Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Configurable Native mHealth Apps for Ecological Momentary Assessment and Intervention: A Systematic Literature Review of Reusable Systems and Frameworks

Carsten Vogel
carsten.vogel@uni-wuerzburg.de
University of Würzburg
Miriam Schlüter
University of Würzburg
Abdul Rahman Idrees
Ulm University
Michael Winter
University of Würzburg
Rüdiger Pryss
University of Würzburg

Article

Keywords: Ecological Momentary Assessment (EMA), Ecological Momentary Intervention (EMI), Just-in-Time Adaptive Interventions (JITAI), mHealth Platforms, Configurable Systems, No-Code Development, Reusable Frameworks, Systematic Literature Review.

Posted Date: April 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9294623/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Configurable Native mHealth Apps for Ecological Momentary Assessment and Intervention: A Systematic Literature Review of Reusable Systems and Frameworks

Carsten Vogel [0000-0002-2190-5349] [1,2,*], **Miriam Schlüter** [1,2,+], **Abdul Rahman Idrees** [0000-0002-9861-1947] [3,+], **Michael Winter** [0000-0003-2561-7923] [1,2,+], and **Rüdiger Pryss** [0000-0003-1522-785X] [1,2,+]

[1]Institute of Clinical Epidemiology and Biometry, University of Würzburg, Würzburg, Germany
[2]Institute of Medical Data Science, University Hospital of Würzburg, Würzburg, Germany
[3]Institute of Databases and Information Systems, Ulm University, Ulm, Germany
[4]Department of Clinical Psychology and Psychotherapy, Ulm University, Ulm, Germany
[*]Corresponding author: carsten.vogel@uni-wuerzburg.de
[+]These authors contributed equally to this work.

ABSTRACT

Introduction: Ecological Momentary Assessment (EMA) and Ecological Momentary Intervention (EMI), including Just-In-Time Adaptive Intervention (JITAIs), are increasingly implemented through (native) mobile health (mHealth) applications. Existing reviews emphasize methodological and clinical aspects, while the technical architectures and reusable framework characteristics of native EMA/EMI systems remain fragmented across disciplines. Given the heterogeneity and rapid evolution of these systems, a literature review was conducted to systematically map this emerging field. **Objective:** This review systematically maps configurable and reusable native mHealth platforms that integrate EMA and EMI/JITAI, focusing on architectural patterns, configurability mechanisms, and the operationalization of adaptive interventions, as well as gaps in technical reporting and evaluation maturity. **Methods:** Following PRISMA guidelines and a preregistered protocol (OSF), we searched ACM Digital Library, APA PsycINFO, PubMed, IEEE Xplore, Web of Science, and Google Scholar for studies published from 2009 to July 11, 2024. Peer-reviewed publications describing implemented native mobile applications with integrated EMA and EMI functionality were included. Three reviewers screened 3,724 records at title–abstract level, and two reviewers assessed 399 full texts. Data were extracted using a structured form and synthesized descriptively. **Results:** Forty-one publications describing 34 distinct systems were included. Most platforms were presented as configurable frameworks enabling multi-study deployment and runtime configuration of assessment and intervention logic. EMA functionality was typically implemented through hybrid sampling strategies combining signal-based prompting (40/41), event-contingent sampling (34/41), and passive sensing (24/41). Automated intervention delivery was widely supported (39/41), whereas explicit AI-driven decision logic was rare (4/41), indicating that adaptivity was predominantly implemented through rule-based or configurable mechanisms rather than data-driven models. Evaluation maturity was limited, with feasibility and pilot studies predominating and 16 publications lacking empirical outcomes. Visual reporting was common but varied in completeness, with 27/41 publications including user interface screenshots, 21/41 system architecture figures, 10/41 code architecture visualizations, and all 41/41 presenting primarily tabular data visualizations. Technical reporting was frequently incomplete—particularly for backend architecture, databases, cloud infrastructure, and interoperability—limiting comparability and reproducibility across systems Overall, the field shows substantial architectural heterogeneity without clear convergence toward standardized design patterns. **Conclusions:** Configurable native EMA/EMI frameworks and reusable systems represent a diverse and expanding design space characterized by broad configurability and automated intervention capabilities, but still a minority of systems and with uneven evaluation maturity and insufficient architectural transparency. Standardized reporting of system architectures, clearer definitions of adaptivity and AI, and stronger alignment between platform capabilities and empirical validation are needed to improve reproducibility, comparability, and practical reuse.

