

Exploring the definition, measurement, and reporting of engagement in remote measurement studies for physical and mental health symptom tracking: A systematic review

Katie M White (✉ katie.white@kcl.ac.uk)

Charlotte Williamson

Nicol Bergou

Carolin Oetzmann

Valeria de Angel

Faith Matcham

Claire Henderson

Matthew Hotopf

Systematic Review

Keywords: Remote Measurement Technologies, Engagement, Smartphone Apps, Wearables, Smartphones, Digital health, mHealth, Mobile health, Wearable devices, Health technology, Remote measurement technologies

Posted Date: May 5th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1625589/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Remote Measurement Technologies (RMTs) could revolutionise management of chronic health conditions by providing real-time symptom tracking. However, the promise of RMTs relies on user engagement, which at present is variably reported in the field. This review aimed to synthesise the RMT literature to identify how and to what extent engagement is defined, measured, and reported, and to present recommendations for the standardisation of future work.

Methods

Seven databases (Embase, MEDLINE and PsycINFO (via Ovid), PubMed, IEEE Xplore, Web of Science, and Cochrane Central Register of Controlled Trials) were searched in July 2020 for papers using RMT apps for symptom monitoring in adults with a health condition, prompting users to track at least three times during the study period. Data were synthesised using critical interpretive synthesis.

Results

A total of 76 papers met the inclusion criteria. Sixty five percent of papers did not include a definition of engagement. Thirty five percent included both a definition and measurement of engagement. Four synthetic constructs were developed for measuring engagement: i) engagement with the research protocol, ii) objective RMT engagement, iii) subjective RMT engagement, and iv) interactions between objective and subjective RMT engagement.

Discussion

The field is currently impeded by incoherent measures and a lack of consideration for engagement definitions. A process for implementing the reporting of engagement in study design is presented, alongside a framework for definition and measurement options available. Future work should consider engagement with RMTs as distinct from the wider eHealth literature, and measure objective versus subjective RMT engagement.

Registration

This review has been registered on PROSPERO [CRD42020192652].

Introduction

Remote Measurement Technologies (RMTs) provide real-time, longitudinal health tracking by collecting frequent questionnaire, physiological and behavioural data outside and between traditional clinic or research assessments¹. RMTs form a subsection of the electronic health (eHealth) movement, which utilises mobile or online technologies to improve patient outcomes². Active RMT (aRMT) measures

comprise smartphone applications (apps) for symptom reporting and can be combined with mobile or wearable sensors for passive RMT (pRMT) data collection³⁻⁵. RMTs hold great potential for the management of chronic health conditions. First, they can provide unbiased data on symptom fluctuations and clinical state. Remote symptom assessments have been validated against in-person measures across physical and mental health conditions⁶⁻⁸. By measuring symptoms remotely, it is also possible to gain more detailed longitudinal information than is possible through less frequent clinic or research assessments. Second, multi-parametric RMT data may provide the temporal resolution needed to detect indicators of future relapse or remission in long term conditions⁹. Preliminary work has suggested that remote monitoring of depression and anxiety may prove valuable in predicting future symptoms¹⁰. The use of RMTs could potentially revolutionise research, clinical practice and self-management in chronic diseases¹¹.

The promise of RMTs depends almost entirely on user engagement. Engagement can be broadly understood as a multi-stage construct indicating the extent to which a resource is actively used¹². In research, low initial engagement in RMT studies, i.e. uptake, increases the risk of selection bias, lowers statistical power, and reduces the external generalisability of results¹³. Sustained engagement, i.e. retention and ongoing adherence to research protocols, is essential to ensure that the resulting datasets in research studies are complete and therefore allow identification of patterns of symptom change¹⁴. Describing and understanding the drivers of participants' engagement with RMTs in research is necessary to determine the success of future real-world implementation of RMTs in clinical services.

The current state of engagement with RMTs is unclear. Qualitative research suggests acceptance of the use of technology for symptom monitoring in conditions such as depression¹⁵, multiple sclerosis¹⁶, epilepsy¹⁷, arthritis¹⁸ and fibromyalgia¹⁹. In practice however, reported engagement statistics are hugely heterogeneous. Dropout rates for studies of remote monitoring of various lengths in depression have been estimated at 3.6%²⁰ and 26.2%²¹ in separate reviews. A large caveat to synthesizing findings is the current lack of standardization in the measurement and reporting of participant engagement. Simblett et al.²² found a variation of idiosyncratic, non-comparable engagement measures was reported across 33 studies, which included dropout, adherence rates, and usage statistics. This heterogeneity was noted to have 'severely limited quantified conclusions'. Two recent systematic reviews on using RMT data to monitor symptom change cited methodological differences and inadequate missing data reporting as key barriers to performing meta-analyses^{23,24}. There is a clear need to understand how engagement is reported in the current RMT literature, in order to promote reproducibility of results and ensure that future meta-analyses can compare like with like.

A body of work in the wider eHealth field has made significant progress in the standardization of engagement reporting. The publication of guidelines such as CONSORT-EHEALTH²⁵ and STROBE²⁶ has considerably improved reproducibility, with regards to the reporting of recruitment, adherence and attrition. Systematic reviews of eHealth studies have further explored how authors report on engagement. Sieverink et al.²⁷ recommended that the inclusion of a 'definition of intended engagement, justification

with theory or rationale, and corresponding measurement' was crucial for comparing engagement across technologies²⁷. Perski et al.²⁸ conducted a systematic review using critical interpretive synthesis to further explore how engagement with digital behaviour change interventions (DBCIs) had been defined and measured. They proposed an integrative definition, comprising *objective* (the extent of usage, e.g. adherence) and *subjective* (experiential factors, e.g. attention, interest) engagement concepts. These subjective and objective components were underpinned by corresponding measurements. This framework has since been used to synthesise the reporting of engagement in eHealth interventions for depression²⁹. Thus, an exploration of varying definitions and measurements of engagement has been key to standardizing the findings of the wider eHealth literature. It is currently unclear how far these findings apply to RMT work specifically.

This systematic review aims to explore the current state of engagement reporting in studies using RMTs to track symptoms in physical and mental health conditions. Our broad aim is to describe the extent to which engagement is defined, measured, and reported in the literature. There are three main objectives:

1. To present a quantitative synthesis of the proportion of studies which report on any form of engagement definition or measurement.
2. To synthesise studies reporting on engagement with regards to the following questions:
 - (a) How has engagement been defined in the selected literature?
 - (b) How has engagement been measured in the selected literature?
3. To present recommendations for the standardisation of future work in this field

Our findings will aid in understanding the current trends in the field, and promote standardization for future work, syntheses, and evaluation of RMT research.

Results

1. Summary of search results

The electronic database search yielded 6,929 articles, reduced to 4,772 after removing duplicates across databases. Of the full texts screened, 76 met the inclusion criteria and were included in the data synthesis (Fig. 1). Sixty studies were described across the 76 papers.

[Insert Fig. 1]

Characteristics of the included papers are described in Electronic Supplementary Material 2. The most common conditions of study were bipolar disorder (N = 23), depression and/or anxiety disorder (N = 11), chronic pain (N = 8), cancer (N = 5), and psychosis (N = 4). Forty one papers (54%) reported on RMT use in a research context, 22 (29%) on RMTs for self-management and 13 (17%) on the use of RMTs for clinical

support. aRMT only was used in 45 (59%) papers, where 31 (45%) papers used a combination of aRMT and pRMT methods. Symptoms were tracked for a median of 56 days (range = 794 days). Common symptoms tracked included mood, pain, sleep, fatigue, exercise, and daily activity levels.

2. To what extent has engagement been reported in the literature?

Of the 76 papers included, a large majority (n = 74, 97%) reported on engagement in some form. The two papers that did not report on engagement^{30,31} are secondary analyses and cite previous feasibility papers. Of the 74 papers that did report on engagement, 48 (65%) did not include a definition of engagement. Twenty-six (35%) papers included both a definition and corresponding measurement of engagement. All but one of the papers that included a definition and measurement noted engagement as a main aim of the paper; Greer et al.³² aimed to test the effect of a smartphone app, of which symptom monitoring was one component, on adherence to cancer treatment. Of the 30 papers (39.5%) that noted engagement as a main aim, 5 (17%) did not include a definition of engagement.

