RADAR-base Platform: Digital phenotyping of mental and physical conditions through remotely collected wearables and smartphone data

Zulqarnain Rashid, Amos A Folarin, Yuezhou Zhang, Yatharth Ranjan, Pauline Conde, Heet Sankesara, Shaoxiong Sun, Callum Stewart, Petroula Laiou, Richard JB Dobson

Submitted to: JMIR Mental Health
on: July 27, 2023

Table of Contents

Original Manuscript **5**
Supplementary Files **31**
Multimedia Appendixes **32**
Multimedia Appendix 1 ..... **32**

RADAR-base Platform: Digital phenotyping of mental and physical conditions through remotely collected wearables and smartphone data

Zulqarnain Rashid[1] PhD; Amos A Folarin[1, 2, 3, 4, 5] PhD; Yuezhou Zhang[1] PhD; Yatharth Ranjan[1] MSc; Pauline Conde[1] BSc; Heet Sankesara[1] BSc; Shaoxiong Sun[1] PhD; Callum Stewart[1] PhD; Petroula Laiou[1] PhD; Richard JB Dobson[1, 2, 3, 4, 5] PhD

[1]Institute of Psychiatry, Psychology and Neuroscience, King's College London King's College London London GB
[2]Institute of Health Informatics, University College London London GB
[3]NIHR Biomedical Research Center at South London and Maudsley NHS Foundation Trust and King's College London London GB
[4]Health Data Research UK London, University College London, London London GB
[5]NIHR Biomedical Research Center at University College London Hospitals, NHS Foundation Trust London GB

Corresponding Author:
Zulqarnain Rashid PhD
Institute of Psychiatry, Psychology and Neuroscience, King's College London
King's College London
De Crespigny Park, Denmark Hill
London
GB

Abstract

Background: The integration of digital biomarkers and remote patient monitoring offers valuable and timely insights into a patient's management of their condition, including aspects such as disease progression and treatment response. This serves as a complementary resource to traditional healthcare settings leveraging mobile technology to improve scale and lower latency, cost and burden.

Objective: Smartphones with embedded and connected sensors have immense potential for improving healthcare through various apps and mHealth (mobile health) platforms. This capability could enable the development of reliable digital biomarkers from long-term longitudinal data collected remotely from patients.

Methods: We built an open-source platform, RADAR-base, to support large-scale data collection in remote monitoring studies. RADAR-base is a modern remote data collection platform built around Confluent's Apache Kafka, to support scalability, extensibility, security, privacy and quality of data. It provides support for study design and set-up, active (eg PROMs) and passive (e.g. phone sensors, wearable devices and IoT) remote data collection capabilities with feature generation (e.g. behavioural, environmental and physiological markers). The backend enables secure data transmission, and scalable solutions for data storage, management and data access.

Results: The platform has been used to successfully collect longitudinal data for various cohorts in a number of disease areas including Multiple Sclerosis, Depression, Epilepsy, ADHD, Alzheimer, Autism and Lung diseases. Digital biomarkers developed through collected data are providing useful insights into different diseases.

Conclusions: RADAR-base offers a contemporary, open-source solution driven by the community for remotely monitoring, collecting data, and digitally characterising both physical and mental health conditions. Clinicians have the ability to enhance their insight through the utilisation of digital biomarkers, enabling improved prevention, personalisation, and early intervention in the context of disease management.

(JMIR Preprints 27/07/2023:51259)
DOI: https://doi.org/10.2196/preprints.51259

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✔ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

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

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

RADAR-base Platform: Digital phenotyping of mental and physical conditions through remotely collected wearables and smartphone data

Zulqarnain Rashid[1], PhD; Amos A Folarin[12345], PhD; Yuezhou Zhang[1], PhD; Yatharth Ranjan[1], MSc; Pauline Conde[1], BSc; Heet Sankesara[1], BSc; Shaoxiong Sun[1], PhD; Callum Stewart[1], PhD; Petroula Laiou[1], PhD; Richard JB Dobson[12345], PhD

1 Institute of Psychiatry, Psychology and Neuroscience, King's College London.
2 Institute of Health Informatics, University College London, London, United Kingdom.
3 NIHR Biomedical Research Center at South London and Maudsley NHS Foundation Trust and King's College London, London, United Kingdom.
4 Health Data Research UK London, University College London, London, United Kingdom.
5 NIHR Biomedical Research Center at University College London Hospitals, NHS Foundation Trust, London, United Kingdom.

Corresponding Author

Amos A Folarin
Institute of Psychiatry, Psychology and Neuroscience
King's College London
SGDP Centre, IoPPN, Box PO 92
De Crespigny Park, Denmark Hill, London , SE5 8AF
Tel: +44(0)207 848 0924
email: amos.folarin@kcl.ac.uk

Abstract

Background

The use of digital biomarkers through remote patient monitoring offers valuable and timely insights into a patient's condition, including aspects such as disease progression and treatment response. This serves as a complementary resource to traditional healthcare settings leveraging mobile technology to improve scale and lower latency, cost and burden.

Objectives

Smartphones with embedded and connected sensors have immense potential for improving healthcare through various apps and mHealth (mobile health) platforms. This capability could enable the development of reliable digital biomarkers from long-term longitudinal data collected remotely from patients.

Methods

We built an open-source platform, RADAR-base, to support large-scale data collection in remote monitoring studies. RADAR-base is a modern remote data collection platform built around Confluent's Apache Kafka, to support scalability, extensibility, security, privacy and quality of data. It provides support for study design and set-up, active (eg PROMs) and passive (e.g. phone sensors, wearable devices and IoT) remote data collection capabilities with feature generation (e.g. behavioural, environmental and physiological markers). The backend enables secure data transmission, and scalable solutions for data storage, management and data access.

Results

The platform has been used to successfully collect longitudinal data for various cohorts in a number of disease areas including Multiple Sclerosis, Depression, Epilepsy, ADHD, Alzheimer's disease, Autism and Lung diseases. Digital biomarkers developed through collected data are providing useful insights into different diseases.

Conclusion

RADAR-base offers a contemporary, open-source solution driven by the community for remotely monitoring, collecting data, and digitally characterising both physical and mental health conditions. Clinicians have the ability to enhance their insight through the utilisation of digital biomarkers, enabling improved prevention, personalisation, and early intervention in the context of disease management.

Keywords: digital biomarkers; mHealth; mobile apps; Internet of Things (IoT); remote data collection; wearables; real-time monitoring; platform;

Introduction

Digital biomarkers offer a host of advantages for measuring our health over traditional assessment approaches that are typically confined to clinical settings, including decentralisation, scalability, sampling frequency and real-time measurement, and affordability. However, significant challenges remain with implementing digital biomarkers.

Digital biomarkers developed from sensor data can help with prevention and early intervention to better diagnose and manage disease. Collected data should be reliable and of high quality reflecting the true condition of the patients and many studies have attempted to measure the effectiveness of digital biomarkers for various clinical use cases [1].

High quality longitudinal data collected for long periods and at scale is a key requirement for digital biomarker development. The widespread availability of smartphones, more capacious mobile networks and the development of new wearable sensors has enabled measurement of a growing set of physiological and phenomenological parameters relevant to physical and mental diseases. To facilitate wearable and smartphone data remote collection at scale and digital biomarker development, the RADAR-base platform was developed and released under the open-source Apache 2 licence in January 2018 [2] [3]. RADAR-base comprises an Apache Kafka-based back-end deployed onto Kubernetes infrastructure and two mobile apps. The cross-platform (Android, iOS) Cordova Active Remote Monitoring App (aRMT) for active monitoring of participants, requiring conscious action (e.g. questionnaires, audio questions, timed tests etc.), and the Passive Remote Monitoring App (pRMT), a native Android app for passive monitoring of participants via phone, wearable devices and IoT. A high-level overview of the platform is shown in Figure 1.

The aRMT app renders questionnaires using JSON definition files, which store data in key-value pairs. Within a single questionnaire file, exists a collection of questions, each containing attributes such as field name, label or text, input type (checkbox, free text, etc.), choices, and more. Subsequently, the aRMT app employs these files to display the questionnaire within the user interface. This facilitates the dynamic deployment of questionnaires for wide ranging project requirements. Example questionnaires used in the aRMT app in existing projects include measures of self-esteem (Rosenberg Self-Esteem Scale or RSES), depression (Patient Health Questionnaire-8 or PHQ8), and ecological momentary assessment (ESM). These questionnaires include traditional question sets, while others take a different approach. For instance, the speech questionnaire requires users to record themselves reading a specific text or answering a question, rather than responding to written prompts.

