. Mov Disord. Feb 2024;39(2):328-338. [doi: 10.1002/mds.29689] [Medline: 38151859]

23. Bruno E, Biondi A, Böttcher S, Vértes G, Dobson R, Folarin A, et al. Remote assessment of disease and relapse in epilepsy: protocol for a multicenter prospective cohort study. JMIR Res Protoc. Dec 16, 2020;9(12):e21840. [FREE Full text] [doi: 10.2196/21840] [Medline: 33325373]

24. Siddi S, Bailon R, Giné-Vázquez I, Matcham F, Lamers F, Kontaxis S, et al. RADAR-CNS consortium. The usability of daytime and night-time heart rate dynamics as digital biomarkers of depression severity. Psychol Med. Jun 2023;53(8):3249-3260. [doi: 10.1017/S0033291723001034] [Medline: 37184076]

25. Zhang Y, Folarin AA, Sun S, Cummins N, Ranjan Y, Rashid Z, et al. Predicting depressive symptom severity through individuals' nearby bluetooth device count data collected by mobile phones: preliminary longitudinal study. JMIR Mhealth Uhealth. Jul 30, 2021;9(7):e29840. [FREE Full text] [doi: 10.2196/29840] [Medline: 34328441]

26. Georgescu AL, Koehler JC, Weiske J, Vogeley K, Koutsouleris N, Falter-Wagner C. Machine learning to study social interaction difficulties in ASD. Front Robot AI. 2019;6:132. [FREE Full text] [doi: 10.3389/frobt.2019.00132] [Medline: 33501147]

27. Zampella CJ, Bennetto L, Herrington JD. Computer vision analysis of reduced interpersonal affect coordination in youth with autism spectrum disorder. Autism Res. Dec 2020;13(12):2133-2142. [FREE Full text] [doi: 10.1002/aur.2334] [Medline: 32666690]

28. Parish-Morris J, Sariyanidi E, Zampella C, Bartley GK, Ferguson E, Pallathra AA, et al. Oral-motor and lexical diversity during naturalistic conversations in adults with autism spectrum disorder. Proc Conf. Jun 2018;2018:147-157. [FREE Full text] [doi: 10.18653/v1/w18-0616] [Medline: 33073267]

29. Hauser M, Sariyanidi E, Tunc B, Zampella C, Brodkin E, Schultz R. Using natural conversations to classify autism with limited data: age matters. In: Proceedings of the 6th Workshop on Computational Linguistics and Clinical Psychology. 2019. Presented at: CLPsych '19; June 6, 2019:45-54; Minneapolis, MN. URL: https://aclanthology.org/W19-3006.pdf [doi: 10.18653/v1/w19-3006]

30. Egger HL, Dawson G, Hashemi J, Carpenter KL, Espinosa S, Campbell K, et al. Automatic emotion and attention analysis of young children at home: a ResearchKit autism feasibility study. NPJ Digit Med. 2018;1:20. [FREE Full text] [doi: 10.1038/s41746-018-0024-6] [Medline: 31304303]

31. Xiong M, Li F, Liu Z, Xie X, Shen H, Li W, et al. Efficacy of melatonin for insomnia in children with autism spectrum disorder: a meta-analysis. Neuropediatrics. Jun 2023;54(3):167-173. [FREE Full text] [doi: 10.1055/s-0043-1761437] [Medline: 36827993]

32. Keogh S, Bridle C, Siriwardena NA, Nadkarni A, Laparidou D, Durrant SJ, et al. Effectiveness of non-pharmacological interventions for insomnia in children with Autism Spectrum Disorder: a systematic review and meta-analysis. PLoS One. 2019;14(8):e0221428. [FREE Full text] [doi: 10.1371/journal.pone.0221428] [Medline: 31437211]

33. Morgan B, Nageye F, Masi G, Cortese S. Sleep in adults with autism spectrum disorder: a systematic review and meta-analysis of subjective and objective studies. Sleep Med. Jan 2020;65:113-120. [doi: 10.1016/j.sleep.2019.07.019] [Medline: 31739229]