Keywords: Ecological Momentary Assessment (EMA); Ecological Momentary Intervention (EMI); Just-in-Time Adaptive Interventions (JITAI); mHealth Platforms; Configurable Systems; No-Code Development; Reusable Frameworks; Systematic Literature Review.

1 Introduction

Ecological Momentary Assessment (EMA) and related Experience Sampling Methods (ESMs), which both capture experiences, symptoms, and behaviors repeatedly in everyday life, are widely used to study processes in naturalistic settings[1]. Ecological Momentary Interventions (EMIs), including Just-In-Time Adaptive Interventions (JITAIs), complement EMA by enabling data collection and intervention delivery in real-world contexts, capturing experiences in the moment and providing timely, context-sensitive support, rather than relying solely on retrospective reports or clinic-based encounters[2,3]. As smartphones have become ubiquitous, they now serve as the primary infrastructure for implementing EMA and EMI in research and practice, combining continuous availability, sensing capabilities, and timely notifications[3]. Across the EMA literature, delivery modes range from paper diaries and telephones to handheld and smartphone technologies, reflecting a heterogeneous methodological landscape, while reviews commonly emphasize study design and reporting rather than reusable software architecture[1].

As a result, a substantial body of literature has reviewed EMA and EMI from methodological, technical, and clinical perspectives[4–8]. Many EMA- and ESM-focused reviews synthesize sampling strategies, measurement characteristics, adherence, and feasibility, often treating the concrete software implementation as secondary[4,5]. Similarly, EMI and JITAI reviews frequently emphasize intervention components, tailoring variables, decision rules, and effectiveness outcomes rather than the software systems used to operationalize them[3,6,7]. This emphasis has clarified how EMA and EMI can be designed and evaluated, yet it offers limited guidance for researchers seeking reusable technological infrastructures supporting multiple projects[9]. Tool-oriented publications and tutorials provide valuable insights into specific systems and selection criteria, but they often cover broad tool landscapes without restricting to configurable native frameworks and reusable systems with integrated EMI support[9,10].

1.1 Rationale

Over the past decade, numerous mobile systems have been developed to support these methods, ranging from single-study applications to more generalizable platforms. Despite this growth, there is no consolidated overview of reusable, configurable platforms that integrate both EMA and EMI. Relevant technical information on system architectures, configurability mechanisms, and reuse strategies is fragmented across publications and often embedded within intervention-focused studies. Consequently, it remains difficult to determine which published systems qualify as configurable native mobile health (mHealth) frameworks designed for reuse and integrating both EMA and EMI, and to systematically compare their architectural and configurability approaches across studies. This complicates the identification, comparison, and practical selection of such systems in research and implementation contexts.

Native implementation is a key practical dimension for EMA and EMI because it affects access to OS-level notification scheduling, sensing integration, offline behavior, and performance constraints[3,11]. In this review, we use the term framework to denote systems intended for reuse across studies or applications, for example by providing multi-study support, runtime configuration, administrative tooling, and extensibility mechanisms[9].

To our knowledge, prior reviews have not specifically mapped configurable, native reusable systems or frameworks that integrate both EMA and EMI, nor have they focused on this intersection[7–9]. This review addresses this gap by identifying and characterizing studies that present such reusable mHealth systems[11–13]. By focusing on configurability, native implementation, and the integration of assessment and intervention within reusable systems or frameworks, this review complements prior syntheses centered on study methodology and intervention outcomes[4,6,7]. The resulting overview aims to support informed platform selection, guide future framework development, and improve transparency and comparability in the reporting of EMA and EMI infrastructures[9,10].