The passive application runs in the background, requiring minimal or no input from participants. Data is collected from smartphone "sensors" and from integrated wearable devices. The catalogue of devices currently integrated into the pRMT app includes onboard Android smartphone sensors, Empatica E4, Pebble 2 smartwatch, Biovotion Everion, Faros 180 and 360, Fitbit, Garmin Vivosmart, and Oura Ring. Pluggable capability is provided to integrate new wearable devices offering a native SDK (e.g. Empatica E4) or through 3rd party vendor's REST API (e.g. Fitbit via the backend REST Collector + REST-Authorizer for OAuth-2 Flows).

A common task is the exploration of collected raw data. RADAR-base includes capability for data aggregation, management of studies, and real-time visualisation in Grafana dashboards [4]. In addition to the near real-time visualisation provided by the dashboard, the RADAR-base platform includes a Python package designed for data processing, feature generation and visualisation. This package offers a range of standard tools for exploratory visualisations of collected data. It also simplifies the implementation of feature generation pipelines, allowing users to take data exported from a RADAR-base project and generate processed data (high-level features), along with any associated labels, in a format suitable for use with commonly used machine learning libraries.

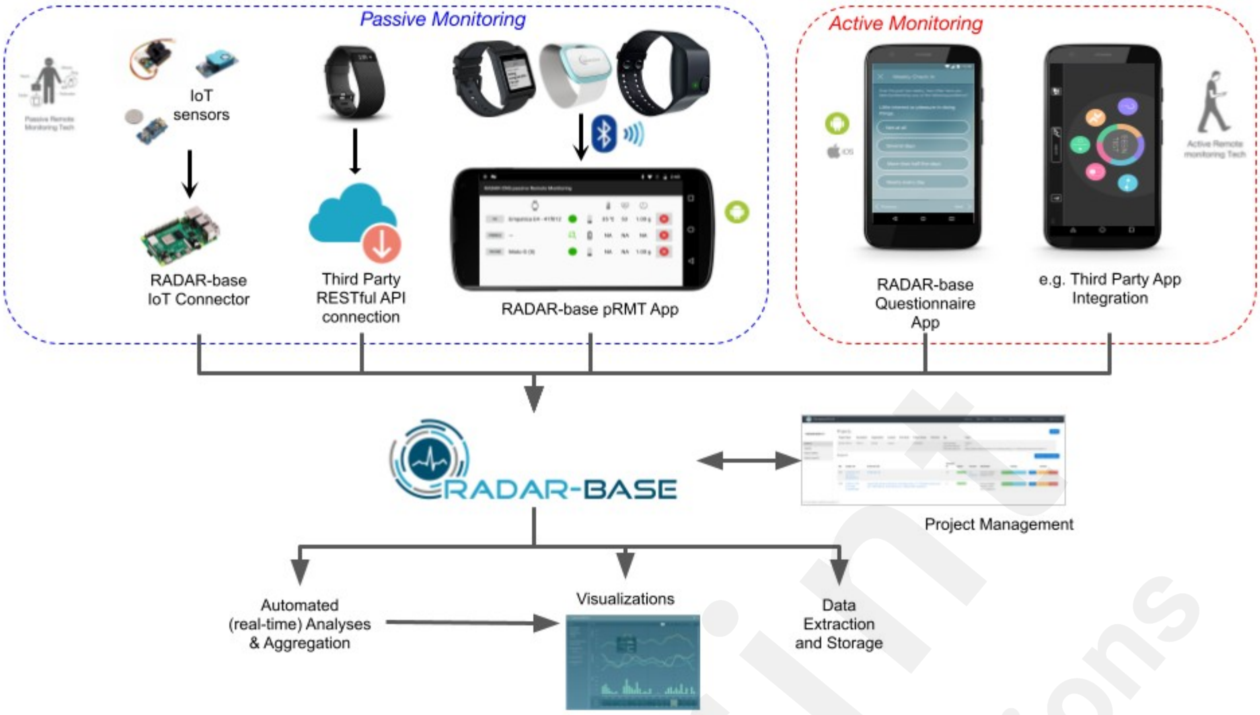

Figure 1. Overview of the RADAR-base Platform. Current data sources: Empatica E4, Pebble 2, Fitbit, Biovotion, Faros, Garmin; Active aRMT Questionnaire app, and Passive pRMT app.

The Platform is used in a wide range of research and clinical studies and is available under a range of service models, depending on the requirement of the project. This paper extends the foundation laid by the initial iteration of the RADAR-base platform [3], describing its architecture and technical components in thorough detail. The present work aims to encapsulate and outline a pertinent assortment of research and clinical studies that leverage the RADAR-base platform for data collection and digital phenotyping. The focus of this paper will be on distilling prevalent usage patterns and addressing the challenges encountered in employing the platform for diverse research endeavours.

Related Work and Comparison with other platforms

Numerous studies are validating digital biomarkers for disorders and their effectiveness, studies [5] and [6] were conducted to assess the usefulness of digital biomarkers for mood and depression. Both studies had smaller cohorts of only 60 and 59 participants respectively which limits the studies findings. Digital biomarkers are being studied with a view to replacing, or augmenting traditional markers for disorders and a number of barriers have been identified, these barriers include standardisation and regulation, studies are undergoing to address challenges and streamline digital biomarkers in healthcare [7]. Open source software platform for end-to-end digital biomarker development Digital Biomarker Discovery Pipeline (DBDP) was developed to standardise digital biomarker development. DBDP modules calculate and utilise resting heart rate (RHR), glycemic variability, insulin sensitivity status, exercise response, inflammation, heart rate variability, activity, sleep, and circadian patterns to predict health outcomes [8]. mCerebrum: A Mobile Sensing Software Platform for Development and Validation of Digital Biomarkers and Interventions supports high-rate data collections from multiple sensors with real time assessment of data quality and development of digital biomarkers [9]. Guidelines for developing digital biomarkers are proposed in the study [10].
Intel Context Sensing SDK is a library for Android and Windows with specific context states, it

however only provides front-end components [11]. The EmotionSense app is developed by the University of Cambridge to sense emotions with implications for psychological therapy and improving well-being, however, it is only focused on depression [12]. Medopad provides solutions for different healthcare issues with symptom tracking, this is a commercial solution and mainly focuses on phone sensors and active monitoring methods [13]. PHIT allows users to build health apps based on existing infrastructure [14]. ResearchKit, an open-source framework for building apps specifically for iOS, ResearchKit makes it easier to enrol participants and conduct studies, however, new wearable device integration requires strong programming skills and it does not include a data management solution [15].
ResearchStack is an SDK and UX framework for building research study apps on Android, with a similar application domain as ResearchKit [16] Both ResearchKit and ResearchStack provide software libraries, frameworks, and development tools that require extensive programming skills to create apps. A framework to create observational medical studies for mobile devices without extensive programming skills was presented [17]. LAMP (Learn Assess Manage and Prevent) platform [18] provides an app and backend infrastructure for clinical relevant studies, app can adapt to different studies with input from patients.

One of the distinctive features of the RADAR-base platform is its utilisation of Confluent platform technologies [19], which are built around Apache Kafka. This choice forms the foundation for a highly scalable end-to-end solution for event-driven messaging, capable of addressing diverse use cases such as high throughput, low latency messaging, real-time data processing, and fault tolerance. Deployable as microservices with Kubernetes cluster [20], the platform offers seamless integration for new sensors and data sources with minimal effort. Number of new sensors added to the platform since its inception for clinical studies. Minimal effort to integrate new sensors and devices offers opportunity to capture signals of various diseases and use cases.

There is a need for a systematic approach to assess the quality and utility of digital biomarkers to ensure an appropriate balance between their safety and effectiveness. Study [21] outlines key considerations for the development and evaluation of digital biomarkers, examining their role in clinical research and routine patient care. RADAR-base provides a secure and effective digital biomarker ecosystem ensuring transparency of the algorithms, interoperable components with open interfaces to accelerate the development of new multicomponent systems, and high integrity measurement systems.