3. How has engagement been defined and measured in the literature?

We identified five synthetic constructs for engagement definition: i) objective definition, e.g. feasibility, ii) subjective definition, e.g. usability, iii) utility, iv) app usage, and v) engagement with the mobile research protocol. Four synthetic constructs were developed for engagement measurement: i) engagement with the research protocol, ii) objective RMT engagement, iii) subjective RMT engagement, and iv) interactions between objective and subjective RMT engagement. Table 1 describes the number of papers that included each of the definition and measurement constructs.

[Insert Table 1]

What follows comprises a narrative synthesis of each synthetic construct for the definition and measurement of engagement across the included papers.

3.1 How has engagement been defined in the literature?

Papers used a range of, often interchangeable, definitions for engagement. The most cited term was 'feasibility'. For many studies, feasibility was conceptualised as objective compliance with, or adherence to, symptom monitoring³³⁻³⁷, such that higher completion rates signalled greater feasibility. Depending on the context of RMT use, feasibility either represented i) the ability of users to use an app to complete the monitoring, or ii) the ability of RMTs to collect sufficient data for assessment of the cohort in question. Terms such as 'usability', 'acceptability' and 'satisfaction' were often used in addition to feasibility. These tended to be more subjective concepts, often paired with rating scales and questionnaires to understand how far users felt the technologies were appropriate for use. Two studies further understood engagement as 'perceived utility' or 'usefulness' of the RMT for disease management^{38,39}. Other studies defined app usage data as a proxy for engagement.³⁸ sought to understand 'patient engagement' on the basis of the frequency and quality of interaction (active vs. passive) with components of the self-monitoring system, and⁴⁰ identified patient preference for RMT type by usage

data. One paper defined engagement as “participant willingness to engage in mobile research protocols”⁴¹ (pg. 14). Overall, a large proportion of studies used multiple conceptualisations of, and terms for, engagement in the same paper, often with little justification for why they had been chosen or how they complemented each other.

3.2 How has engagement been measured and reported in the literature?

There was a high level of heterogeneity in the measurement of engagement across the 74 studies. Overall, four synthetic constructs were identified: i) engagement with the research protocol, ii) objective RMT engagement, iii) subjective RMT engagement, and iv) interactions between objective and subjective RMT engagement.

3.2.1 Engagement with the research protocol

A large majority of papers (88%, N = 65) reported on engagement in reference to recruitment, retention and withdrawal from RMT studies. Where generic reasons for declining research participation, such as lack of time and/or lack of interest, were widely reported, some studies also reported on RMT-specific reasons: unwilling to use/switch to trial smartphone⁴²⁴³³²⁴⁴⁴⁵, concerns about remote data collection⁴²⁴⁵⁴⁶, having ‘no need’ to monitor symptoms alongside treatment⁴⁷. One study⁴⁸ calculated the percentage of participants who declined to participate explicitly based on the RMT nature of the research, based on previous non-response to clinic and research opportunities.

Reasons for participant dropout followed a similar vein. Technological issues, for example an inability to download the study app⁴⁹, phone updates³³ and software malfunctions with study apps^{33 3637}, were the most reported reasons for non-completion of a study. However, many papers did not differentiate these from non-RMT related reasons for dropout, e.g. death or hospitalization, or failed to provide reasons at all. Seven studies^{2,35,41, 49–52} investigated whether dropout was associated with sociodemographic or clinical variables, and one further explored the association between dropout and daily access to data, as a main aim of the study⁵⁰. Two studies reported on likelihood of dropout over time, finding opposing results^{35,53}. Understanding the factors associated with decline and dropout over time provided insights in gauging interest in RMTs for support alongside clinical care⁴⁷⁴⁸.

There was little consensus over what constituted researcher-initiated withdrawal from the study. Again, there was a lack of distinction made between RMT and non-RMT related factors. Some studies reported on the percentage of participants completing the full number of days of remote monitoring^{34,54,55}, where others focused on completion of a sufficient, often arbitrary, number of daily assessments. Participants were withdrawn or excluded from analyses if they did not submit any daily assessments^{33,37,56} or never used the app³⁶. Two papers^{51,57} defined withdrawal based on failure to complete a minimum amount of ESM assessments during the study period: 33% and 30% respectively. A failure to attend appointments for additional, in-person questionnaires was also a reason for withdrawal^{36,43,58}. Retention rates were largely discussed in terms of the resulting impact on the final dataset for analyses. Overall, a lack of

standardisation in the measurement of engagement with the research protocol, coupled with the merging of technological and non-technological factors, suggests that attempts to provide accurate information on dropout and retention rates in RMT studies would be difficult.

3.2.2 Objective engagement with RMTs

Alongside engagement with the research protocol, studies also reported engagement with the RMTs themselves. 87% (N = 64) of papers conceptualised engagement with RMTs as an objective measure, e.g. direct behaviour whilst using the device. First, studies reported on the total, raw number of assessments sent and responded to, or the number of symptoms tracked, during the study period in the sample as a whole. This was further broken down into 'per participant' measures: the mean or median average number of daily aRMT assessments completed across the study period^{32,33,49,53,54,59,60}, or the average number of assessments completed within each day⁶¹⁻⁶³, depending on the frequency of the aRMT tasks. Some studies also reported the number of days in which participants provided self-monitoring data^{37,46,53,54}, though at times it was unclear whether this referenced only the days where data was recorded, or the total number of days during which the participant remained in the study.

Second, symptom tracking compliance rates were reported. There was little standardisation in how these were measured. 'Compliance rate' was generally conceptualised as the percentage of aRMT assessments completed over the total available to be completed and was used interchangeably with the terms 'response rate', 'adherence', and 'completion'. Many studies simply reported the average percentage compliance, often with accompanying measures of variance, across the sample for aRMT assessments and associated clinical measures. Some instead reported on the percentage of days in which self-monitoring was adhered to⁶⁴⁻⁶⁶ (assuming that self-monitoring was assessed daily). Some also measured compliance against set criteria, for example the percentage of participants who completed all assessments^{42, 67-69}, or percentage completion at longitudinal time points;³³ reported the percentage of participants who completed at least 30, 60 and 90 daily assessments respectively. Four studies assessed the variation in compliance across time of day⁶⁹⁻⁷². Of the 26 studies which included passive tracking, five reported on some form of pRMT compliance^{33,40,56,73,74}. These included the percentage of participants logging daily Fitbit data and the percentage of days in which 'complete' actigraphy/activity data was logged. One study defined multi-parametric adherence as the proportion of study days with at least 50% of daily self-report questionnaires completed and 12 hours of logged activity tracking, chosen to align with future research and clinical goals⁴⁰. Crucially however, most studies gave little or no justification for their measurement of compliance, or the conclusions that were drawn from it.

Third, app usage statistics were used when discussing engagement. For studies which required participants to remotely enrol, percentage of successful study app downloads was a fundamental indicator of engagement^{33,49,50}. Other reported usage statistics throughout the study period included total app use (per day, per week or per the study period)^{32,42,52,71,75,76} and number of times the app was launched by the user, be that self-initiated⁷⁵ or in response to prompts^{36,60}. One multi-parametric study

directly compared the number of days spent engaging with the app with that of the Fitbit⁴⁰. Studies which included RMT symptom tracking as a component of a behaviour change app also reported on in-app module viewing^{38,47,77,78}. The impact of app usage was considered by three studies: i) minutes and days of app use accounted for a large percentage of variance in an 'app engagement factor'³², ii) viewing in-app symptom visualisations correlated with aRMT and pRMT adherence⁴⁰, and iii) longitudinal app use was considered to reflect 'satisfaction and interest'⁷⁶.

3.2.3 Subjective engagement with RMTs

An exploration of subjective engagement, e.g. indirect user experience during or after interacting with the device, was included as an adjunct to objective engagement in 27 (37%) papers. Many studies used quantitative measures, usually in the form of Likert scales, to measure participant experiences of engaging with remote self-monitoring. These were administered at study end, apart from one study⁷² which issued a usability scale after week one in order to assess early technological problems. Four studies also assessed corresponding clinicians^{33,44,68,79}. Scales included a wide range of questions surrounding two main themes: i) usability of, or satisfaction with, the technology itself, and ii) utility of the technology for symptom management. Technology usability assessed ease of use, helpfulness of reminders and navigation of the app interface. Utility of the technology focused more on the use of the technology for symptom management and communication with clinical care teams, and intentions for future use. Most studies included questions covering both themes to some extent, however crucially these were generally assessed as a combined 'satisfaction' variable (apart from⁴² which used two distinct scales to measure technology obtrusiveness and clinical utility respectively). Studies either combined several, validated scales in one paper, including the System Usability Scale (SUS)^{36,69,75,80} and Technology Assessment Model (TAM)^{75,81}, or used author-developed scales, due to a lack of 'psychometrically evaluated measures of user acceptance' of RMTs⁸¹.