34. Ness SL, Bangerter A, Manyakov NV, Lewin D, Boice M, Skalkin A, et al. An observational study with the Janssen Autism Knowledge Engine (JAKE) in individuals with autism spectrum disorder. Front Neurosci. 2019;13:111. [FREE Full text] [doi: 10.3389/fnins.2019.00111] [Medline: 30872988]

35. Bangerter A, Chatterjee M, Manyakov NV, Ness S, Lewin D, Skalkin A, et al. Relationship between sleep and behavior in autism spectrum disorder: exploring the impact of sleep variability. Front Neurosci. 2020;14:211. [FREE Full text] [doi: 10.3389/fnins.2020.00211] [Medline: 32265629]

36.   Jagannatha S, Sargsyan D, Manyakov NV, Skalkin A, Bangerter A, Ness S, et al. A practical application of data mining methods to build predictive models for autism spectrum disorder based on biosensor data from Janssen Autism Knowledge Engine (JAKE®). Stat Biopharm Res. May 28, 2019;11(2):111-117. [doi: 10.1080/19466315.2018.1527247]

37.   Bangerter A, Manyakov NV, Lewin D, Boice M, Skalkin A, Jagannatha S, et al. Caregiver daily reporting of symptoms in autism spectrum disorder: observational study using web and mobile apps. JMIR Ment Health. Mar 26, 2019;6(3):e11365. [FREE Full text] [doi: 10.2196/11365] [Medline: 30912762]

38.   Zervogianni V, Fletcher-Watson S, Herrera G, Goodwin M, Pérez-Fuster P, Brosnan M, et al. A framework of evidence-based practice for digital support, co-developed with and for the autism community. Autism. Aug 2020;24(6):1411-1422. [FREE Full text] [doi: 10.1177/1362361319898331] [Medline: 32026715]

39.   Gringras P, Nir T, Breddy J, Frydman-Marom A, Findling RL. Efficacy and safety of pediatric prolonged-release melatonin for insomnia in children with autism spectrum disorder. J Am Acad Child Adolesc Psychiatry. Nov 2017;56(11):948-57.e4. [FREE Full text] [doi: 10.1016/j.jaac.2017.09.414] [Medline: 29096777]

40.   Yavuz-Kodat E, Reynaud E, Geoffray MM, Limousin N, Franco P, Bourgin P, et al. Validity of actigraphy compared to polysomnography for sleep assessment in children with autism spectrum disorder. Front Psychiatry. 2019;10:551. [FREE Full text] [doi: 10.3389/fpsyt.2019.00551] [Medline: 31428003]

41.   Lord C, Rutter M, DiLavore P, Risi S, Gotham K, Bishop S. Autism Diagnostic Observation Schedule (ADOS-2). 2nd edition. San Antonio, TX. Pearson Assessments; 2012.

42.   Green J, Leadbitter K, Ellis C, Taylor L, Moore HL, Carruthers S, et al. Combined social communication therapy at home and in education for young autistic children in England (PACT-G): a parallel, single-blind, randomised controlled trial. Lancet Psychiatry. Apr 2022;9(4):307-320. [FREE Full text] [doi: 10.1016/S2215-0366(22)00029-3] [Medline: 35305746]

43.   Fletcher-Watson S, Petrou A, Scott-Barrett J, Dicks P, Graham C, O'Hare A, et al. A trial of an iPad™ intervention targeting social communication skills in children with autism. Autism. Oct 2016;20(7):771-782. [FREE Full text] [doi: 10.1177/1362361315605624] [Medline: 26503990]

44.   Le J, Zhang L, Zhao W, Zhu S, Lan C, Kou J, et al. Infrequent intranasal oxytocin followed by positive social interaction improves symptoms in autistic children: a pilot randomized clinical trial. Psychother Psychosom. 2022;91(5):335-347. [FREE Full text] [doi: 10.1159/000524543] [Medline: 35545057]

45.   Provenzani U, Fusar-Poli L, Brondino N, Damiani S, Vercesi M, Meyer N, et al. What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. Autism. Feb 2020;24(2):274-284. [doi: 10.1177/1362361319854641] [Medline: 31269800]

46.   Mathersul D, McDonald S, Rushby JA. Autonomic arousal explains social cognitive abilities in high-functioning adults with autism spectrum disorder. Int J Psychophysiol. Sep 2013;89(3):475-482. [doi: 10.1016/j.ijpsycho.2013.04.014] [Medline: 23628291]