Digital Phenotyping of Disease

A key feature of RADAR-base is its extensibility allowing new wearables and sensors to be readily integrated into the platform to collect new modes of data depending on the study requirements. Collected raw data from phone and wearable sensors can then be aggregated and converted into low level features, and subsequently high-level features, representing digital biomarkers . Figure 2 exemplifies the process for Major Depressive Disorder (MDD) for wearable and phone sensor collected data. E.g. phone microphone collected audio data can provide different speech features and respiratory acoustics which could help identify respiration and physiological stress, similarly low level acceleration provided actigraphy features can be used to identify psychomotor retardation. Developed digital biomarkers may provide insight into the disorder and could be used in clinical trials to ascertain their usefulness in management of the disease.

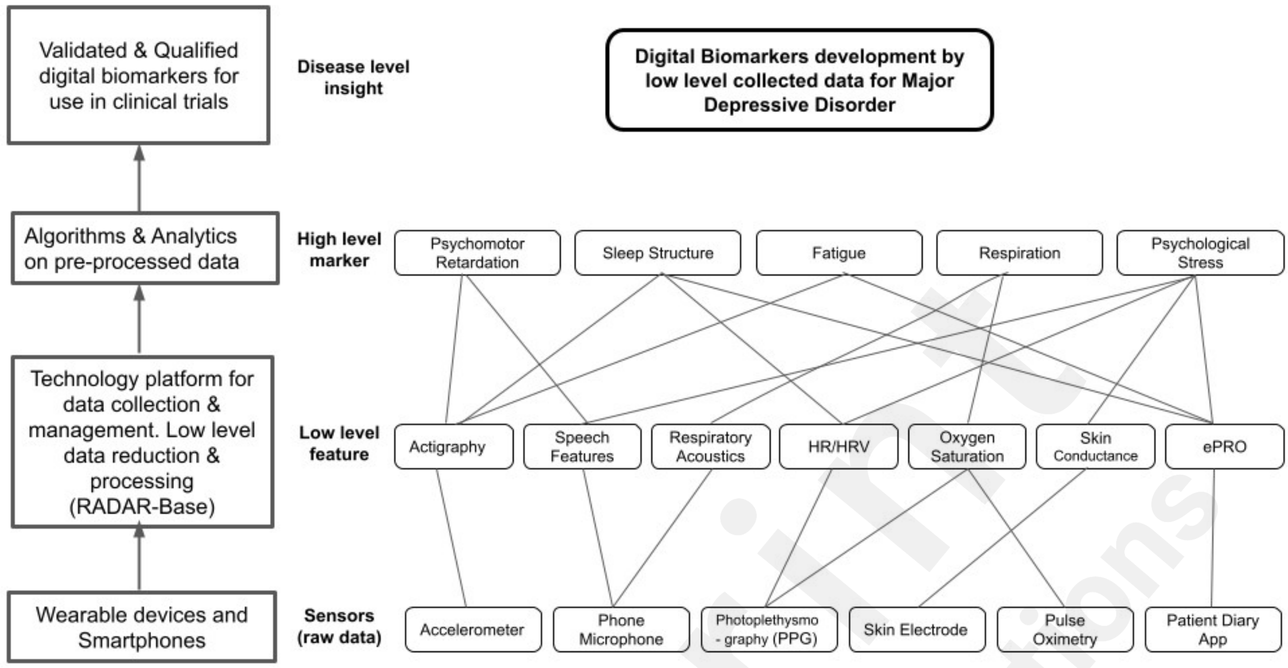

Figure 2. Collected raw data transition into high level features which gives insight into Major Depressive Disorder.

The next two sections discuss the development of high level features for MDD and Epilepsy and how these features are being used to explore and manage disease. Table 1 shows selected features extracted from sensor data and Table 2 presents digital biomarkers developed for different disorder areas [22] [23] [24]. An open-source feature generation pipeline has been created to enhance and standardise the analysis of data generated by RADAR-base. This pipeline facilitates the extraction of features and biomarkers, enabling cross-disease symptom analysis. With its capabilities to ingest, analyse, visualise, and export RADAR-base data, the pipeline simplifies and establishes a convention for data scientists. This streamlines the process of feature-based analysis, ensuring consistency and enabling researchers to gain valuable insights from the data. Pipelines are readily extended and published on the RADAR-base pipeline catalogue [25].

Major Depressive Disorder

Major Depressive Disorder (MDD) is associated with a wide range of negative outcomes including: premature mortality, reduced quality-of-life, loss of occupational function and is often experienced alongside physical comorbidity and approximately 55% will go on to develop chronic depression, characterised by periods of recovery and relapse [26].

The pRMT app provides a comprehensive solution for activity monitoring by utilising wearable device sensors and smartphones to collect data without requiring any input from the wearer. It leverages a range of sensors, including Global Positioning System (GPS), accelerometer, gyroscope, communication logs, ambient noise and light levels, and screen interactions. Through this approach, the app can effectively and passively gather diverse data streams, enabling a seamless and unintrusive data collection process for various applications. These sensors along with the Fitbit watch have the potential to identify changes in sleep,

communication and activity patterns associated with depressive episodes.
The smartphone embedded Bluetooth sensor can be used to record individuals' local proximity information, such as the nearby Bluetooth device count (NBDC) that includes the Bluetooth signal of other phone users. The NBDC data have the potential to reflect changes in people's behaviours and statuses during the depressive state [27]. An illustration is given in Figure 3.

Speech characteristics, such as speaking rate, pitch, pause duration and energy, collected via smartphone microphones, have been used to detect depression with a prediction accuracy of 81.3% [28]. aRMT app delivers validated questionnaires, cognitive games, speech tasks or electronic diaries using the experience sampling method (ESM) to provide fine grained understanding of mood changes and stressors in the context of daily life. The aRMT app has been used to measure effect, cognition, mood and behaviour in real time, with evidence highlighting the increased validity of this methodology in comparison to traditional retrospective reports. The aRMT assessments of positive and negative affect have also been found to be reliably indicative of mood state and have been associated with MDD symptoms [29].