Qualitative measures of subjective engagement also explored participants' experiences. Five studies conducted semi-structured interviews with participants^{34,40,42,78,82}, and one with care managers⁴². Other studies reported more vaguely on 'evaluative feedback'^{49,65,83}, 'response to open-ended questions'^{36,39,84} or 'in-app free text'⁷⁸ collected throughout follow-up. Qualitative methods were nearly always collected in conjunction with quantitative usability scales, or as a section of the usability scale itself, however one study⁸⁰ recorded questions and difficulties reported by participants to the researchers during the initial demonstration of the technology. Emerging themes covered overviews of experience, challenges and benefits of RMTs for symptom management, and suggestions for improvement.

3.2.4 Interaction between objective and subjective engagement with RMTs

Most studies reported on objective and subjective engagement either in isolation, or as one, combined 'engagement' variable. However, 22 papers (30%) explored the link between the objective and subjective measures used.

Some studies reported on the resulting, subjective effects of using RMTs for symptom monitoring in the study (either as a main aim or an additional outcome). Where the technology took the form of an intervention, quantitative analyses explored associations between RMT use and changes in main outcome variables, including symptom severity^{32,33,37,43,49,62,76,85}, physical activity^{33,37,49}, receipt of treatment^{32,86}, and medication uptake³². Other studies reported on the impacts of remote self-monitoring from a more exploratory standpoint. The following qualitative themes were suggested: identification, self-awareness and mindfulness of symptoms and/or emotional health^{34,36,40,42}, adapting self-management strategies^{42,65}, being seen 'as a person'⁷⁸, having access to a 'safety net' of professional support^{51,78}. Only one study suggested a perceived negative effect of RMT use; viewing real-time mood forecasting might lead to a self-fulfilling prophecy, worsening mood⁸⁷.

Five studies conducted correlational analyses on the relationship between objective and subjective variables^{33,37,49,53,84}. Participants who reported 'liking' an RMT app, or a higher degree of satisfaction, tended to submit more daily assessments⁸⁴ and have a higher duration of compliance (consecutive days with recorded data)⁵³. Two studies from the same author both found significant Pearson product moment correlations between higher daily assessment count and app satisfaction, across both a 3-month and 6-month time point^{33,37}. A consideration of the interaction between both forms of engagement was considered essential for encouraging future use of RMTs for remote medical and psychological assistance³⁷.

3.3 An integrative framework for the reporting of engagement

The synthetic constructs for measuring engagement were combined to develop an overarching synthetic argument: an integrative framework for measuring engagement with RMTs. This is depicted in Fig. 2. Definitions were considered too heterogenous to include in the framework, and instead recommendations are included within the discussion section of this review.

[Insert Fig. 2]

Discussion

Principal Findings

Understanding how engagement is reported in the RMT literature is imperative in order to ensure reproducibility in the field, allow studies to build upon results of previous work, and to implement findings in real world settings. The first and second aims of this review were to explore the current state of defining and measuring engagement by using both quantitative and CIS data synthesis methods. The majority of papers reported on engagement in some form. However, these represented a large range of incoherent and often unjustified measures. Many studies employed several measures in one paper, resulting in a lack of distinction between measure types, explanation of why some were chosen over others, and understanding of what conclusions might be drawn from the engagement findings. A much

lower proportion of studies included a corresponding definition of engagement, including many of those which described engagement as a main aim of the study. Where engagement was defined, concepts and phrases were interchangeable across papers. Indeed, even across papers with similar engagement definitions, such as 'feasibility', cut-offs used to measure RMT compliance differed hugely. Thus, though there is potential for the field to evaluate engagement, a lack of standardised reporting is impeding progress.

The third aim of this review was to present recommendations for the standardization of future work. The first step towards this has been to provide clarity on the engagement measures that are currently used in the literature. The integrative framework depicted in Fig. 2 is split into two core themes of engagement measures: engagement with the research protocol and engagement with RMTs themselves. This distinction is important given that many studies did not differentiate between these concepts when reporting on engagement; using correlates of dropout from RMT studies as a proxy for gauging wider interest in RMT implementation is only relevant if the reasons for dropout were specific to the RMT aspects of the study. The engagement with RMTs section is further split into objective and subjective measurements. Interestingly, a wide range of measures were used here, but very few studies acknowledged a distinction between objective and subjective engagement or the possible interactions between the two. Under each section of the framework lies a series of options for measuring each engagement type. This framework should aid in the classification and selection of measurements when reporting on an RMT study.

The second step towards standardising future work lies in establishing 'best practice' guidelines for reporting on engagement. This review has highlighted several drawbacks in the current literature: a dearth of clearly conceptualised engagement definitions and corresponding measures, and, as a result, an inability to make concrete conclusions on the extent of engagement in the field. Figure 3 depicts how the process of assessing engagement with RMTs should begin as early as study development. Authors designing RMT studies are encouraged to explore the reasons for examining engagement, e.g. understanding the feasibility of using a symptom tracking app, and why it is important to know this, e.g. in order to understand the extent of data collection that could be expected with clinical implementation, or to inform approaches to missing data. Authors should then pre-define a definition, measurement(s) and applicable cut-offs to be used in answering this question. Such data could be reported in the main paper or a subsequent engagement paper. Following these guidelines could provide the necessary foundations for reproducible results in the field.

[Insert Fig. 3]

Links with previous work

To our knowledge, this is the first review that has explored the extent of engagement reporting across the entirety of the RMT literature. This work complements and expands on previous reviews that have found heterogeneity in reporting of RMT studies²²⁻²⁴. It also might provide explanation as to why previous

studies have found a wide range of dropout rates^{20,21}. A lack of provision of engagement definitions in this work reflects the findings of Sieverink and colleagues in the eHealth literature²⁷; a small minority of studies included an operationalization of the engagement measure used, making comparisons difficult at best and futile at worst.

A key link to previous work is the parallel between the objective and subjective engagement concepts highlighted in Perski and colleagues' work with DBCIs²⁸. Both the definition and measurement synthetic constructs found in this review loosely map on to this distinction, with two exceptions. First, subjective engagement with RMTs is largely focused on the usability or utility of the technology, or on the subjective effects of tracking symptoms, e.g. self-awareness of symptoms, feeling monitored by a 'safety net'. This is in comparison to the typical measures of 'flow' or immersion that are seen as a mechanism of action towards behavior change in DBCIs. Second, engagement with the research protocol is a novel finding that has not been acknowledged in the DBCI literature. This is likely owing to the simultaneous use of RMTs as both a method for data collection and a tool for symptom self-management. Such differences warrant the exploration of the RMT field in its own right, as separate from DBCI or general eHealth literature.

Strengths and limitations

This review was deliberately extensive in nature, including papers spanning physical and mental health conditions and aRMT or multiparametric RMT measures. It did not exclude papers for not reporting on engagement. This allowed for an overview of the scope of reporting across the field, in contrast to previous reviews which pre-defined engagement for the purpose of inclusion^{20,22,27,88}. However, one limitation of such a broad focus is the lack of sub-typing by condition, journal type, RMT type or RMT purpose. For example, objective RMT engagement might be more important to consider in studies using RMTs for research purposes, whereas subjective RMT engagement might be more insightful for RMTs implemented into clinical practice. A specific focus on pRMT studies might also uncover additional measures of objective and subjective engagement. A second limitation is the decision to include secondary analyses papers in data synthesis. Six of the included papers reviewed analyses from the MONARCA I trial^{43,45,46,51,58,89}, and a subsequent 5 on the MONARCA II trial^{64,66,85,87,90}. This was justified given the tendency to report on feasibility in standalone papers, however, may have resulted in an over-representation of definitions or measures chosen by these authors. It might also be the case that papers reporting on the same dataset publish multiple analysis papers but only one engagement paper, resulting in a skew towards measures of engagement with the research protocol in these findings. It should also be considered that the searches for this review were undertaken in July 2020. However, the authors have no reason to believe that engagement reporting standards have radically changed between this date and publication of this manuscript.