47.   Hollocks MJ, Howlin P, Papadopoulos AS, Khondoker M, Simonoff E. Differences in HPA-axis and heart rate responsiveness to psychosocial stress in children with autism spectrum disorders with and without co-morbid anxiety. Psychoneuroendocrinology. Aug 2014;46:32-45. [doi: 10.1016/j.psyneuen.2014.04.004] [Medline: 24882156]

48.   Muscatello RA, Andujar J, Taylor JL, Corbett BA. Exploring key physiological system profiles at rest and the association with depressive symptoms in autism spectrum disorder. J Autism Dev Disord. Jan 2021;51(1):15-29. [FREE Full text] [doi: 10.1007/s10803-020-04516-1] [Medline: 32350791]

49.   Zhang Y, Folarin AA, Sun S, Cummins N, Bendayan R, Ranjan Y, et al. RADAR-CNS Consortium. Relationship between major depression symptom severity and sleep collected using a wristband wearable device: multicenter longitudinal observational study. JMIR Mhealth Uhealth. Apr 12, 2021;9(4):e24604. [FREE Full text] [doi: 10.2196/24604] [Medline: 33843591]

50.   Matcham F, Barattieri di San Pietro C, Bulgari V, de Girolamo G, Dobson R, Eriksson H, et al. RADAR-CNS consortium. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. BMC Psychiatry. Feb 18, 2019;19(1):72. [FREE Full text] [doi: 10.1186/s12888-019-2049-z] [Medline: 30777041]

51.   Langener AM, Stulp G, Kas MJ, Bringmann LF. Capturing the dynamics of the social environment through experience sampling methods, passive sensing, and egocentric networks: scoping review. JMIR Ment Health. Mar 17, 2023;10:e42646. [FREE Full text] [doi: 10.2196/42646] [Medline: 36930210]

52.   Doryab A, Villalba DK, Chikersal P, Dutcher JM, Tumminia M, Liu X, et al. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and Fitbit data. JMIR Mhealth Uhealth. Jul 24, 2019;7(7):e13209. [FREE Full text] [doi: 10.2196/13209] [Medline: 31342903]

53.   Eskes P, Spruit M, Brinkkemper S, Vorstman J, Kas MJ. The sociability score: app-based social profiling from a healthcare perspective. Comput Human Behav. Jun 2016;59:39-48. [doi: 10.1016/j.chb.2016.01.024]

54.   Charman T, Loth E, Tillmann J, Crawley D, Wooldridge C, Goyard D, et al. The EU-AIMS Longitudinal European Autism Project (LEAP): clinical characterisation. Mol Autism. 2017;8:27. [FREE Full text] [doi: 10.1186/s13229-017-0145-9] [Medline: 28649313]

55. Loth E, Charman T, Mason L, Tillmann J, Jones EJ, Wooldridge C, et al. The EU-AIMS Longitudinal European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. Mol Autism. 2017;8:24. [FREE Full text] [doi: 10.1186/s13229-017-0146-8] [Medline: 28649312]

56. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, Hyve, et al. RADAR-CNS Consortium. RADAR-base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR Mhealth Uhealth. Aug 01, 2019;7(8):e11734. [FREE Full text] [doi: 10.2196/11734] [Medline: 31373275]

57. Milstein N, Gordon I. Validating measures of electrodermal activity and heart rate variability derived from the Empatica E4 utilized in research settings that involve interactive dyadic states. Front Behav Neurosci. 2020;14:148. [FREE Full text] [doi: 10.3389/fnbeh.2020.00148] [Medline: 33013337]

58. O'Sullivan J, Bogaarts G, Schoenenberger P, Tillmann J, Slater D, Mesgarani N, et al. Automatic speaker diarization for natural conversation analysis in autism clinical trials. Sci Rep. Jun 24, 2023;13(1):10270. [FREE Full text] [doi: 10.1038/s41598-023-36701-4] [Medline: 37355730]