1.2 Objectives

Given the rapid pace of technological development and the cross-disciplinary dispersion of technical descriptions, a systematic mapping of this field is needed. The objective of this review is to systematically identify and map configurable native mHealth platforms that support EMA in combination with EMI and/or JITAI, with a focus on reusable systems and frameworks. Accordingly, the primary research question (PRQ) is:

PRQ: What configurable native mHealth frameworks and reusable systems have been developed to support EMA in combination with EMI and/or JITAI?

To further characterize these systems and their evolution, the following secondary research questionss (SRQs) are investigated:

SRQ1 What structural architectures, implementation strategies, configurability mechanisms, and reusable technical design patterns characterize these systems?

SRQ2 How do these platforms enable and operationalize health-related assessment and intervention delivery at runtime, including triggering logic, personalization mechanisms, and adaptive decision processes?

SRQ3 What temporal patterns can be observed in the reporting of architectural approaches, configurability strategies, and system capabilities since 2009?

By addressing these questions, the review seeks to clarify current capabilities, identify structural and reporting gaps, and inform future system development and research in configurable mHealth infrastructures.

2 Methods

This review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysesc (PRISMA)[14] and the Extension for Scoping Reviews (PRISMA-ScR)[15].

2.1 Protocol and Registration

The review protocol[16] was registered on the Open Science Framework (OSF)[17] prior to data extraction, synthesis, and final analysis.

2.2 Eligibility criteria

The unit of analysis was implemented mobile software systems rather than individual study participants or clinical outcomes.

Publication characteristics. We included peer-reviewed journal articles and peer-reviewed full conference papers reporting primary empirical, technical, or system-development contributions. We excluded secondary and tertiary literature (e.g., systematic reviews, scoping reviews, meta-analyses, narrative reviews), complete books, full conference proceedings volumes, dissertations, theses, and other non–peer-reviewed publications. Only articles published in English were considered to ensure feasibility and consistency of full-text assessment.

Time frame. We restricted inclusion to publications from 2009 onward. This cutoff reflects the widespread adoption of modern smartphones, mobile operating systems, and centralized app distribution platforms (e.g., app stores), which collectively enabled technically viable and scalable native EMA and EMI systems. Studies published prior to 2009 were unlikely to describe mobile architectures relevant to contemporary configurable mHealth infrastructures.

Concept and system requirements. Eligible publications had to describe a technically implemented mobile application integrating EMA with EMI and/or JITAI functionality within a health-related or behavioral context. The system had to be technically implemented and described as functional, rather than being purely conceptual, theoretical, or protocol-based. Evidence of user evaluation was not required for inclusion. During full-text eligibility assessment, systems were required to meet all of the following criteria:

- The system operated as a standalone application on participants' mobile devices.
- The system implemented both assessment and intervention functionality within the same mobile application.
- The publication described the technical architecture or platform beyond a superficial mentioning of EMA/EMI terminology.
- The system included at least one configurable component, such as editable assessment schedules, intervention content, decision rules, parameter settings, modular components, or developer-level configuration options.
- The system demonstrated architectural abstraction or modularization (e.g., separation of content, logic, or configuration from hard-coded components) indicating potential for reuse beyond a single-purpose implementation.

Systems were excluded if no clear EMA component could be identified during the full-text review, if intervention functionality was absent, or if the publication referenced EMA or EMI without describing the underlying mobile system. If intervention functionality was not explicitly described in the publication, the system was considered not to include EMI capability. Single-purpose implementations without configurable or reusable architectural elements were excluded following full-text assessment. Health-related or behavioral contexts included clinical, preventive, wellness, mental health, lifestyle, and behavior-change applications targeting physical, psychological, or behavioral outcomes, irrespective of formal medical diagnosis. General-purpose EMA/EMI platforms were eligible if intended for health-related or behavioral applications.