Implications for future work and conclusions

This review provides the foundation for future RMT studies to define, measure and report on engagement in a standardized and reproducible way. Authors should use both the engagement reporting process

(Fig. 3) and the integrative framework (Fig. 2) as a basis for assessing engagement. At the same time, further work should be undertaken to understand how engagement reporting might differ by condition, journal type, RMT type and RMT purpose. Such findings will build on the proposed framework, making it applicable to a wider range of RMT work as the field continues to grow.

This review suggests that where authors are generally interested in measuring engagement with RMTs, the primary emphasis is placed on engaging with the research protocol or objectively with the RMTs themselves. This is logical for two reasons. First, the CONSORT-EHEALTH guidelines²⁵ focus mainly on reporting of engagement with the research protocol, e.g. participant attrition, and technology process outcomes, e.g. metrics of use. Second, these types of engagement directly impact on missing data and resulting statistical analyses. However, a deeper exploration of subjective engagement with RMTs might aid further understanding. A handful of papers in this review have already examined the effects of monitoring symptoms remotely on both clinical outcome variables and feelings of self-awareness and safety. A few have also begun to acknowledge the interaction between RMT usability and future use. Future work should focus on uncovering the links between objective and subjective RMT engagement, which could have huge implications for understanding how and why users engage with these technologies for research, clinical practice, or self-management.

To conclude, the current review provides an exploration of the current state of engagement reporting in studies which use RMTs for symptom tracking in physical or mental health conditions. Where there is clearly interest, the growth of the field is currently impeded by a lack of engagement definitions, incoherent measures, and an inability to compare findings. Recommendations for the standardization of reporting guidelines and the integration of engagement definitions and measures into the study design process have been put forward. In extending existing reporting guidelines, engagement with RMTs should also be considered as distinct from general eHealth interventions.

Future work should aim to contribute to the ongoing framework that has been provided, moving towards a unified understanding of the impact of engagement in this field, and what can be done to promote it.

Methods

Design

Following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines⁹¹, we report on the development of the search strategy, exclusion criteria, and selection and data extraction process for this systematic review. Data synthesis was informed by critical interpretive synthesis (CIS)⁹². CIS is an analysis technique that combines both quantitative and qualitative data, and promotes an inductive approach to the development of a theoretical framework from the available literature⁹³. There is currently a lack of understanding of reporting of engagement in RMTs, thus CIS was deemed an appropriate method. CIS has previously been used in reviews of engagement with eHealth^{28,92}. This review has been registered on PROSPERO [CRD42020192652].

Information sources & search strategy

A systematic search of the following seven databases was conducted in July 2020: Embase, MEDLINE and PsycINFO (all via Ovid), PubMed, IEEE Xplore, Web of Science Core Collection (comprising Science Citation Index Expanded and Social Sciences Citation Index), and Cochrane Central Register of Controlled Trials (CENTRAL). No date limits were imposed on the searches. Search terms included a combination of synonyms of i) symptom monitoring (e.g., self-monitor*), ii) remote measurement technology (e.g., smartphone app*), iii) health disorder (e.g., chronic disease*), and iv) core symptomatology across physical and mental health conditions (mood, depress*, pain, fatigue, mobility) using the Boolean operator 'AND'. Terms were searched for as free text in the title, abstract, or full text if available, using subject headings where appropriate (see Electronic Supplementary Material 1).

Eligibility criteria

Papers were included if they met the following criteria: 1) a peer-reviewed, full-text publication that observes the use of aRMT or aRMT and pRMT for the purpose of symptom monitoring for a physical or mental health condition, 2) participants of 18 + years recruited on the basis of having a physical or mental health condition (clinical cohorts with diagnosed disorders or non-clinical samples with validated measures), and 3) an RMT smartphone app that prompts users to track symptoms at least three times during the study period (in order to capture at least one time point in between enrolment and study end). Studies with or without comparison groups were included. Papers were excluded for the following reasons:

- Not written in English, Spanish, or German (languages spoken by co-authors);
- A sole focus on *pre-implementation* of RMT use, e.g. purely qualitative user-experience workshops;
- Description of an RMT system without actual observed data;
- General population cohorts, e.g. fitness tracking, wellbeing studies in non-clinical samples;
- RMT properties: use of a generic symptom monitoring app not developed specifically for the research on which the paper reports, e.g. PsyMate™⁹⁴, only pRMT;
- Symptom tracking properties: data inputted through a web-link or SMS, non-prompted, user-initiated symptom monitoring only.

Papers reporting secondary analyses were included, given that they have the potential to report on engagement information that might not have been presented in the initial paper²⁹. General wellbeing or fitness studies were excluded as users might have different motivations to engage than if the symptom tracking was linked with specific disease management. Generic apps were excluded on the assumption that they might carry pre-defined engagement metrics. Engagement reporting was not included in the search strategy or eligibility criteria in order to evaluate the consistency of reporting across the literature²⁹. No papers were excluded on the basis of quality, in order to describe the overall state of the research field⁹³.

Study selection process

Papers identified through the search process were merged using EndNote 20⁹⁵. Duplicate records were removed. Two pairs of reviewers independently completed the title and abstract screening stage, whereby each record was screened by at least two reviewers. Two reviewers (KW and CW) completed the full text screening stage. Any disagreements at each stage were resolved through discussion or by consulting a third reviewer.

Data extraction and management

A data extraction form was developed and refined, informed by recent systematic reviews on engagement in eHealth^{28,29} and a preliminary review of included papers. Due to the large volume of eligible studies, three team members performed data extraction; KW independently checked the extraction data for accuracy. For each paper, data were extracted on i) study characteristics, ii) symptom monitoring, and iii) engagement reporting. 'RMT use' reflected the main aim of the RMT and was coded as 1) self-management (used to manage a condition by the individual), 2) clinical support (used in conjunction with clinical care), or 3) research (asked to use as part of a research study only). Engagement reporting was split into definition, measurement and reporting sections. For the purpose of data extraction, 'engagement' was loosely defined as any definition or measure pertaining to how or why participants interacted with any element of the study. No pre-determined codes were enforced onto the engagement data, in keeping with the inductive approach of CIS.

Quality appraisal

No papers were excluded on the basis of quality. Following the principles of CIS, the process of data synthesis itself represents a critique of the literature and interpretations of credibility or methodological standards⁹². The data synthesis in this review therefore treats the literature as an 'object of scrutiny in its own right'⁹², acting as a form of quality appraisal.

Synthesis method

The ultimate goal of CIS is the formulation of one or more *synthesising arguments* that integrate evidence from *synthetic constructs* in the literature into a coherent theoretical framework⁹². A detailed process for conducting CIS can be found in⁹² and²⁸. We took the following steps in synthesising the qualitative constructs of engagement definition and measurement:

1. Text identified in the 'engagement reporting' stage of data extraction for each individual paper was coded using nVivo software⁹⁶. The specific research aims were used to guide the coding frame.
2. Codes were reviewed to develop *synthetic constructs*, whereby similar codes were grouped together to form themes.
3. Each paper was retrospectively categorised on the presence of each synthetic construct. This gave the opportunity to present a complementary quantitative overview of the data.

4. An overarching *synthesising argument* was developed (i.e., an integrative framework for reporting on engagement with RMTs) by combining all synthetic constructs.
5. The synthesising argument was refined through discussion between co-authors.

Declarations

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

This study represents independent research funded by the National Institute for Health Research (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.

Competing interests

MH is principal investigator of the RADAR-CNS programme, a precompetitive public–private partnership funded by the Innovative Medicines Initiative and European Federation of Pharmaceutical Industries and Associations. The programme receives support from Janssen, Biogen, MSD, UCB and Lundbeck.

Author contributions

KMW, MH and CH conceived and designed the review. KMW generated the search strategy and searched the databases. Studies were screened for eligibility by KMW, CW, NB and CO. KMW, CW and VdA carried out data extraction. KMW wrote the first draft and all authors contributed to subsequent edits of the manuscript. All authors approved the final manuscript.