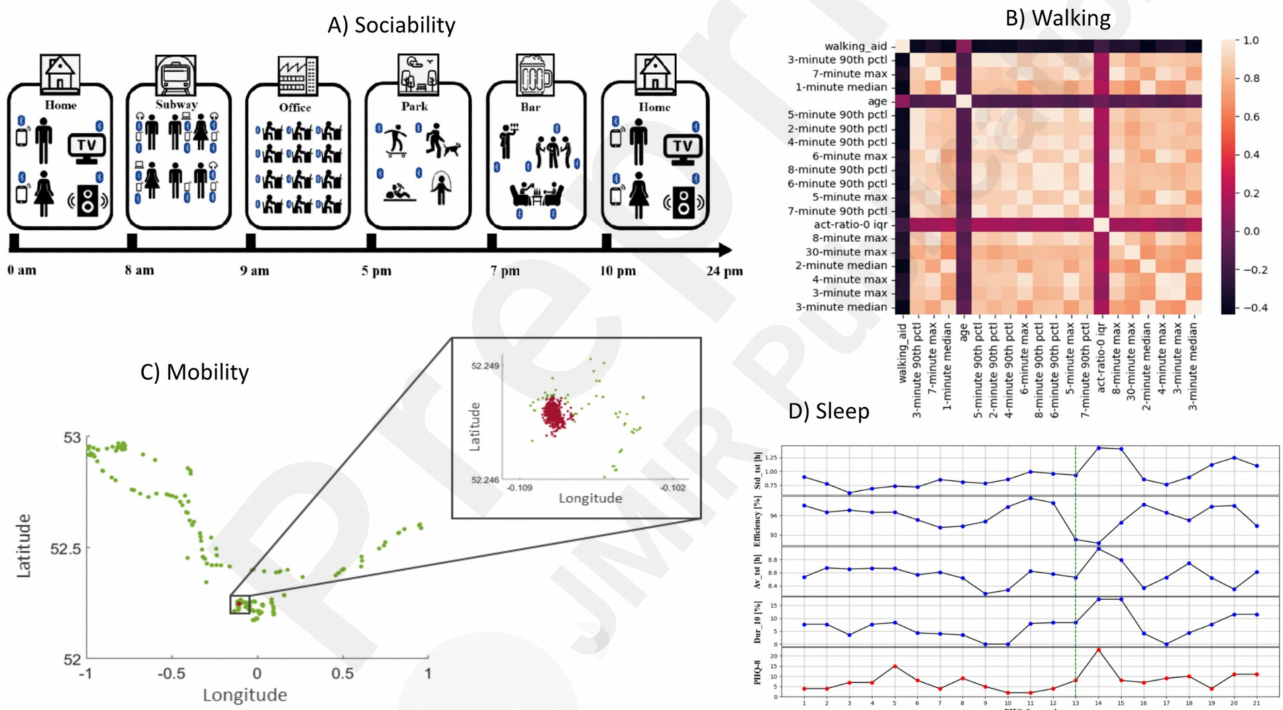

Figure 3. (a) A schematic diagram showing an individual's nearby Bluetooth devices count (NBDC) in different scenarios in daily activities and life. (b) Pearson correlation heatmap for top 20 features of walking data (median rankings from all models). (c) Exemplar geolocation data correspond to a biweekly segment of a study participant, The red dots denote an individual's home location cluster, whereas longitude and latitude along the axes are expressed in decimal degrees. (d) The PHQ-8 scores and a select 4 sleep features of one participant with an obvious increasing trend in PHQ-8 score at 13th PHQ-8 record.

Epilepsy

Numerous epilepsy research studies based on epilepsy monitoring units (EMUs), have shown the possibility of capturing characteristic movement associated with myoclonic seizure manifestations using wearable sensors [30]. Using pRMT app integrated wearable sensors, it is possible to record several signals associated with seizure including motor components, using inertial sensors such as accelerometry and surface electromyography, various features of Heart rate variations captured by wearable electrocardiogram (ECG) and photoplethysmography (PPG), and alteration of the autonomic nervous system with ECG, PPG, and electrodermal activity (EDA) sensors, with different levels of signal and seizure detection accuracy. Figure 4 shows seizures detected with the collected data using the Empatica E4 wearable.

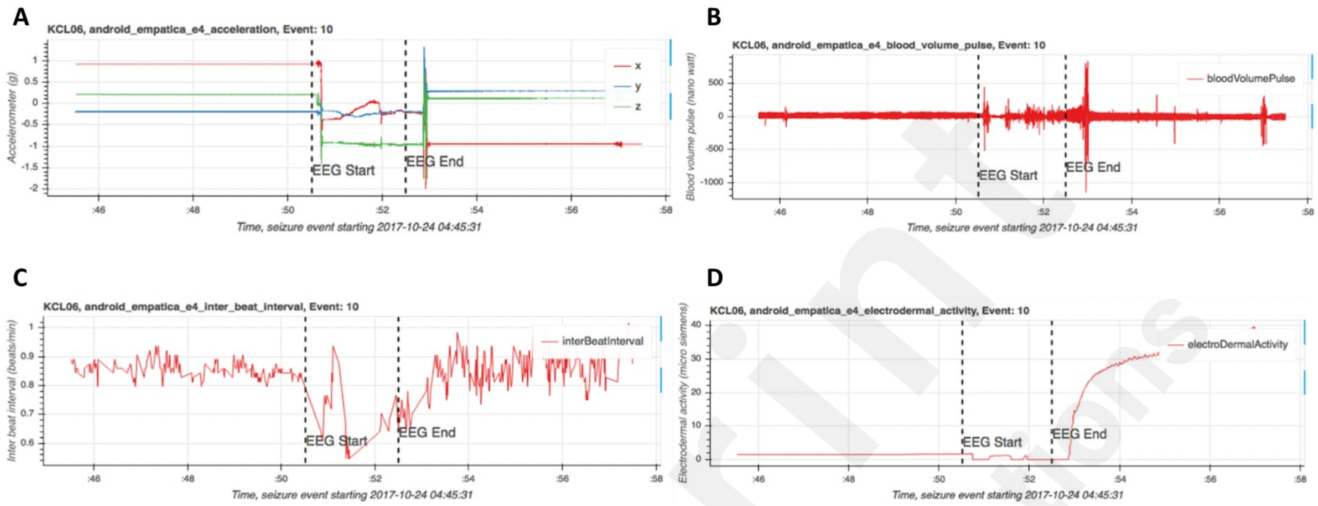

Figure 4. Empatica E4 sensor data. The area between the vertical dashed lines indicates a focal seizure with a motor component. (a) Accelerometer, (b) PPG blood volume pulse, (c) PPG inter-beat-interval, (d) Electrodermal activity.

Table1. Features extracted from the sensors integrated into the RADAR-base platform.

Apps	Devices	Sensors/Raw Data	Selected Features
aRMT pRMT app	Empatica E4 Biovotion Everion Bittium Faros 180 Faros 360 Fitbit Garmin Smart Phone Oura Ring	Acceleration Blood Volume Pulse Electrodermal Activity (EDA) Inter Beat Interval Temperature Blood Pulse Wave Galvanic Skin Response Heart Rate Oxygen Saturation Led Current PPPG Raw ECG Gyroscope Light Magnetic Field Location Microphone Step Count	Sleep Duration Sleep Architecture Sleep Stability Sleep Quality Sleep Efficiency Sleep Fragmentation Index Sleep Onset Latency Sleep Onsent Latency Variance Sleep Midpoint Sleep Midpoint Variance Insomnia Hypersomnia Unlocktimes/Duration Unlock Duration Min/Max Median Interval Between Two Unlocks Step Epoch Daily Step Sum Moderate walking duration Maximum non-stop duration

		Usage Event User Interaction Activity Levels Activity Log Record Intraday Calories Intraday Steps Resting Heart Rate Sleep Classic Sleep Stage SMS Unread Bluetooth Devices Phone Battery Level Phone Contact List Garmin Stress Tracking Garmin Relaxation Breathing timer Garmin Vo2 max Garmin Body Battery Energy Monitor SpO2	Maximum non-stop step count Activity level Actigraphy Respiratory Acoustics Heart Rate Heart Rate Variation Oxygen Saturation Skin Conductance ePRO Ambient Light Activity Phone Use Bluetooth Max/Min/Mean/SD of nearby Bluetooth Devices (NBDC) NBDC Entropy NBDC Frequency Features Location Location variance Moving time Moving distance Number of Location Clusters Location Entropy Homestay Location Frequency Features Gait Median Gait Cycles Frequency of gait Median Force Change in Total Sleep Social Jet Lag

Table 2. Digital biomarkers generated from extracted features, and associated studies.

Digital Biomarker	Device	Sensor/Raw data	RADAR-CNS	RADAR-AD	AIMS-2-TRIALS	ART	BigData@Heart	RALPMH	COVID-Collab
Total Sleep	Fitbit/Garmin	Sleep stage	✔	✔	✔	✔	✔	✔	✔
Social interactions	Smartphone	Bluetooth	✔		✔	✔			
Gait patterns	Smartphone	Acceleration	✔	✔		✔			
Respiration	Garmin	Garmin Respiration Rate						✔	✔
Psychological Stress	Garmin	Garmin Stress			✔		✔	✔	✔
Phone Use	Smartphone	User Interaction	✔		✔	✔	✔		
Ambulatory Mobility	Smartphone	Location	✔			✔			✔
Fatigue	Garmin	Garmin Body Battery Energy						✔	✔
Seizures	Empatica E4/Biovotion Everion	Acceleration, EDA, PPG	✔						
Step	Fitbit/Garmin	Step count	✔	✔		✔			✔
Activity	Fitbit/Garmin	Activity logs	✔	✔	✔	✔		✔	✔
Speech	Smartphone	Microphone	✔		✔			✔	

Resting HR	Fitbit/Garmin	PPG	✔	✔		✔		✔	✔
HR Variability	Fitbit/Garmin	PPG	✔	✔		✔		✔	✔
Sleep Variance	Fitbit/Garmin	Sleep stage	✔	✔		✔		✔	✔
Mobility Variance	Smartphone	Location	✔	✔		✔		✔	✔
Respiratory Acoustics	Smartphone	Microphone	✔	✔				✔	

Participant Recruitment Process

Participants are recruited using different methods, including through clinical services, hospitals, and remotely including through a citizen science approach, depending on the study requirements. A number of recruitment strategies are supported by the platform including:

- All participants are recruited at once, and the study starts simultaneously.
- Participants enter the study in a "batch" mode.
- Participants are recruited continuously until the desired sample size or date is reached ("stream mode").