Platform and deployment context. Eligible systems were required to operate as standalone applications on participants' mobile devices in real-world contexts. Progressive web applications were included if they were installable on mobile devices and functioned as self-contained, app-like systems from the end-user perspective. We excluded:

- web-only platforms and platforms primarily accessed through desktop browsers,
- systems that mainly redirected users to external websites or tools for assessment or intervention delivery,
- interventions delivered primarily through third-party applications (e.g., messaging or social media platforms),
- systems relying predominantly on non-automated human-mediated delivery (e.g., telephone calls or postal communication), and
- applications restricted to controlled laboratory or inpatient infrastructures that were not deployable outside those controlled settings.

Assessment of configurability and reuse. Due to limited reporting in titles and abstracts, configurability features and reusable architectural characteristics were not evaluated during initial screening. Instead, these attributes were systematically assessed during full-text eligibility assessment using a structured extraction framework. Final inclusion required confirmation of at least one configurable component and evidence of architectural abstraction or modularization supporting potential reuse. This approach ensured comprehensive identification of relevant systems while enabling structured analysis of reusable and configurable design patterns in the final corpus.

2.3 Information sources

A comprehensive search of interdisciplinary bibliographic databases was conducted to identify literature on native mHealth applications supporting EMA and EMI. Five bibliographic databases were searched to identify relevant studies: **ACM Digital Library**, **APA PsycINFO (via Ovid)**, **PubMed**, **IEEE Xplore**, and **Web of Science Core Collection**. In addition, Google Scholar was searched as a supplementary source using Publish or Perish. The first 300 records ranked by relevance were screened. The selected databases collectively cover mobile computing, human–computer interaction, behavioral science, psychology, digital health, and biomedical research.

Each database was searched independently using a shared conceptual search strategy (see Section 2.4), with syntax and field adaptations tailored to the requirements of each platform. All database searches were completed on **2024-07-11**. A publication-date filter restricted results to records published from 2009 onward; no additional methodological or design filters were applied. Each database was searched across its full available coverage. Search results from all databases were exported separately and later merged during record-processing, prior to screening. Database searches were not supplemented by additional sources.

2.4 Search

The search strategy was structured around three core concepts: (1) EMA and related assessment approaches, (2) EMI, JITAI, and related digital intervention mechanisms, and (3) mobile applications, smartphones, sensing technologies, and mHealth platforms. The strategy was developed iteratively by the lead author in consultation with the review team and informed by core PRESS principles[18]. Candidate terms were identified through preliminary scoping searches, key-paper harvesting, terminology mapping across disciplines (behavioral science, HCI, digital health), and controlled vocabulary inspection (e.g., MeSH). For each concept, extensive synonym sets were constructed using keyword variants, truncation, phrase searching, and controlled vocabulary where available. The core Boolean structure followed:

```
<Assessment Terms> AND <Intervention Terms>
    AND <Mobile App Technology Terms>
```

Searches were intentionally optimized for recall rather than precision to minimize the risk of missing eligible reusable systems and frameworks in this heterogeneous literature. Broad generic terms (e.g., "intervention*") were tested but removed when they produced excessive irrelevant retrieval. More specific formulations (e.g., "app-based intervention*", "mHealth intervention*") were retained to preserve sensitivity while reducing noise.

A consistent Boolean structure combining the three concept groups was maintained across databases Queries were adapted only to accommodate database-specific syntax, indexing systems, field tags, controlled vocabularies, character limits, and interface constraints. Where necessary, long strings were split into multiple executions due to platform restrictions. The following was specific to individual databases:

- **Web of Science**: full Boolean query using "Topic" fields.
- **Scopus**: adapted to Scopus field tags (TITLE-ABS-KEY).
- **ScienceDirect**: free-text query executed in title–abstract–keyword fields.
- **PubMed**: combination of MeSH terms and All Fields terms.
- **IEEE Xplore**: exact-match with plural-variant expansion.
- **Google Scholar**: searched as a pragmatic supplement using *Publish or Perish*; a simplified core-concept query accommodated platform constraints, and the first 300 ranked results were screened to capture relevant interdisciplinary and gray literature while limiting noise and redundancy.