References

1. Matcham, F. *et al.* Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): A multi-centre prospective cohort study protocol 11 Medical and Health Sciences 1103 Clinical Sciences 11 Medical and Health Sciences 1117 Public Health and Health Services. *BMC Psychiatry* **19**, 1–11 (2019).
2. Naslund, J. A., Marsch, L. A., McHugo, G. J. & Bartels, S. J. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. *J. Ment. Health* **24**, 321–332

- (2015).
3. Yuezhou *et al.* Predicting Depressive Symptom Severity Through Individuals' Nearby Bluetooth Device Count Data Collected by Mobile Phones: Preliminary Longitudinal Study. *JMIR Mhealth Uhealth* 2021;9(7)e29840 <https://mhealth.jmir.org/2021/7/e29840> **9**, e29840 (2021).
 4. Zhang, Y. *et al.* Relationship Between Major Depression Symptom Severity and Sleep Collected Using a Wristband Wearable Device: Multicenter Longitudinal Observational Study. *JMIR Mhealth Uhealth* **9**, e24604 (2021).
 5. Sun, S. *et al.* Using Smartphones and Wearable Devices to Monitor Behavioral Changes During COVID-19. *J Med Internet Res* **22**, e19992 (2020).
 6. Myers, D. R., Weiss, A., Rollins, M. R. & Lam, W. A. Towards remote assessment and screening of acute abdominal pain using only a smartphone with native accelerometers. *Sci. Rep.* **7**, 1–12 (2017).
 7. Burchert, S., Kerber, A., Zimmermann, J. & Knaevelsrud, C. Screening accuracy of a 14-day smartphone ambulatory assessment of depression symptoms and mood dynamics in a general population sample: Comparison with the PHQ-9 depression screening. *PLoS One* **16**, 1–25 (2021).
 8. Pombo, N., Garcia, N., Bousson, K., Spinsante, S. & Chorbev, I. Pain Assessment– Can it be Done with a Computerised System? A Systematic Review and Meta-Analysis. *Int. J. Environ. Res. Public Health* **13**, 415 (2016).
 9. Jones, M. & Johnston, D. Understanding phenomena in the real world: The case for real time data collection in health services research. *Journal of Health Services Research and Policy* vol. 16 172–176 (2011).
 10. Moshe, I. *et al.* Predicting Symptoms of Depression and Anxiety Using Smartphone and Wearable Data. *Front. Psychiatry* **12**, 1–12 (2021).
 11. Liao, Y., Thompson, C., Peterson, S., Mandrola, J. & Beg, M. S. The Future of Wearable Technologies and Remote Monitoring in Health Care. *Am. Soc. Clin. Oncol. Educ. book. Am. Soc. Clin. Oncol. Annu. Meet.* **39**, 115–121 (2019).
 12. O'Brien, H.L, Toms, E. G. What is User Engagement? A Conceptual Framework for Defining User Engagement with Technology. *J. Am. Soc. Inf. Sci. Technol.* **59**, 938–955 (2008).
 13. Teague, S. *et al.* Retention strategies in longitudinal cohort studies: A systematic review and meta-analysis. *BMC Med. Res. Methodol.* **18**, 1–22 (2018).
 14. Druce, K. L., Dixon, W. G. & McBeth, J. Maximizing Engagement in Mobile Health Studies: Lessons Learned and Future Directions. *Rheumatic Disease Clinics of North America* vol. 45 159–172 (2019).
 15. Simblett, S. *et al.* Barriers to and Facilitators of Engagement With mHealth Technology for Remote Measurement and Management of Depression: Qualitative Analysis. *JMIR mHealth uHealth* **7**, e11325 (2019).
 16. Simblett, S. K. *et al.* Engaging across dimensions of diversity: A cross-national perspective on mHealth tools for managing relapsing remitting and progressive multiple sclerosis. *Mult. Scler. Relat. Disord.* **32**, 123–132 (2019).

17. Simblett, S. K. *et al.* Patient perspectives on the acceptability of mHealth technology for remote measurement and management of epilepsy: A qualitative analysis. *Epilepsy Behav.* **97**, 123–129 (2019).
18. White, K. M. *et al.* Remote Measurement in Rheumatoid Arthritis: Qualitative Analysis of Patient Perspectives. *JMIR Form. Res.* **5**, e22473 (2021).
19. Vanderboom, C. E., Vincent, A., Luedtke, C. A., Rhudy, L. M. & Bowles, K. H. Feasibility of interactive technology for symptom monitoring in patients with fibromyalgia. *Pain Management Nursing* vol. 15 557–564 (2014).
20. Girolamo, G. *et al.* The acceptability of real-time health monitoring among community participants with depression: A systematic review and meta-analysis of the literature. *Depress. Anxiety* da.23023 (2020) doi:10.1002/da.23023.
21. Torous, J., Lipschitz, J., Ng, M. & Firth, J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. *J. Affect. Disord.* **263**, 413–419 (2020).
22. Simblett, S. *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.* **20**, (2018).
23. Dogan, E., Sander, C., Wagner, X., Hegerl, U. & Kohls, E. Smartphone-based monitoring of objective and subjective data in affective disorders: Where are we and where are we going? Systematic review. *J. Med. Internet Res.* **19**, (2017).
24. de Angel, V. *et al.* Digital health tools for the passive monitoring of depression: a systematic review of methodsTitle. *npj Digit. Med.* **5**, (2022).
25. Eysenbach, G. & Group, C. CONSORT - EHEALTH checklist (V. 1. 6. 1): Information to include when reporting ehealth / mhealth trials (web - based / Internet - based intervention and decision aids, but also social media, serious games, DVDs, mobile applications, certain te. 1–13 (2011).
26. von Elm, E. *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J. Clin. Epidemiol.* **61**, 344–349 (2008).
27. Sieverink, F., Kelders, S. M. & Gemert-Pijnen, V. Clarifying the concept of adherence to ehealth technology: Systematic review on when usage becomes adherence. *J. Med. Internet Res.* **19**, 1–15 (2017).
28. Perski, O., Blandford, A., West, R. & Michie, S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl. Behav. Med.* **7**, 254–267 (2017).
29. Molloy, A. & Anderson, P. L. Engagement with mobile health interventions for depression: A systematic review. *Internet Interv.* **26**, 100454 (2021).
30. Li, H. *et al.* Use of ecological momentary assessment to detect variability in mood, sleep and stress in bipolar disorder. *BMC Res. Notes* **12**, 791 (2019).

31. Juengst, S. B., Terhorst, L., Kew, C. L. & Wagner, A. K. Variability in daily self-reported emotional symptoms and fatigue measured over eight weeks in community dwelling individuals with traumatic brain injury. *Brain Inj.* **33**, 567–573 (2019).
32. Greer, J. A. *et al.* Randomized Trial of a Smartphone Mobile App to Improve Symptoms and Adherence to Oral Therapy for Cancer. *JNCCN J. Natl. Compr. Cancer Netw.* **18**, 143–150 (2020).
33. Jamison, R. N., Jurcik, D. C., Edwards, R. R., Huang, C.-C. C. & Ross, E. L. A Pilot Comparison of a Smartphone App with or Without 2-Way Messaging among Chronic Pain Patients: Who Benefits from a Pain App? *Clin. J. Pain* **33**, 676–686 (2017).
34. Schwartz, S., Schultz, S., Reider, A. & Saunders, E. F. H. Daily mood monitoring of symptoms using smartphones in bipolar disorder: A pilot study assessing the feasibility of ecological momentary assessment. *J. Affect. Disord.* **191**, 88–93 (2016).
35. Bove, R. *et al.* Evaluating more naturalistic outcome measures. *Neurol. Neuroimmunol. NeuroInflammation* **2**, (2015).
36. Zia, J. *et al.* Feasibility and Usability Pilot Study of a Novel Irritable Bowel Syndrome Food and Gastrointestinal Symptom Journal Smartphone App. *Clin. Transl. Gastroenterol.* **7**, e147 (2016).
37. Jamison, R. N., Mei, A. & Ross, E. L. Longitudinal trial of a smartphone pain application for chronic pain patients: Predictors of compliance and satisfaction. *J. Telemed. Telecare* **24**, 93–100 (2018).
38. Selter, A. *et al.* An mhealth app for self-management of chronic lower back pain (Limbr): Pilot study. *JMIR mHealth uHealth* **6**, 1–12 (2018).
39. Rijsbergen, M. *et al.* Mobile e-diary application facilitates the monitoring of patient-reported outcomes and a high treatment adherence for clinical trials in dermatology. *Journal of the European Academy of Dermatology and Venereology* vol. 34 633–639 (2020).
40. Van Til, K., McInnis, M. G. & Cochran, A. A comparative study of engagement in mobile and wearable health monitoring for bipolar disorder. *Bipolar Disord.* **22**, 182–190 (2020).
41. Anguera, J. A., Jordan, J. T., Castaneda, D., Gazzaley, A. & Areán, P. A. Conducting a fully mobile and randomised clinical trial for depression: Access, engagement and expense. *BMJ Innov.* **2**, 14–21 (2016).
42. Bauer, A. M. *et al.* Acceptability of mHealth augmentation of Collaborative Care: A mixed methods pilot study. *Gen. Hosp. Psychiatry* **51**, 22–29 (2018).
43. Faurholt-Jepsen, M. *et al.* Daily electronic self-monitoring in bipolar disorder using smartphones - The MONARCA i trial: A randomized, placebo-controlled, single-blind, parallel group trial. *Psychol. Med.* **45**, 2691–2704 (2015).
44. Faurholt-Jepsen, M. *et al.* Smartphone-based self-monitoring in bipolar disorder: evaluation of usability and feasibility of two systems. *Int. J. Bipolar Disord.* **7**, 1–11 (2019).
45. Faurholt-Jepsen, M. *et al.* Voice analysis as an objective state marker in bipolar disorder. *Transl. Psychiatry* **6**, e856 (2016).