59. Bredin H, Yin R, Coria JM, Gelly G, Korshunov P, Lavechin M. Pyannote.audio: neural building blocks for speaker diarization. In: Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. 2020. Presented at: ICASSP '20; May 4-8, 2020:7124-7128; Barcelona, Spain. URL: https://ieeexplore.ieee.org/document/9052974 [doi: 10.1109/icassp40776.2020.9052974]

60. Radford A, Kim JW, Xu T, Brockman G, McLeavey C, Sutskever I. Robust speech recognition via large-scale weak supervision. In: Proceedings of the 40th International Conference on Machine Learning. 2023. Presented at: ICML '23; July 23-29, 2023:28492-28518; Honolulu, HI. URL: https://dl.acm.org/doi/10.5555/3618408.3619590 [doi: 10.48550/arXiv.2212.04356]

61. Chang JP, Chiam C, Fu L, Wang AZ, Zhang J, Danescu-Niculescu-Mizil C. ConvoKit: a toolkit for the analysis of conversations. In: Proceedings of the 21th Annual Meeting of the Special Interest Group on Discourse and Dialogue. 2020. Presented at: SIGDIAL '20; July 1-3, 2020:57-60; Virtual Event. URL: https://aclanthology.org/2020.sigdial-1.8.pdf [doi: 10.18653/v1/2020.sigdial-1.8]

62. Barami T, Manelis-Baram L, Kaiser H, Ilan M, Slobodkin A, Hadashi O, et al. Automated analysis of stereotypical movements in videos of children with autism spectrum disorder. JAMA Netw Open. Sep 03, 2024;7(9):e2432851. [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.32851] [Medline: 39264628]

63. Kleckner IR, Jones RM, Wilder-Smith O, Wormwood JB, Akcakaya M, Quigley KS, et al. Simple, transparent, and flexible automated quality assessment procedures for ambulatory electrodermal activity data. IEEE Trans Biomed Eng. Jul 2018;65(7):1460-1467. [FREE Full text] [doi: 10.1109/TBME.2017.2758643] [Medline: 28976309]

64. Dutta A, Zisserman A. The VIA annotation software for images, audio and video. In: Proceedings of the 27th ACM International Conference on Multimedia. 2019. Presented at: MM '19; October 21-25, 2019:2276-2279; Nice, France. URL: https://dl.acm.org/doi/10.1145/3343031.3350535

65. Cao Z, Hidalgo G, Simon T, Wei S, Sheikh Y. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell. Jan 2021;43(1):172-186. [doi: 10.1109/TPAMI.2019.2929257] [Medline: 31331883]

66. Duan H, Zhao Y, Chen K, Lin D, Dai B. Revisiting skeleton-based action recognition. In: Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022. Presented at: CVPR '22; June 18-24, 2022:2959-2968; New Orleans, LA. URL: https://ieeexplore.ieee.org/document/9879048 [doi: 10.1109/cvpr52688.2022.00298]

67. Baltrusaitis T, Zadeh A, Lim YC, Morency LP. OpenFace 2.0: facial behavior analysis toolkit. In: Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition. 2018. Presented at: FG '18; May 15-19, 2018:59-66; Xi'an, China. URL: https://ieeexplore.ieee.org/document/8373812 [doi: 10.1109/fg.2018.00019]

68. Drimalla H, Scheffer T, Landwehr N, Baskow I, Roepke S, Behnia B, et al. Towards the automatic detection of social biomarkers in autism spectrum disorder: introducing the simulated interaction task (SIT). NPJ Digit Med. 2020;3:25. [FREE Full text] [doi: 10.1038/s41746-020-0227-5] [Medline: 32140568]

69. Tarvainen MP, Niskanen JP, Lipponen JA, Ranta-Aho PO, Karjalainen PA. Kubios HRV--heart rate variability analysis software. Comput Methods Programs Biomed. 2014;113(1):210-220. [doi: 10.1016/j.cmpb.2013.07.024] [Medline: 24054542]

70. Föll S, Maritsch M, Spinola F, Mishra V, Barata F, Kowatsch T, et al. FLIRT: a feature generation toolkit for wearable data. Comput Methods Programs Biomed. Nov 2021;212:106461. [FREE Full text] [doi: 10.1016/j.cmpb.2021.106461] [Medline: 34736174]