Full database-specific search strings, including exact syntax and execution details, are available in the OSF repository[16].

A publication-year restriction (2009 onward) was imposed to reflect the emergence and maturation of smartphone-based mHealth ecosystems. No methodological or other filters were applied at the database-interface level. All records were exported in full without database-level deduplication. Deduplication, metadata normalization, and bibliographically operationalizable eligibility constraints (e.g., language and publication-type exclusions) were applied after merging all sources into a unified reference management workflow prior to screening.

2.5 Selection of sources of evidence

Records retrieved from each database were exported in full, without applying database-interface filters other than the publication-year restriction (2009 onward). Each export contained all available metadata (e.g., identifiers, titles, abstracts, authors, venue, and subject tags). All exports were merged into a single master dataset, and metadata fields were standardized across sources prior to screening. Where abstracts were missing but a DOI was available, abstracts were supplemented when feasible via DOI-based lookups to ensure sufficient information for title–abstract screening. Eligibility constraints that could be operationalized bibliographically were applied during preprocessing. This included verification of the publication-year restriction, removal of non-English publications, exclusion of ineligible publication types (e.g., review-derived records), and removal of publication containers such as full books or proceedings volumes. Records with incomplete metadata were retained unless clearly unusable. Deduplication was conducted after merging all exports. Duplicates were identified primarily using DOI matching and, where DOIs were unavailable, by matching titles, authors, and publication years. Potential duplicates with non-identical metadata were manually inspected to prevent erroneous removal. All preprocessing was completed before reviewers accessed the records.

Title–abstract screening Title–abstract screening was conducted in `Rayyan`[1], a web-based systematic review management tool[19]. Prior to formal screening, a calibration exercise was performed in which all three reviewers independently screened the first 50 records against the core criteria (EMA, EMI/JITAI, and native mobile app implementation). Reviewers then met to compare judgments and align interpretations of eligibility. Screening procedures were deemed feasible; the only procedural adjustment was the introduction of standardized visual tags in `Rayyan` to mark key EMA/EMI and mobile-app indicators and support efficient triage.

After calibration, **three reviewers independently screened all records in blinded mode**, assigning each record to the categories *include*, *exclude*, or *unclear*. Reviewers could add tags and notes to support consistent decision-making and subsequent conflict resolution. No additional machine-learning or automation features in `Rayyan` were used. After merging decisions, conflicts were resolved using predefined inclusion-favoring rules: **include** (≥ 1 *include*, no *exclude*); **exclude** (≥ 2 *exclude*, no *include*); **discussion** (all remaining cases resolved by reviewer consensus). Included records proceeded to full-text screening.

Full-text screening All available full-text PDFs were assessed using a structured Google Forms questionnaire to operationalize eligibility criteria and capture metadata for subsequent charting.

Each paper was assessed by **two independent reviewers**: one reviewer screened all 399 papers, while three additional reviewers independently conducted a second, blinded screening of the full set, divided into subsets of 200, 163, and 36 papers. At the end of the form, reviewers recorded an inclusion, exclusion, or uncertainty judgment along with reasons for inclusion or exclusion.

[1] Rayyan (old version): https://www.rayyan.ai (accessed 2024-08-01)

Questionnaire responses were merged automatically where possible; disagreements were resolved through discussion between the relevant reviewers. In rare cases of irreconcilable conflict (e.g., *include* vs. *exclude*), a third author not involved in the original assessments acted as tie-breaker, with access to both reviewers' form entries and notes. Automated scripts were used for merging and organizing responses.

Two-stage eligibility decision at full text Eligibility determination followed a two-step process.

First, reviewers assessed whether the paper described a native mobile app with EMA/EMI. Papers failing this requirement were excluded (**Reason 1: Not describing a native mobile EMA/EMI application**).