46. Stanislaus, S. *et al.* Mood instability in patients with newly diagnosed bipolar disorder, unaffected relatives, and healthy control individuals measured daily using smartphones. *J. Affect. Disord.* **271**, 336–344 (2020).
47. Gustavell, T., Sundberg, K., Segersvärd, R., Wengström, Y. & Langius-Eklöf, A. Decreased symptom burden following surgery due to support from an interactive app for symptom management for patients with pancreatic and periampullary cancer. *Acta Oncol. (Madr)*. **58**, 1307–1314 (2019).
48. Niendam, T. A. *et al.* Enhancing early psychosis treatment using smartphone technology: A longitudinal feasibility and validity study. *J. Psychiatr. Res.* **96**, 239–246 (2018).
49. Jamison, R. N., Mei, A., Edwards, R. R. & Ross, E. L. Efficacy of vibrating gloves for chronic hand pain due to osteoarthritis. *Pain Med. (United States)* **19**, 1044–1057 (2018).
50. Crouthamel, M. *et al.* Using a researchkit smartphone app to collect rheumatoid arthritis symptoms from real-world participants: Feasibility study. *JMIR mHealth uHealth* **6**, 1–14 (2018).
51. Faurholt-Jepsen, M. *et al.* Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. *Int. J. Methods Psychiatr. Res.* **25**, 309–323 (2016).
52. Hung, S. *et al.* Smartphone-based ecological momentary assessment for Chinese patients with depression: An exploratory study in Taiwan. *Asian J. Psychiatr.* **23**, 131–136 (2016).
53. Rodriguez Hermosa, J. L. *et al.* Compliance and Utility of a Smartphone App for the Detection of Exacerbations in Patients With Chronic Obstructive Pulmonary Disease: Cohort Study. *JMIR mHealth uHealth* **8**, e15699 (2020).
54. Buck, B. *et al.* Capturing behavioral indicators of persecutory ideation using mobile technology. *J. Psychiatr. Res.* **116**, 112–117 (2019).
55. Moukaddam, N., Truong, A., Cao, J., Shah, A. & Sabharwal, A. Findings from a trial of the smartphone and online usage-based evaluation for depression (SOLVD) application: What do apps really tell us about patients with depression? concordance between app-generated data and standard psychiatric questionnaires for de. *J. Psychiatr. Pract.* **25**, 365–373 (2019).
56. Broderick, J. E. *et al.* Patient reported outcomes can improve performance status assessment: a pilot study. *J. Patient-Reported Outcomes* **3**, (2019).
57. Lenaert, B., Neijmeijer, M., van Kampen, N., van Heugten, C. & Ponds, R. Poststroke Fatigue and Daily Activity Patterns During Outpatient Rehabilitation: An Experience Sampling Method Study. *Arch. Phys. Med. Rehabil.* **101**, 1001–1008 (2020).
58. Faurholt-Jepsen, M. *et al.* Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Res.* **217**, 124–127 (2014).
59. Kim, J. *et al.* Depression screening using daily mental-health ratings from a smartphone application for breast cancer patients. *J. Med. Internet Res.* **18**, 1–15 (2016).
60. Band, R., Barrowclough, C., Caldwell, K., Emsley, R. & Wearden, A. Activity patterns in response to symptoms in patients being treated for chronic fatigue syndrome: An experience sampling methodology study. *Heal. Psychol. Off. J. Div. Heal. Psychol. Am. Psychol. Assoc.* **36**, 264–269 (2017).

61. Probst, T. *et al.* Does tinnitus depend on time-of-day? An ecological momentary assessment study with the 'TrackYourTinnitus' application. *Front. Aging Neurosci.* **9**, 1–9 (2017).
62. Kauer, S. D. *et al.* Self-monitoring using mobile phones in the early stages of adolescent depression: Randomized controlled trial. *J. Med. Internet Res.* **14**, 1–17 (2012).
63. Carpenter, R. W., Lane, S. P., Bruehl, S. & Trull, T. J. Concurrent and lagged associations of prescription opioid use with pain and negative affect in the daily lives of chronic pain patients. *J. Consult. Clin. Psychol.* **87**, 872–886 (2019).
64. Faurholt-Jepsen, M. *et al.* Is smartphone-based mood instability associated with stress, quality of life, and functioning in bipolar disorder? *Bipolar Disord.* **21**, 611–620 (2019).
65. Ireland, D. & Andrews, N. Pain ROADMAP: A Mobile Platform to Support Activity Pacing for Chronic Pain. *Stud. Health Technol. Inform.* **266**, 89–94 (2019).
66. Faurholt-Jepsen, M. *et al.* Differences in mood instability in patients with bipolar disorder type I and II: a smartphone-based study. *Int. J. Bipolar Disord.* **7**, (2019).
67. Reid, S. C. *et al.* A mobile phone application for the assessment and management of youth mental health problems in primary care: A randomised controlled trial. *BMC Fam. Pract.* **12**, 131 (2011).
68. Corden, M. E. *et al.* MedLink: A mobile intervention to improve medication adherence and processes of care for treatment of depression in general medicine. *Digit. Heal.* **2**, 2055207616663069 (2016).
69. Lin, W.-C., Burke, L., Schlenk, E. A. & Yeh, C. H. Use of an Ecological Momentary Assessment Application to Assess the Effects of Auricular Point Acupressure for Chronic Low Back Pain. *Comput. Inform. Nurs.* **37**, 276–282 (2019).
70. Yang, Y. S., Ryu, G. W. & Choi, M. Factors associated with daily completion rates in a smartphone-based ecological momentary assessment study. *Healthc. Inform. Res.* **25**, 332–337 (2019).
71. Torous, J. *et al.* Utilizing a Personal Smartphone Custom App to Assess the Patient Health Questionnaire-9 (PHQ-9) Depressive Symptoms in Patients With Major Depressive Disorder. *JMIR Ment. Heal.* **2**, e8 (2015).
72. Suso-Ribera, C. *et al.* Validity, reliability, feasibility, and usefulness of pain monitor. *Clin. J. Pain* **34**, 900–908 (2018).
73. Wu, J. Q. & Cronin-Golomb, A. Temporal Associations between Sleep and Daytime Functioning in Parkinson's Disease: A Smartphone-Based Ecological Momentary Assessment. *Behav. Sleep Med.* **18**, 560–569 (2020).
74. Beiwinkel, T. *et al.* Using smartphones to monitor bipolar disorder symptoms: A pilot study. *JMIR Ment. Heal.* **3**, (2016).
75. Ben-Zeev, D. *et al.* Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophr. Bull.* **40**, 1244–1253 (2014).
76. Prada, P. *et al.* EMOTE0: A Smartphone Application for Monitoring and Reducing Aversive Tension in Borderline Personality Disorder Patients, a Pilot Study. *Perspect. Psychiatr. Care* **53**, 289–298 (2017).