Simultaneous recruitment from multiple sites is possible supporting recruitment of diverse population groups for the same study.

Projects using RADAR-base platform

Table 3 presents a summary of selected projects using the platform with disorders they are focused on, along with the cohort size and sensors being used. It also lists the main objectives of the project. Brief summary of the projects listed in Table 3 are provided in the next sections. In some projects, wearables data collected through the platform augments existing collected data, e.g. historical clinical records, or at baseline assessment.

Table 3: Projects summary with disease area and study size including devices employed and main objectives.

Project	Disease Area	Size	Enrolled	Devices/Data Types	Main Objectives
RADAR-CNS	Depression	600	623	Fitbit, Phone Sensors, Questionnaires	Depressive Relapse
	Epilepsy	200	145	Biovotion Everion, Empatica E4, Questionnaires	Epilepsy seizure and pre-ictal seizure detection
	Multiple Sclerosis	500	430	Fitbit, Bittium Faros, Phone Sensors, Questionnaires	Trajectory of disease, characterisation relapsing/remitting of disease symptoms
ART-CARMA	Cardiometabolic Risk Factors	300	200- Ongoing	Empatica EmbracePlus, Phone Sensors, Questionnaires	Pre-treatment initiation through to treatment initiation, titration and the subsequent period
ART	Attention Deficit Hyperactivity Disorder	40	40	Fitbit, Phone Sensors, Questionnaires	To establish a remote assessment and monitoring system for adults and adolescents with ADHD
RADAR-AD	Alzheimer's Disease	200	229	Fitbit, Phone Sensors, Questionnaires	Feasibility of remote monitoring technologies for AD
AIMS-2-TRIALS	Autism	500	300 - Ongoing	Empatica E4, Fitbit, Phone Sensors, Questionnaires	Biology of autism to tailor treatments and develop new therapies and

					medicines.
BigData@Heart	Atrial Fibrillation	160	160	Phone Sensors, Questionnaires	Comparison of two strategies of rate-control, based either on initial treatment with digoxin or beta-blockers
DynaMORE	Mental Health	-	-	Phone Sensors, Questionnaires	Developing an in silico model of stress resilience
CONVALESCENCE	Long-term Effects of COVID-19	800	363 - Ongoing	Garmin Vivoactive, Phone Sensors, Questionnaires	Characterisation, determinants, mechanisms and consequences of the long-term effects of COVID-19
COVID-Collab	COVID-19	15000	17667	Fitbit, Questionnaires	Behaviour and physical and mental health occur in response to COVID infection, persistent symptoms, and the pandemic in general
RALPMH	Lung Disease	60	60	Garmin Vivoactive, Phone Sensors, Questionnaires	Feasibility of remote monitoring technologies for high-burden pulmonary disorders
EDIFY	Eating Disorder	500	10 - Ongoing	Oura Ring, Phone Sensors, Questionnaires	Delineating illness and recovery trajectories to inform personalised

					prevention and early intervention in young people
UNFOLD	Psychosis	50	-	Questionnaires	To characterise the processes involved in developing an identity as a person in recovery
Jovens na Pandemia & MAAY Study	Depression	280	-	Phone Sensors, Questionnaires	Remotely monitor behavioural and symptom changes associated with behavioural interventions in children and adolescents.

Remote Assessment of Disease And Relapse – Central Nervous System (RADAR-CNS)

RADAR-CNS was a cohort study that developed new ways of monitoring major depressive disorder, epilepsy, and multiple sclerosis using wearable devices and smartphone technology. Patients' data were collected continuously for 24 months [31]. More than 1200 participants took part in the study in different disease areas, and participant recruitment was done via clinics and hospitals. Different study protocols with different wearable devices were used for each disease. Participants were recruited from 6 different sites from different countries.
Digital biomarkers developed through the remotely collected data give a better understanding of the diseases and will help clinicians to manage them timely [32] [33] [27].

ADHD Remote Technology Study of Cardiometabolic Risk Factors and Medication Adherence (ART-CARMA)

ART-CARMA aims to obtain real-world data from the patient's daily life to explore the extent to which ADHD medication and physical activity, individually and jointly, may influence cardiometabolic risks in adults with ADHD. The second objective is to obtain valuable real-world data on adherence to pharmacological treatment and its predictors and correlates. The long-term goal is to use collected data to improve the management of cardiometabolic disease in adults with ADHD, and to improve ADHD medication treatment adherence and the personalisation of treatment [34]. For this cohort, two study sites in London and Barcelona are concurrently recruiting the participants using the platform.

ADHD Remote Technology (ART)

The ART was a pilot project focused on developing a novel remote assessment system for Attention Deficit Hyperactivity Disorder (ADHD). ART assessed the feasibility and validity of remote researcher-led administration and self-administration of modified versions of two cognitive tasks sensitive to ADHD, a four-choice reaction time task (Fast task) and a combined Continuous Performance Test/Go No-Go task (CPT/GNG) [35]. A cohort of 40 participants was recruited, 20 controls and 20 patients with ADHD.

Remote Assessment of Disease And Relapse – Alzheimer's Disease (RADAR-AD)

RADAR-AD aimed to transform Alzheimer's disease patient care through remote assessment using mobile technologies such as smartphones or fitness trackers [36]. The project developed the technology to identify which clinical or physiological features, digital biomarkers, can be measured remotely to predict deterioration in function. RADAR-AD created a pipeline for developing, testing and implementing remote measurement technologies with patients involved at each stage. Complete details of the study protocol and pipeline development is explained in [37]. It was an augmentation study in which 300 participants took part. Three different categories of participants were recruited, controls, mild cognitively impaired (MCI)/prodromal AD and Alzheimer's dementia.

Autism Innovative Medicine Studies-2-Trials (AIMS-2-TRIALS)

The AIMS-2-TRIALS programme includes a range of studies to explore how autism develops, from before birth to adulthood, and how this varies in different people. AIMS-2-TRIALS is looking for biological markers which indicate whether a person has or may develop particular characteristics [38]. AIMS-2-TRIALS collects both active and passive data in clinical assessment settings and in home-based and ambulatory settings. Fitbit is used for remote data collection, and Empatica E4 is used for local data collection at hospitals. Digital markers could help to identify those who may ultimately benefit from particular treatments. Medicines will also be tested to help with social difficulties, repetitive behaviours and sensory processing. Remote monitoring data is augmenting the clinical data.

Rate Control Therapy Evaluation in Permanent Atrial Fibrillation (BigData@Heart)

RATE-AF study was designed to compare two strategies of rate-control, based either on initial treatment with digoxin or beta-blockers in 160 patients with Atrial fibrillation (AF) in need for rate control therapy. Monitoring with wearable devices, phone sensors and questionnaires was conducted over a continuous 6-month period. Objectives of the project included discovering new phenotypes, developing reliable sub-phenotyping and informing new taxonomies of heart failure based on a better understanding of underlying disease processes. Sleep, Heath Rate, Heart Rate Variation, and Activity data were collected to develop new phenotypes. This work is additionally supported by the EU IMI2 BigData@Heart major programme [39].

Dynamic Modelling of Resilience (DynaMORE)

DynaMORE generated the first personalised in silico model of mental health in the face of adversity or stress resilience. The model is based on and validated against unique multiscale longitudinal real-world empirical data sets, collected through neuroimaging, experimental assessments, questionnaires and remote monitoring using the pRMT app and a wearable device. The model will substantially deepen scientific understanding of the mechanisms of resilience, supporting the creation of mechanistically targeted interventions for the primary prevention of stress-related disorders. On this basis, DynaMORE developed an entirely new mobile health (mHealth) product incorporating the RADAR-base platform that will include model-based prognostic tools for real-time and real-life monitoring of at-risk subjects and for automated decision-making about timed, personalised interventions.

CONVALESCENCE

CONVALESCENE is focused on the characterisation, determinants, mechanisms and consequences of the long-term effects of COVID-19, providing the evidence base for health care services [40]. It's an existing large longitudinal cohort being further characterised and augmented by incorporating wearables data. Deep phenotyping and remote assessment using mobile devices and smartphones through the RADAR-base platform is being used to identify subclinical damage or dysfunction in individuals with long-term COVID-19.