Second, remaining papers were evaluated against additional functional and conceptual criteria: **Reason 2 — Lacking EMA features; Reason 3 — Lacking EMI features; Reason 4 — Lacking configurability characteristics; Reason 5 — Lacking reusable system or framework characteristics; Reason 6 — Just a concept without implementation**.

Final verification The final set of included and excluded studies was manually checked by two authors for consistency prior to finalization.

2.6 Data charting process

Data charting was conducted using a structured Google Forms instrument developed by the review team, which also served as the standardized data extraction template during full-text assessment. The form contained single-choice, multiple-choice, and open-text items covering bibliographic metadata, study design, EMA and EMI characteristics, technical features of the mobile application or framework, configurability, reusability, and implementation details.

Prior to full implementation, the form was calibrated by two reviewers who independently charted data from the first 20 full-text records and subsequently aligned interpretations through discussion.

All retrieved full-text records were independently assessed in duplicate using the structured charting instrument. The form was completed in full for each record during full-text assessment, irrespective of the final eligibility decision. For records meeting inclusion criteria, extracted variables were retained for synthesis; for excluded records, charted data were used solely to document eligibility decisions.

The finalized dataset underwent a manual consistency check prior to analysis. Study authors were not contacted for additional information.

2.7 Data items

All variables were predefined within the standardized charting form and derived from the objectives of the review. Variables were grouped into conceptual domains to reflect system-level, intervention-level, and evaluation-level characteristics. For each of the 399 full-text records assessed, we extracted bibliographic metadata (e.g., publication year, venue, and author-provided keywords) and study characteristics (e.g., study design, trial type, participant characteristics, and target population).

EMA-related variables captured whether ecological momentary assessment was implemented and, if so, the type of assessment (self-report, sensor-based, hybrid), sampling logic (fixed, random, event-based), prompting schedules, and contextual triggering mechanisms.

EMI-related variables included intervention type (manual vs. automated), personalization strategies, and context-aware or sensor-driven adaptations.

EMI-related variables captured whether ecological momentary intervention components were implemented, the degree of automation (manual vs. automated delivery), personalization strategies, and context-aware or sensor-driven adaptation mechanisms.

Technical system characteristics included frontend and backend frameworks, architectural descriptions, cloud services, sensing modalities, integration with external devices, testing practices (e.g., CI/CD), and use of advanced technologies (e.g., AR, VR, blockchain).

Artificial intelligence and analytics variables captured whether AI or machine learning techniques were implemented in intervention logic or analytics pipelines, whether AI was discussed conceptually, and which analytic approaches were reported.

Evaluation and reporting variables included reported evaluation approaches; use of standardized instruments (e.g., MARS, SUS); security and privacy considerations (e.g., GDPR, HIPAA compliance); funding sources; and the depth of system and evaluation reporting.

In addition, studies were classified as describing either (a) a general EMA/EMI mobile application or (b) a configurable and/or reusable framework. Configurability was defined as the ability to modify system parameters without altering source code.

Where information was not explicitly reported, variables were coded as *not reported*. No assumptions were made beyond explicitly stated descriptions in the source publications.

2.8 Critical appraisal of individual sources of evidence

A formal critical appraisal of individual sources of evidence was not conducted. This review aimed at mapping and characterizing the technological design, configurability, architectural properties, and implementation features of native EMA/EMI mobile systems. The objective was descriptive and exploratory rather than evaluative. Given the descriptive and technology-focused objective of this review, formal methodological quality or risk-of-bias assessment using established tools was not aligned with its purpose. Therefore, no appraisal instrument was applied, and quality assessments did not inform inclusion or synthesis.

2.9 Synthesis of results

Extracted data were synthesized descriptively to address the primary and secondary research questions. The primary unit of quantitative synthesis was the *publication*, consistent with the structured charting form and the tables presented in the Results section. In parallel, publications were mapped to *distinct systems/frameworks* to report the number of unique platforms represented and to identify instances in which multiple publications described the same system. Unless explicitly stated otherwise, reported counts and proportions refer to publications rather than unique systems.