77. Sengupta, A., Beckie, T., Dutta, K., Dey, A. & Chellappan, S. A mobile health intervention system for women with coronary heart disease: usability study. *JMIR Form. Res.* **4**, (2020).
78. Gustavell, T., Sundberg, K. & Langius-Eklöf, A. Using an interactive app for symptom reporting and management following pancreatic cancer surgery to facilitate person-centered care: Descriptive study. *JMIR mHealth uHealth* **8**, 1–14 (2020).
79. Reid, S. C. *et al.* A mobile phone application for the assessment and management of youth mental health problems in primary care: Health service outcomes from a randomised controlled trial of mobiletype. *BMC Fam. Pract.* **14**, 1 (2013).
80. Suso-Ribera, C. *et al.* Validity, Reliability, Feasibility, and Usefulness of Pain Monitor: A Multidimensional Smartphone App for Daily Monitoring of Adults With Heterogenous Chronic Pain. *Clin. J. Pain* **34**, 900–908 (2018).
81. Bandarian-Balooch, S., Martin, P. R., McNally, B., Brunelli, A. & Mackenzie, S. Electronic-Diary for Recording Headaches, Triggers, and Medication Use: Development and Evaluation. *Headache* **57**, 1551–1569 (2017).
82. Mohr, D. C. *et al.* MedLink: A Mobile Intervention to Address Failure Points in the Treatment of Depression in General Medicine. *Int. Conf. Pervasive Comput. Technol. Healthc. [proceedings]. Int. Conf. Pervasive Comput. Technol. Healthc.* **2015**, 100–107 (2015).
83. Hidalgo-Mazzei, D. *et al.* Psychoeducation in bipolar disorder with a SIMPLE smartphone application: Feasibility, acceptability and satisfaction. *J. Affect. Disord.* **200**, 58–66 (2016).
84. Ross, E. L., Jamison, R. N., Nicholls, L., Perry, B. M. & Nolen, K. D. Clinical Integration of a Smartphone App for Patients With Chronic Pain: Retrospective Analysis of Predictors of Benefits and Patient Engagement Between Clinic Visits. *J. Med. Internet Res.* **22**, e16939 (2020).
85. Faurholt-Jepsen, M. *et al.* The effect of smartphone-based monitoring on illness activity in bipolar disorder: The MONARCA II randomized controlled single-blinded trial. *Psychol. Med.* **50**, 838–848 (2020).
86. Shafran, R. *et al.* Translating the Intention to Seek Treatment into Action: Does Symptom Monitoring Make a Difference? Results from a Randomized Controlled Trial. *Behav. Cogn. Psychother.* **47**, 114–128 (2019).
87. Busk, J. *et al.* Forecasting Mood in Bipolar Disorder from Smartphone Self-assessments: Hierarchical Bayesian approach. *JMIR mHealth uHealth* **8**, 1–14 (2020).
88. Barello, S. *et al.* eHealth for patient engagement: A Systematic Review. *Front. Psychol.* **6**, (2016).
89. Faurholt-Jepsen, M. *et al.* Mood instability in bipolar disorder type I versus type II-continuous daily electronic self-monitoring of illness activity using smartphones. *J. Affect. Disord.* **186**, 342–349 (2015).
90. Busk, J. *et al.* Daily estimates of clinical severity of symptoms in bipolar disorder from smartphone-based self-assessments. *Transl. Psychiatry* **10**, (2020).
91. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, (2021).

92. Dixon-Woods, M. *et al.* Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med. Res. Methodol.* **6**, 1–13 (2006).
93. Depraetere, J., Vandeviver, C., Keygnaert, I. & Beken, T. Vander. The critical interpretive synthesis: an assessment of reporting practices. *Int. J. Soc. Res. Methodol.* (2020) doi:10.1080/13645579.2020.1799637.
94. Maastricht University. PsyMate. <https://www.psymate.eu/>.
95. Team, T. E. EndNote. (2013).
96. QSR International Pty Ltd. nVivo. (2020).

Table 1

Table 1. N (%) of papers that included each of the engagement definition and measurement constructs.

Characteristic	N	%
Engagement as a main aim^a	30	39.5
Gave a definition of engagement^b	26	35.1
Objective, e.g. feasibility ^b	18	24.3
Subjective, e.g. usability ^b	11	14.9
Utility ^b	3	4.1
App usage ^b	4	5.4
Research protocol	1	1.4
Reported a measure of engagement^a	74	97.4
With the research protocol ^b	65	87.8
Objective RMTs ^b	64	86.5
Subjective RMTs ^b	27	36.5
Combined objective and subjective RMTs ^b	22	29.8

Table 2. ^a Denominator of 76. ^b Denominator of 74. Categories are not mutually exclusive.

Figures

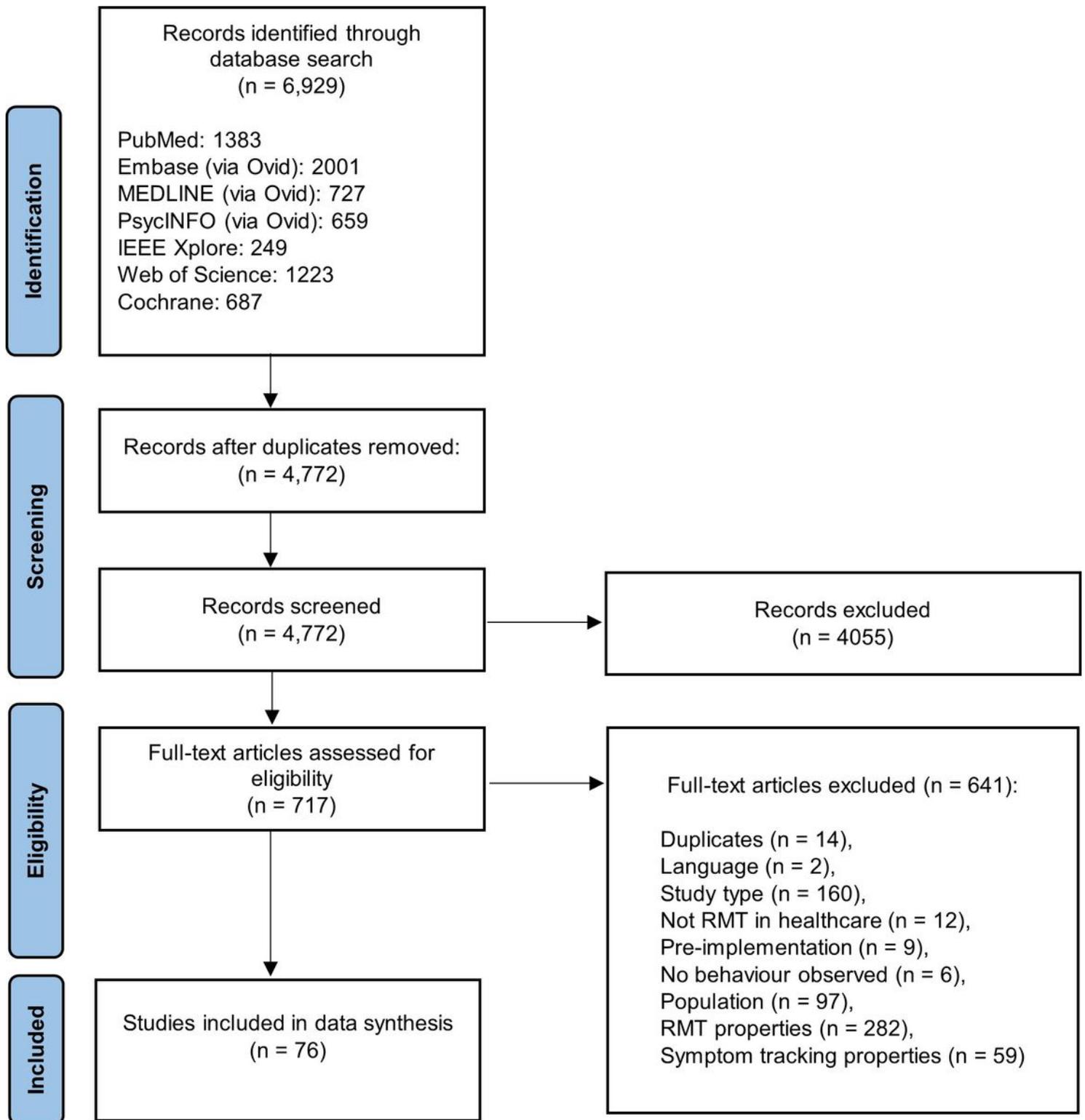


Figure 1

PRISMA flow diagram of included studies. Seven databases were searched to ensure relevant fields were covered. The flow diagram lists reasons for exclusion of articles from the final sample of n=76.