COVID Collab

COVID-Collab is a citizen science project with members of the public volunteering to donate their wearable data and complete diagnosis and symptom questionnaires. The main aim was to investigate the ongoing COVID-19 outbreak - 1) establish if wearable data can be used to diagnose COVID infection and 2) characterise the disease symptoms and evolution. A key feature of the study is the use of wearable data to investigate changes in mental health and physiological measurements such as heart rate during infection with coronavirus [41].

RALPMH: Remote Assessment of Lung Disease and Impact on Physical and Mental Health

Chronic lung disorders like chronic obstructive pulmonary disease (COPD) and idiopathic pulmonary fibrosis (IPF) are characterised by exacerbations and decline over time. 20 participants were recruited in each of three cohorts (COPD, IPF, and post-hospitalisation COVID). Data collection is being done remotely using the RADAR-base platform for different devices, including Garmin wearable devices and smart spirometers, mobile app questionnaires, surveys, and finger pulse oximeters. The RALPMH project focuses on the feasibility of remote monitoring in chronic lung disorders and provide a reference infrastructure for future studies [42]. Figure 5 shows an overview of the RALPHM study.

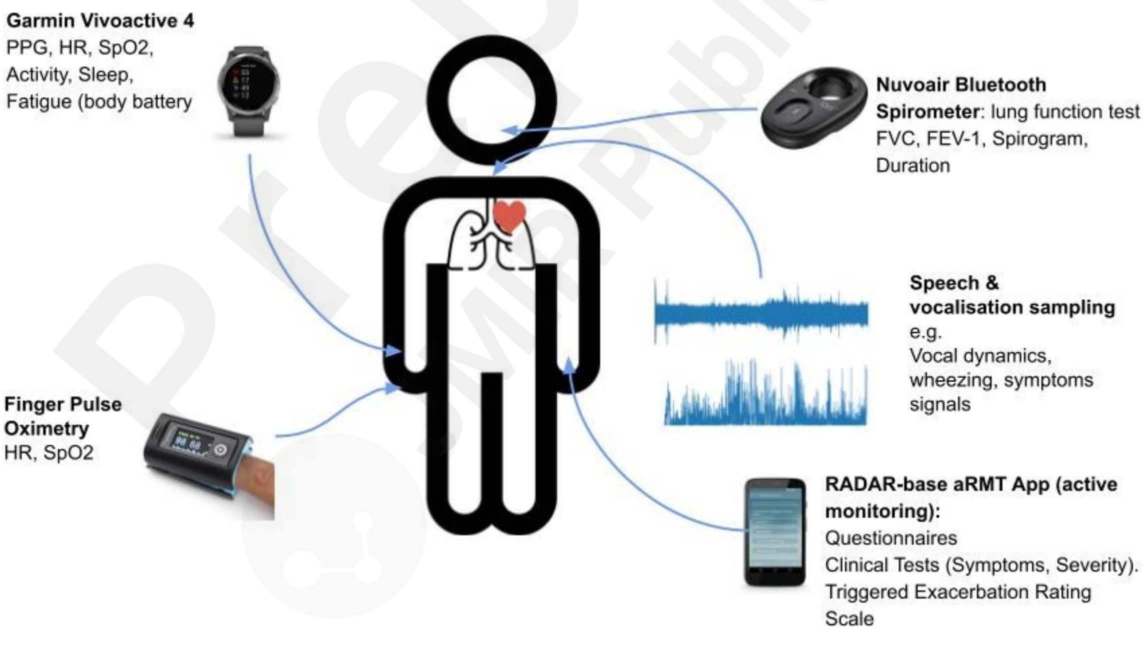

Figure 5. Various data sources will be used to collect both active and passive data to gain a unique perspective into patient health in the RALPMH lung disorder study.

Eating Disorder (EDIFY)

The objective of the EDIFY study is to undertake a longitudinal comparison of the biopsychosocial symptom profiles of those with early and late-stage Eating Disorders

(EDs), and recovery trajectories of those with early-stage EDs. This will provide evidence that will help inform decision-making of targeted intervention and preventative treatments across EDs for those with early and more progressed forms of illness. ED patients do not like to wear any wearable device or have any information displayed on their phone regarding their calories and daily workout. The Oura ring is being experimented for the first time with ED patients at scale as it provides no feedback to the patients during the study.

Ethical Considerations

Prior to initiating each project, we obtained approval from the Research Ethics Committee (REC) and conducted a Data Protection Impact Assessment (DPIA), which is equivalent to an Institutional Review Board (IRB) in some countries. Each study adhered to its unique data collection protocol, and both the GDPR/DPIA and the protocol were approved by the REC and the Data Protection Officer (DPO), respectively. In [3], comprehensive information regarding data protection, privacy, and the process of pseudonymization is expounded.

Data Availability

The collected data which is used to develop biomarkers in the mentioned studies are typically available on reasonable request from the respective principle investigators, requests may be directed to corresponding authors, AF, RD. The data is not publicly accessible due to the ethical constraints of the study. Corresponding authors will formally request approval from the project's principal investigator before sharing the data.

Using RADAR-base platform and Hosting Models

RADAR-base is of interest to a wide range of mHealth communities, from academic research to industry and sensor/wearable vendors interested in collecting data remotely or integrating new data sources into the platform. The platform is freely available as an open-source (Apache 2 Licence) GitHub repository [43] more details of the platform can be found on the official RADAR-base website [44]. A detailed quickstart, deployment details and developer documentation are made available on the platform Confluence Wiki [45]. Docker images for all the components are available at Docker Hub [46] and a Kubernetes stack is also available for the deployment [47]. aRMT app questionnaires and protocol implementation is explained in [48]. An exemplar Data Protection Impact Assessment (DPIA) which explains the data collection procedure is shared in Appendix 1. In principle three hosting models are available for using the platform:

Self Hosting

The platform, along with its accompanying apps and related elements, is open source, allowing users to host it on private or local servers and tailor it to suit their study requirements. Under the self-hosting model, users possess full autonomy over deployment, infrastructure, and data collection.

Supported Hosting

Under the Supported Hosting model, a third party provider of RADAR-base can deploy the platform on local or public cloud such as Amazon Web Services (AWS). The costing and support arrangements will be determined by factors such as the study's duration, participant count, and data throughput, primarily based on sensor types selected. In this setup, users will share control over the infrastructure and data with the RADAR-base team. Integration with eCRFs such as REDCap also allow more sensitive data to be segregated from the managed service and be retained entirely under the user's control.

Fully-Managed Hosting

Option is available to completely outsource the deployment and hosting to a third party, provides the platform deployed on their infrastructure and hosts projects as a service to e.g. researchers.

Results and Discussion

A Large number of mental and physical health research studies have employed the RADAR-Base platform for remote data collection with funding from many major funding agencies. This includes over 50 use cases exploring more than 30 disorder areas with more than 150,000 participants enrolled to date. Major disease areas using the platform are Major Depressive Disorder, Eating Disorder, Multiple Sclerosis, ADHD, Autism, Epilepsy, Atrial Fibrillation, Alzheimer's Disease, and COVID-19. Projects using the platform are collecting various health parameters depending on the disease area requirement. Data collected relates to cognition, mood, voice, digital usage, geolocation, heart rate, to name a few. Figure 6 and 7 shows examples of the status of collected data, its compliance and quality for different studies. Numerous challenges addressed by the platform include completeness of data, quality/accuracy of data, participant engagement and remote data collection.

The RADAR-base platform has effectively transformed low-level sensor data into digital biomarkers through feature generation pipelines. This process involves extracting relevant characteristics and patterns from the raw data, enabling the creation of meaningful and actionable insights. These digital biomarkers hold immense potential in various disease areas, aiding clinicians in making informed decisions, facilitating early intervention, and contributing to the prevention of relapse.