Charted variables were organized into four analytical domains aligned with the research questions: (1) identification and classification of configurable or reusable EMA/EMI systems; (2) architectural structures, implementation strategies, and configurability mechanisms; (3) runtime operationalization of the assessment and intervention delivery, including triggering logic, personalization mechanisms, and adaptive decision processes; and (4) temporal patterns in the reporting of architectural approaches, configurability strategies, and system capabilities since 2009.

Synthesis combined structured tabulation with narrative comparison. Binary and categorical variables were summarized using descriptive counts and proportions and presented in domain-specific tables covering system mappings, study characteristics, EMA sampling strategies, intervention mechanisms, JITAI and AI-related features, engagement mechanisms, and selected software development characteristics.

Temporal patterns were examined descriptively using publication-year distributions and by comparing the occurrence of selected architectural and functional features across publication periods where appropriate. Aggregation and data consistency checks were supported by structured scripts. No inferential statistical testing, trend modeling, or machine learning methods were applied.

In addition to quantitative aggregation, qualitative comparison of charted variables was conducted to identify recurring architectural patterns, degrees of configurability and adaptivity, and common reporting practices. Particular attention was paid to whether technical components (e.g., frontend/backend stack, database technologies, cloud infrastructure, interoperability standards, and security or regulatory positioning) were explicitly reported. Variables coded as *not reported* were treated as missing descriptive information and were not interpreted as absence of implementation.

For contextual illustration of the identified design space, selected system exemplars were described narratively to represent distinct architectural and methodological archetypes (e.g., configurable platform frameworks, multimodal sensing infrastructures, healthcare-integrated systems, and experimentally optimized JITAI implementations). These exemplars served illustrative purposes only and did not influence quantitative summaries.

3 Results

3.1 Selection of sources of evidence

The database search identified 7,282 records. Prior to screening, 3,558 records were removed through automated filtering and deduplication. Automated tools excluded 1,745 records based on predefined criteria (publication date < 2009, n = 44; language, n = 25; publication type, n = 1,676). Additionally, 1,813 duplicate records were removed either by DOI-based deduplication (n = 1,700) or in two rounds of manual deduplication using the duplicate-detection function of `Rayyan` (n = 113). The remaining 3,724 records were subjected to title and abstract screening, of which 411 reports were sought for retrieval. The most common, non-mutually exclusive exclusion reasons were: `not app` (n = 2203), `not EMI` (n = 1625), `not EMA` (n = 1236), `wrong publication type` (n = 1229), `wrong publication focus` (n = 816), `wrong technology` (n = 109), and `only concept` (n = 33). Twelve reports could not be obtained, and 399 were assessed for eligibility by full-text review. Of the 399 assessed reports, 358 were excluded iteratively for the following reasons: not describing a native mobile EMA/EMI application (n = 110), lacking EMA features (n = 13), lacking EMI features (n = 20), lacking configurability characteristics (n = 156), lacking reusable system or

framework characteristics (n = 58) or just a concept without implementation (n = 1). Forty reports describing 34 distinct mHealth systems were included in the final dataset. Figure 1 shows the complete selection process according to the PRISMA statement[14].