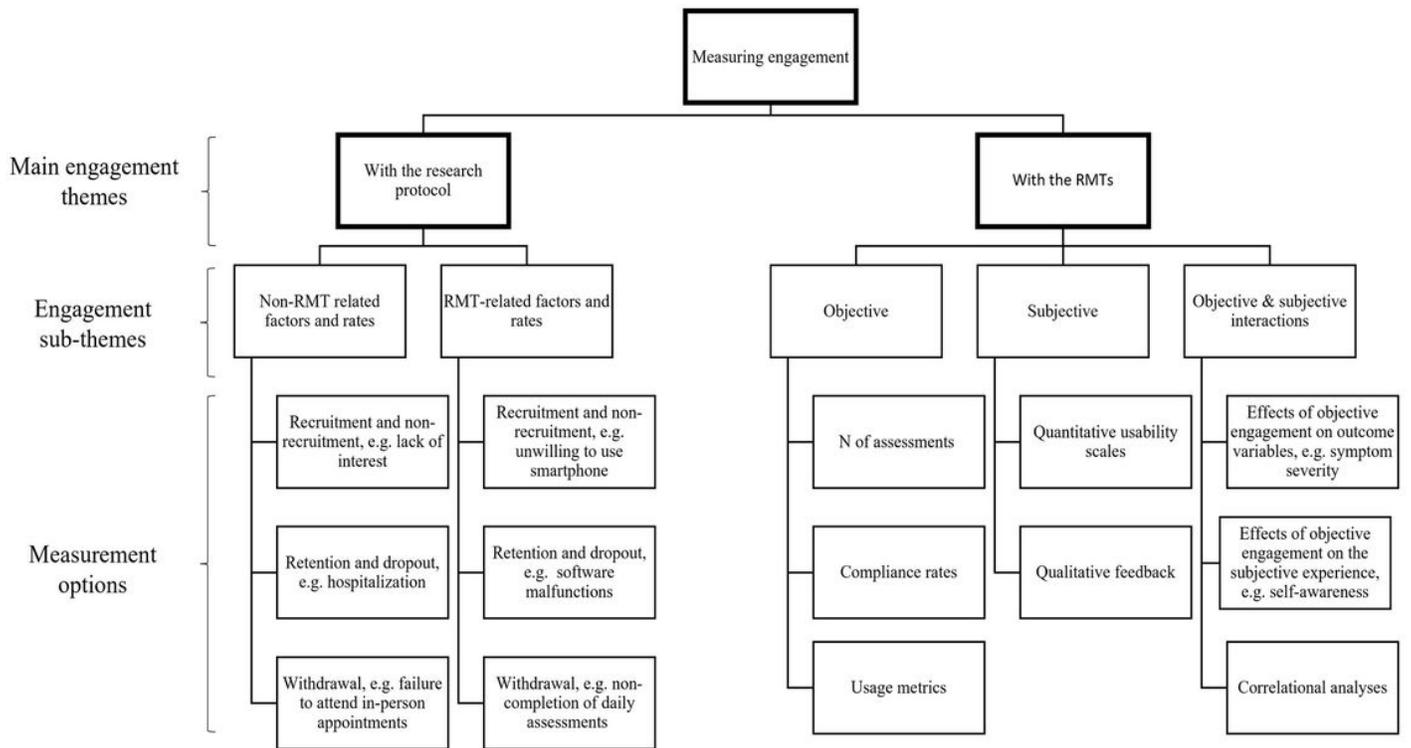


Figure 2

An integrative framework for the measurement of engagement with RMTs, based on the four synthetic constructs found. The main engagement themes cover ‘engagement with the research protocol’ and ‘engagement with RMTs’. Further engagement sub-themes correspond to ways in which engagement can be conceptualised within each of the two main themes. The third section outlines several available options for measurement within each sub-theme.

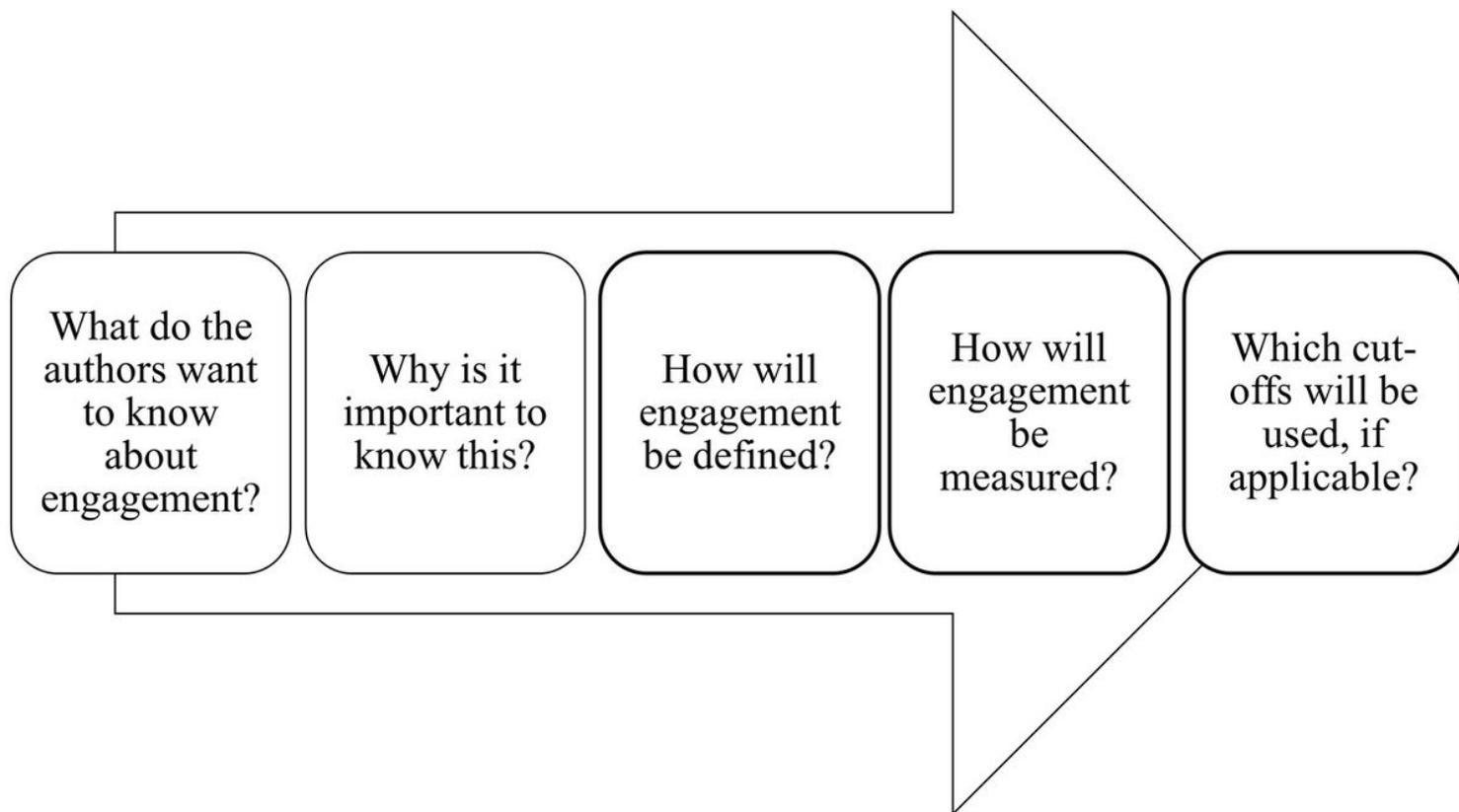


Figure 3

Implementing the reporting of engagement into the study design process for RMT studies. Authors are encouraged to conceptualise and define key measurement strategies for engagement during the study development phase.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM1searchstrategyrevised030522.docx](#)
- [ESM2includedstudies.docx](#)