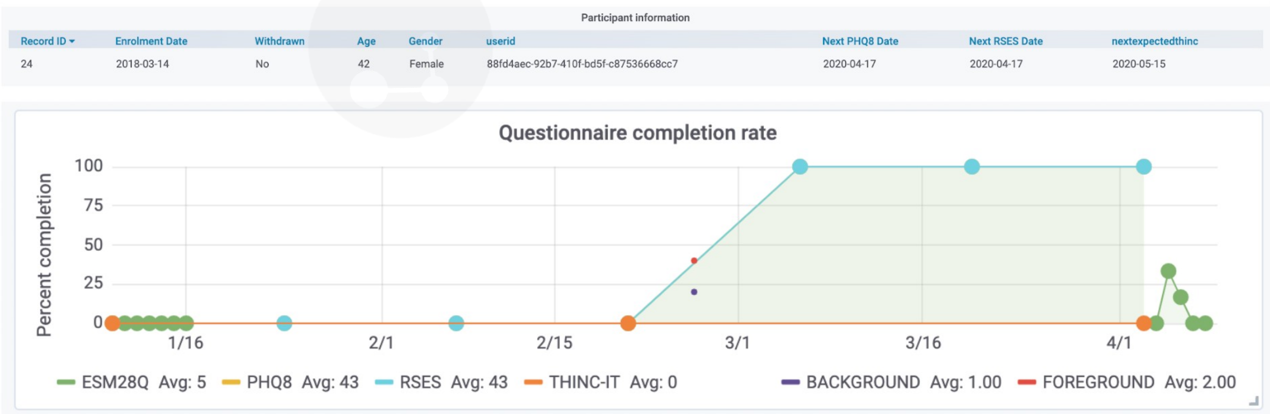

Figure 6. aRMT app questionnaire completion rate from a single patient from the RADAR-

CNS Major Depression Study.

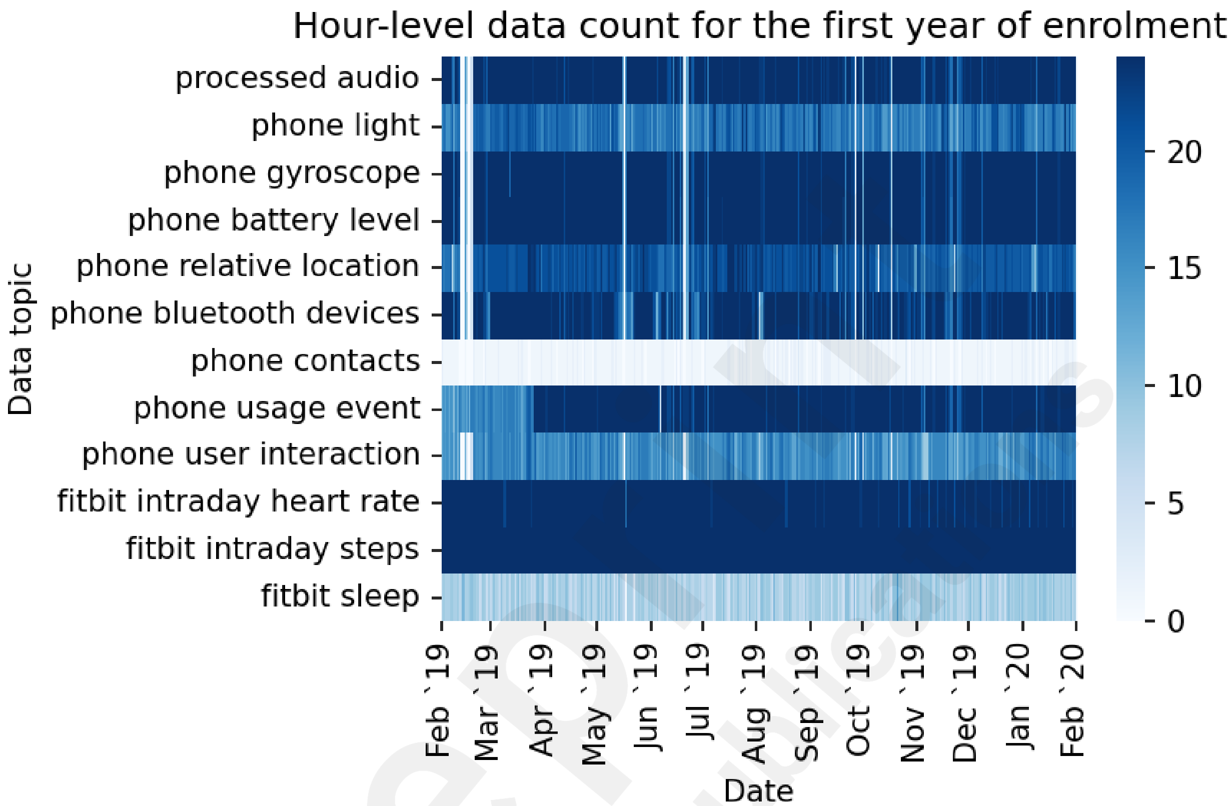

Figure 7. Contiguity of phone sensor data over the first year of enrolment collected through RADAR-base for a patient in the RADAR-CNS major depressive disorder study. The intensity of the colour represents how many hours in a day a particular modality is present.

Studies utilising RADAR-base platform for data collection and feature extraction have reported insights into data quality, participant engagement and retention [49] [50] . In the Major Depressive Disorder Study, 623 participants were enrolled, with 497 (79.8%) of them participating for the maximum study duration (11 to 24 months). Additionally, further analysis revealed that wearable datastream had the highest data available rate (407 participants with over 90% data completeness and 99 participants with over 50% completeness) across all data streams and found several indicators impacting participant engagement and retention, such as age and mental health status [50]. These findings illustrate the feasibility of remote data collection for clinical applications and provide the insights and experiences for future mobile health studies.

RADAR-base is a digital biomarker ecosystem that addresses these considerations. It provides a platform that focuses on safety and effectiveness by ensuring transparency in the algorithms used to generate biomarkers. Additionally, the interoperable components with open interfaces in RADAR-base facilitate the development of new multicomponent systems. This interoperability ensures that various digital biomarker sources can be integrated into a comprehensive and cohesive platform for healthcare purposes. Moreover, RADAR-base emphasises high integrity measurement systems, which means that the data collected and the biomarkers generated are reliable and accurate. This emphasis on data quality is essential for building trust in the digital biomarker ecosystem and encouraging its adoption in clinical research and routine patient care.

Overall, the systematic approach and emphasis on safety, effectiveness, transparency, and interoperability offered by RADAR-base can contribute to the advancement of digital biomarkers and their integration into healthcare systems, ultimately benefiting patient outcomes and medical research.

Abbreviations

PROM: Patient Reported Outcome Measures
IoT: Internet of Things
RSES: Rosenberg Self-Esteem Scale
PHQ8: Patient Health Questionnaire Depression Scale
ESM: Experience Sampling Method
SDK: Software Development Kit
JSON: JavaScript Object Notation
API: Application Programming Interface
REST: Representational State Transfer
DPO: Data Protection Officer

Acknowledgements

This study has received support from the EU/EFPIA IMI Joint Undertaking 2 (RADAR-CNS grant No 115902). This communication reflects the views of the RADAR-CNS consortium, and neither IMI nor the European Union and EFPIA are liable for any use that may be made of the information contained herein. The authors receive funding support from the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. The authors also acknowledge the support of NIHR University College London Hospitals Biomedical Research Centre.