71. Hu X, Sgherza TR, Nothrup JB, Fresco DM, Naragon-Gainey K, Bylsma LM. From lab to life: evaluating the reliability and validity of psychophysiological data from wearable devices in laboratory and ambulatory settings. Behav Res Methods. Oct 2024;56(7):1-20. [doi: 10.3758/s13428-024-02387-3] [Medline: 38528248]

72. Benedek M, Kaernbach C. A continuous measure of phasic electrodermal activity. J Neurosci Methods. Jun 30, 2010;190(1):80-91. [FREE Full text] [doi: 10.1016/j.jneumeth.2010.04.028] [Medline: 20451556]

73. Benedek M, Kaernbach C. Decomposition of skin conductance data by means of nonnegative deconvolution. Psychophysiology. Jul 01, 2010;47(4):647-658. [FREE Full text] [doi: 10.1111/j.1469-8986.2009.00972.x] [Medline: 20230512]

74. pandas-dev/pandas: Pandas. Zenodo. URL: https://zenodo.org/records/10957263 [accessed 2025-05-29]

75. Rahimi-Eichi H, Coombs Iii G, Vidal Bustamante CM, Onnela J, Baker JT, Buckner RL. Open-source longitudinal sleep analysis from accelerometer data (DPSleep): algorithm development and validation. JMIR Mhealth Uhealth. Oct 06, 2021;9(10):e29849. [FREE Full text] [doi: 10.2196/29849] [Medline: 34612831]

76. Johnson KT, Picard RW. Advancing neuroscience through wearable devices. Neuron. Oct 14, 2020;108(1):8-12. [FREE Full text] [doi: 10.1016/j.neuron.2020.09.030] [Medline: 33058768]

77. Goodwin MS, Mazefsky CA, Ioannidis S, Erdogmus D, Siegel M. Predicting aggression to others in youth with autism using a wearable biosensor. Autism Res. Aug 2019;12(8):1286-1296. [FREE Full text] [doi: 10.1002/aur.2151] [Medline: 31225952]

78. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). Br J Clin Psychol. Sep 1992;31(3):301-306. [doi: 10.1111/j.2044-8260.1992.tb00997.x] [Medline: 1393159]

79. Grant C, Widnall E, Cross L, Simonoff E, Downs J. Informing the development of an E-platform for monitoring wellbeing in schools: involving young people in a co-design process. Res Involv Engagem. 2020;6:51. [FREE Full text] [doi: 10.1186/s40900-020-00219-0] [Medline: 32908677]

80. Monitoring mental health. Digital Marmalade. URL: https://www.digitalmarmalade.co.uk/case-studies/king-s-college-london/monitoring-mental-health [accessed 2024-10-31]

81. Rutherford M, Baxter J, Grayson Z, Johnston L, O'Hare A. Visual supports at home and in the community for individuals with autism spectrum disorders: a scoping review. Autism. Feb 2020;24(2):447-469. [doi: 10.1177/1362361319871756] [Medline: 31451016]

82. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of Wristband Fitbit models in assessing sleep: systematic review and meta-analysis. J Med Internet Res. Nov 28, 2019;21(11):e16273. [FREE Full text] [doi: 10.2196/16273] [Medline: 31778122]

83. Remotely monitoring autistic characteristics and co-occurring difficulties with RADAR-base in AIMS-2-TRIALS. The Hyve. URL: https://www.thehyve.nl/cases/remotely-monitoring-autistic-characteristics-and-co-occurring-difficulties-aims-2-trials [accessed 2024-06-21]

84. Sturm A, Kuhfeld M, Kasari C, McCracken JT. Development and validation of an item response theory-based Social Responsiveness Scale short form. J Child Psychol Psychiatry. Sep 2017;58(9):1053-1061. [doi: 10.1111/jcpp.12731] [Medline: 28464350]

85. Hennig T, Krkovic K, Lincoln TM. What predicts inattention in adolescents? An experience-sampling study comparing chronotype, subjective, and objective sleep parameters. Sleep Med. Oct 2017;38:58-63. [doi: 10.1016/j.sleep.2017.07.009] [Medline: 29031757]

86. Snyder E, Cai B, DeMuro C, Morrison MF, Ball W. A new single-item sleep quality scale: results of psychometric evaluation in patients with chronic primary insomnia and depression. J Clin Sleep Med. Nov 15, 2018;14(11):1849-1857. [FREE Full text] [doi: 10.5664/jcsm.7478] [Medline: 30373688]

87. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol. Jun 1988;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]

88. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Weerdmeester BA, McClelland AL, editors. Usability Evaluation In Industry. London, UK. Taylor and Francis; 1996:11-26.

89. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact. Jul 30, 2008;24(6):574-594. [doi: 10.1080/10447310802205776]

90. Greer B, Newbery K, Cella M, Wykes T. Predicting inpatient aggression in forensic services using remote monitoring technology: qualitative study of staff perspectives. J Med Internet Res. Sep 19, 2019;21(9):e15620. [FREE Full text] [doi: 10.2196/15620] [Medline: 31538943]

91. Benevides TW, Shore SM, Palmer K, Duncan P, Plank A, Andresen M, et al. Listening to the autistic voice: mental health priorities to guide research and practice in autism from a stakeholder-driven project. Autism. May 2020;24(4):822-833. [FREE Full text] [doi: 10.1177/1362361320908410] [Medline: 32429818]

92. Wallace S, Parr J, Hardy A. One in a hundred: putting families at the heart of autism research. Autistics. 2013. URL: https://media.tghn.org/medialibrary/2015/07/One-in-a-Hundred-Autisticas-Report.pdf [accessed 2024-12-20]

93. McIntosh DN, Miller LJ, Shyu V, Dunn W. Overview of the short sensory profile. In: Dunn W, editor. Sens Profile Man. San Antonio, TX. Psychological Corporation; 1999:59-73.

94. Böttcher S, Vieluf S, Bruno E, Joseph B, Epitashvili N, Biondi A, et al. Data quality evaluation in wearable monitoring. Sci Rep. Dec 10, 2022;12(1):21412. [FREE Full text] [doi: 10.1038/s41598-022-25949-x] [Medline: 36496546]

95. Campanella S, Altaleb A, Belli A, Pierleoni P, Palma L. PPG and EDA dataset collected with Empatica E4 for stress assessment. Data Brief. Apr 2024;53:110102. [FREE Full text] [doi: 10.1016/j.dib.2024.110102] [Medline: 38328286]

96. Montesinos V, Dell'Agnola F, Arza A, Aminifar A, Atienza D. Multi-modal acute stress recognition using off-the-shelf wearable devices. Annu Int Conf IEEE Eng Med Biol Soc. Jul 2019;2019:2196-2201. [doi: 10.1109/EMBC.2019.8857130] [Medline: 31946337]

97. Brown VA. An introduction to linear mixed-effects modeling in R. Adv Methods Pract Psychol Sci. Mar 25, 2021;4(1):251524592096035. [doi: 10.1177/2515245920960351]

98. Sun S, Folarin AA, Zhang Y, Cummins N, Garcia-Dias R, Stewart C, et al. RADAR-CNS Consortium. Challenges in using mHealth data from smartphones and wearable devices to predict depression symptom severity: retrospective analysis. J Med Internet Res. Aug 14, 2023;25:e45233. [FREE Full text] [doi: 10.2196/45233] [Medline: 37578823]

99. Dong Y, Peng CY. Principled missing data methods for researchers. Springerplus. Dec 2013;2(1):222. [FREE Full text] [doi: 10.1186/2193-1801-2-222] [Medline: 23853744]

100. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods. Jun 2002;7(2):147-177. [Medline: 12090408]

101. Díaz-Román A, Zhang J, Delorme R, Beggiato A, Cortese S. Sleep in youth with autism spectrum disorders: systematic review and meta-analysis of subjective and objective studies. Evid Based Ment Health. Nov 2018;21(4):146-154. [FREE Full text] [doi: 10.1136/ebmental-2018-300037] [Medline: 30361331]

102. Rohani DA, Faurholt-Jepsen M, Kessing LV, Bardram JE. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. JMIR Mhealth Uhealth. Aug 13, 2018;6(8):e165. [FREE Full text] [doi: 10.2196/mhealth.9691] [Medline: 30104184]

103. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. Depress Anxiety. Jan 2019;36(1):72-81. [FREE Full text] [doi: 10.1002/da.22822] [Medline: 30129691]

104. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. Jun 1961;4:561-571. [doi: 10.1001/archpsyc.1961.01710120031004] [Medline: 13688369]

105. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol. Dec 1988;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]

106. Fulford D, Mote J, Gonzalez R, Abplanalp S, Zhang Y, Luckenbaugh J, et al. Smartphone sensing of social interactions in people with and without schizophrenia. J Psychiatr Res. May 2021;137:613-620. [FREE Full text] [doi: 10.1016/j.jpsychires.2020.11.002] [Medline: 33190842]

107. Welch V, Wy TJ, Ligezka A, Hassett LC, Croarkin PE, Athreya AP, et al. Use of mobile and wearable artificial intelligence in child and adolescent psychiatry: scoping review. J Med Internet Res. Mar 14, 2022;24(3):e33560. [FREE Full text] [doi: 10.2196/33560] [Medline: 35285812]

108. Oakley B, Boatman C, Doswell S, Dittner A, Clarke A, Ozsivadjian A, et al. Molehill Mountain Advisory Group. Molehill Mountain feasibility study: protocol for a non-randomised pilot trial of a novel app-based anxiety intervention for autistic people. PLoS One. 2023;18(7):e0286792. [FREE Full text] [doi: 10.1371/journal.pone.0286792] [Medline: 37406026]

109. Picard RW. Future affective technology for autism and emotion communication. Philos Trans R Soc Lond B Biol Sci. Dec 12, 2009;364(1535):3575-3584. [FREE Full text] [doi: 10.1098/rstb.2009.0143] [Medline: 19884152]

110. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth. Aug 09, 2018;6(8):e10527. [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]

111. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Performance assessment of new-generation Fitbit technology in deriving sleep parameters and stages. Chronobiol Int. Jan 2020;37(1):47-59. [doi: 10.1080/07420528.2019.1682006] [Medline: 31718308]

112. Charman T, Pickles A, Simonoff E, Chandler S, Loucas T, Baird G. IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). Psychol Med. Mar 2011;41(3):619-627. [doi: 10.1017/S0033291710000991] [Medline: 21272389]

113. Puli A, Kushki A. Toward automatic anxiety detection in autism: a real-time algorithm for detecting physiological arousal in the presence of motion. IEEE Trans Biomed Eng. Mar 2020;67(3):646-657. [doi: 10.1109/TBME.2019.2919273] [Medline: 31144623]

114. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res. Jul 12, 2018;20(7):e10480. [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]

115. Shandhi MM, Singh K, Janson N, Ashar P, Singh G, Lu B, et al. Assessment of ownership of smart devices and the acceptability of digital health data sharing. NPJ Digit Med. Feb 22, 2024;7(1):44. [FREE Full text] [doi: 10.1038/s41746-024-01030-x] [Medline: 38388660]

116. Bietz MJ, Bloss CS, Calvert S, Godino JG, Gregory J, Claffey MP, et al. Opportunities and challenges in the use of personal health data for health research. J Am Med Inform Assoc. Apr 2016;23(e1):e42-e48. [FREE Full text] [doi: 10.1093/jamia/ocv118] [Medline: 26335984]

117. Kristiansson E, Fridolfsson J, Arvidsson D, Holmäng A, Börjesson M, Andersson-Hall U. Validation of Oura ring energy expenditure and steps in laboratory and free-living. BMC Med Res Methodol. Feb 24, 2023;23(1):50. [FREE Full text] [doi: 10.1186/s12874-023-01868-x] [Medline: 36829120]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder

ADOS-2: Autism Diagnostic Observation Schedule-2
aRMT: active remote measurement technology
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)
DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)
EDA: electrodermal activity
ICD-10: International Statistical Classification of Diseases, Tenth Revision
JAKE: Janssen Autism Knowledge Engine
KCL: King's College London
LEAP: AIMS Longitudinal European Autism Project
LMEM: linear mixed effects model
PANAS: Positive and Negative Affect Scale
pRMT: passive remote measurement technology
RM: remote measurement
RUMC: Radboud University Medical Centre, Nijmegen
UCAM: University of Cambridge

XSL•FO
RenderX