References

1 Motahari-Nezhad, H. *et al.* (2022) Digital Biomarker-Based Studies: Scoping Review of Systematic Reviews. *JMIR Mhealth Uhealth* 10, e35722
2 Ranjan, Y. *et al.* (2018) , RADAR-base: A Novel Open Source m-Health Platform. , in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers - UbiComp '18*, New York, New York, USA, pp. 223–226
3 Ranjan, Y. *et al.* (2019) RADAR-Base: Open Source Mobile Health Platform for Collecting, Monitoring, and Analyzing Data Using Sensors, Wearables, and Mobile Devices. *JMIR Mhealth Uhealth* 7, e11734

4 Grafana: The open observability platform | Grafana Labs. . [Online]. Available: https://grafana.com/. [Accessed: 26-Jul-2023]
5 Opoku Asare, K. *et al.* (2022) Mood ratings and digital biomarkers from smartphone and wearable data differentiates and predicts depression status: A longitudinal data analysis. *Pervasive Mob. Comput.* 83, 101621
6 Jacobson, N.C. *et al.* (2020) Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors. *J. Med. Internet Res.* 22, e16875
7 Babrak, L.M. *et al.* (2019) Traditional and digital biomarkers: two worlds apart? *Digit. Biomark.* 3, 92–102
8 Bent, B. *et al.* (2020) The digital biomarker discovery pipeline: An open-source software platform for the development of digital biomarkers using mHealth and wearables data. *J. Clin. Transl. Sci.* 5, e19
9 Hossain, S.M. *et al.* (2017) mCerebrum: A Mobile Sensing Software Platform for Development and Validation of Digital Biomarkers and Interventions. *Proc. Int. Conf. Embed. Netw. Sens. Syst.* 2017,
10 Torous, J. *et al.* (2017) The new digital divide for digital biomarkers. *Digit. Biomark.* 1, 87–91
11 Intel® Context Sensing SDK | Intel® Software. . [Online]. Available: https://software.intel.com/en-us/context-sensing-sdk. [Accessed: 15-Nov-2018]
12 Mood-tracking app paves way for pocket therapy | University of Cambridge. . [Online]. Available: https://www.cam.ac.uk/research/news/mood-tracking-app-paves-way-for-pocket-therapy. [Accessed: 15-Nov-2018]
13 Medopad | Building Healthier Societies. . [Online]. Available: https://medopad.com/. [Accessed: 15-Nov-2018]
14 Eckhoff, R.P. *et al.* (2015) A platform to build mobile health apps: the personal health intervention toolkit (PHIT). *JMIR Mhealth Uhealth* 3, e46
15 Jardine, J. *et al.* (2015) Apple's ResearchKit: smart data collection for the smartphone era? *J. R. Soc. Med.* 108, 294–296
16 ResearchStack. , *ResearchStack*. [Online]. Available: http://researchstack.org/. [Accessed: 28-Aug-2018]
17 Zens, M. *et al.* (2017) Development of a modular research platform to create medical observational studies for mobile devices. *JMIR Res. Protoc.* 6, e99
18 Torous, J. *et al.* (2019) Creating a digital health smartphone app and digital phenotyping platform for mental health and diverse healthcare needs: an interdisciplinary and collaborative approach. *J. technol. behav. sci.* DOI: 10.1007/s41347-019-00095-w
19 Confluent: Apache Kafka & Streaming Platform for the Enterprise. . [Online]. Available: https://www.confluent.io/. [Accessed: 19-Jun-2018]
20 Kubernetes. . [Online]. Available: https://kubernetes.io/. [Accessed: 10-Apr-2024]
21 Coravos, A. *et al.* (2019) Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *npj Digital Med.* 2,
22 Sun, S. *et al.* (2022) The utility of wearable devices in assessing ambulatory impairments of people with multiple sclerosis in free-living conditions. *Comput. Methods Programs Biomed.* 227, 107204
23 Sun, S. *et al.* (2022) Challenges in the analysis of mobile health data from smartphones and wearable devices to predict depression symptom severity. *arXiv* DOI: 10.48550/arxiv.2212.10540

24 Zhang, Y. *et al.* (2022) Longitudinal Relationships Between Depressive Symptom Severity and Phone-Measured Mobility: Dynamic Structural Equation Modeling Study. *JMIR Ment Health* 9, e34898
25 Home · RADAR-base/radarpipeline Wiki. . [Online]. Available: https://github.com/RADAR-base/radarpipeline/wiki. [Accessed: 08-Mar-2023]
26 Matcham, F. *et al.* (2019) Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. *BMC Psychiatry* 19, 72
27 Zhang, Y. *et al.* (2021) Predicting depressive symptom severity through individuals' nearby bluetooth device count data collected by mobile phones: preliminary longitudinal study. *JMIR Mhealth Uhealth* 9, e29840
28 Sanchez, M.H. *et al.* (2011) , Using prosodic and spectral features in detecting depression in elderly males. , in *Interspeech 2011*, ISCA, pp. 3001–3004
29 Cummins, N. *et al.* (2022) Multilingual markers of depression in remotely collected speech samples. *Res. Sq.* DOI: 10.21203/rs.3.rs-2183980/v1
30 Bruno, E. *et al.* (2020) Remote assessment of disease and relapse in epilepsy: protocol for a multicenter prospective cohort study. *JMIR Res. Protoc.* 9, e21840
31 RADAR-CNS. , *RADAR-CNS*. [Online]. Available: http://radar-cns.org. [Accessed: 25-May-2018]
32 Sun, S. *et al.* (2020) Using Smartphones and Wearable Devices to Monitor Behavioral Changes During COVID-19. *J. Med. Internet Res.* 22, e19992
33 Zhang, Y. *et al.* (2021) Relationship between major depression symptom severity and sleep collected using a wristband wearable device: multicenter longitudinal observational study. *JMIR Mhealth Uhealth* 9, e24604
34 ART-CARMA - The ADHD Remote Technology study of cardiometabolic risk factors and medication adherence. . [Online]. Available: https://www.kcl.ac.uk/research/art-carma-the-adhd-remote-technology-study-of-cardiometabolic-risk-factors-and-medication-adherence. [Accessed: 21-Sep-2022]
35 ADHD Remote Technology (ART). . [Online]. Available: https://www.kcl.ac.uk/research/adhd-remote-technology-art. [Accessed: 21-Sep-2022]
36 Welcome to RADAR-AD | Radar-AD. . [Online]. Available: https://www.radar-ad.org/. [Accessed: 21-Sep-2022]
37 Muurling, M. *et al.* (2021) Remote monitoring technologies in Alzheimer's disease: design of the RADAR-AD study. *Alzheimers Res Ther* 13, 89
38 AIMS-2-TRIALS - Autism Research For Europe. . [Online]. Available: https://www.aims-2-trials.eu/. [Accessed: 21-Sep-2022]
39 BigData@Heart > Home. . [Online]. Available: https://www.bigdata-heart.eu/. [Accessed: 21-Sep-2022]
40 Convalescence long COVID Study | COVID-19 Longitudinal Health and Wellbeing National Core Study - UCL – University College London. . [Online]. Available: https://www.ucl.ac.uk/covid-19-longitudinal-health-wellbeing/convalescence-long-covid-study. [Accessed: 21-Sep-2022]
41 Homepage - Covid Collab. . [Online]. Available: https://covid-collab.org/. [Accessed: 21-Sep-2022]
42 Ranjan, Y. *et al.* (2021) Remote assessment of lung disease and impact on physical and mental health (RALPMH): protocol for a prospective observational study. *JMIR Res.*

Protoc. 10, e28873
43 RADAR-base · GitHub. . [Online]. Available: https://github.com/RADAR-base. [Accessed: 04-Jun-2018]
44 Home | RADAR-base. . [Online]. Available: https://radar-base.org. [Accessed: 04-Jun-2018]
45 Confluence. . [Online]. Available: https://radar-base.atlassian.net/wiki/spaces/RAD/pages/87359489/Data+Source+Catalog. [Accessed: 04-Jun-2018]
46 GitHub - RADAR-base/RADAR-Docker: Integrated Docker Stack for the RADAR mHealth Streaming Platform Components. . [Online]. Available: https://github.com/RADAR-base/RADAR-Docker. [Accessed: 04-Jun-2018]
47 GitHub - RADAR-base/RADAR-Kubernetes: Kubernetes deployment of RADAR-base. . [Online]. Available: https://github.com/RADAR-base/RADAR-Kubernetes. [Accessed: 26-Sep-2022]
48 RADAR-base/RADAR-aRMT-protocols: aRMT App protocols for questionnaire delivery. . [Online]. Available: https://github.com/RADAR-base/RADAR-aRMT-protocols. [Accessed: 08-Mar-2023]
49 Muurling, M. *et al.* (2024) Feasibility and usability of remote monitoring in Alzheimer's disease. *Digit Health* 10, 20552076241238132
50 Matcham, F. *et al.* (2022) Remote Assessment of Disease and Relapse in Major Depressive Disorder (RADAR-MDD): recruitment, retention, and data availability in a longitudinal remote measurement study. *BMC Psychiatry* 22, 136

Supplementary Files

Multimedia Appendixes

Exemplar RADAR-base platform DPIA.
URL: http://asset.jmir.pub/assets/7a428966b1b13298555c53e6653ccdc3.docx