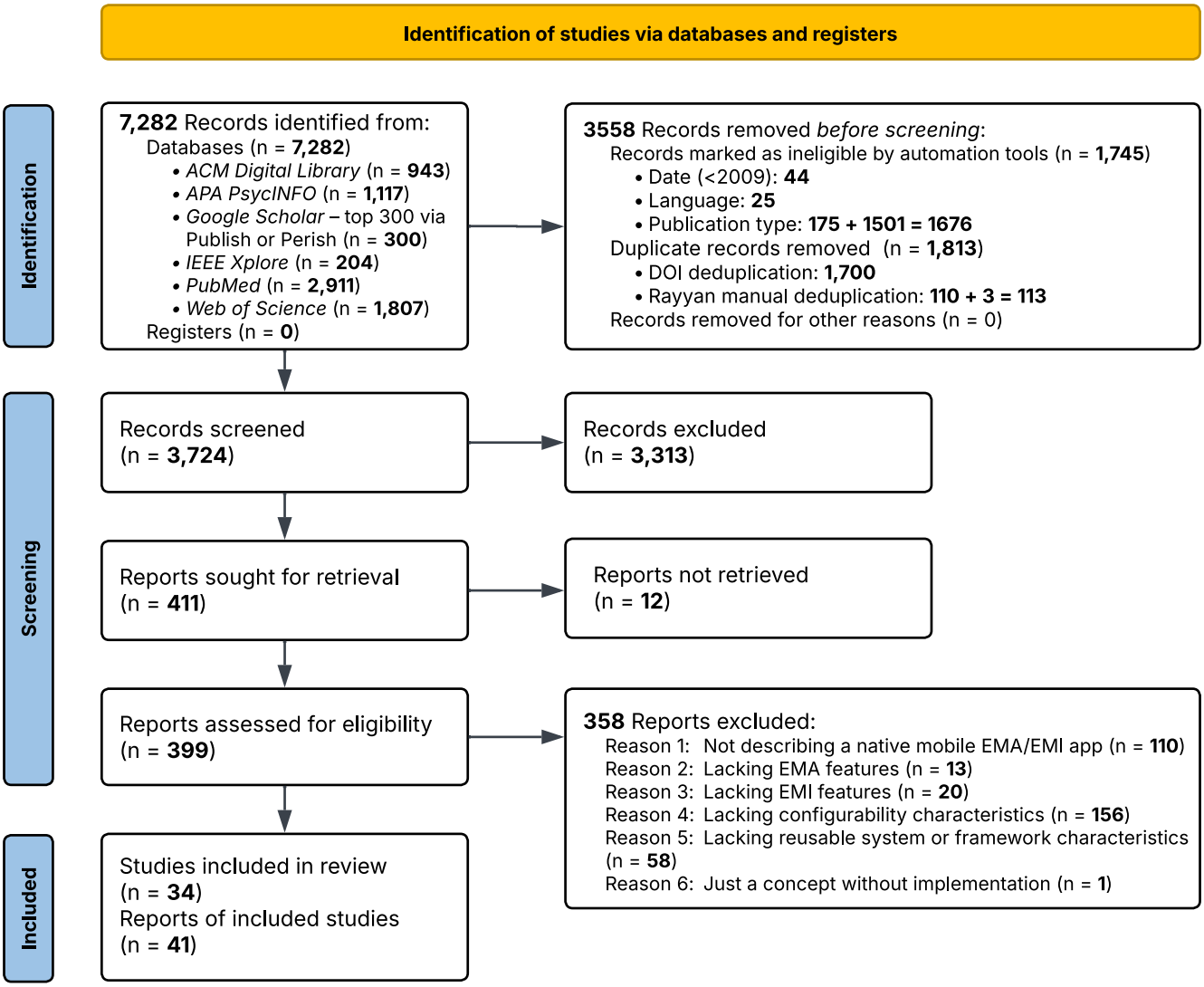


Figure 1. PRISMA 2020 flow diagram illustrating the complete study selection process.

3.2 Characteristics of sources of evidence

Publication years. The included publications spanned from 2014 to 2024, while the whole literature body screened in full-text reached from 2011 to July 11, 2024 (Figure 2). Included publications between 2014 and 2017 were more infrequent. Publication activity increased from 2018 onward and was concentrated between 2020 and 2024, with a peak in 2022. Because the search was conducted up to July 11, 2024, the count for 2024 likely underestimates the full-year publication volume.

Included systems and publication units. Across *41* included publications, *34* (82.93 %) distinct reusable systems/frameworks were identified. Seven (17.07 %) publications reported extensions, follow-up studies, or variants of previously described systems. Table 1 provides a comprehensive overview of the aforementioned items, along with the individual app names created using the identified systems, where applicable.

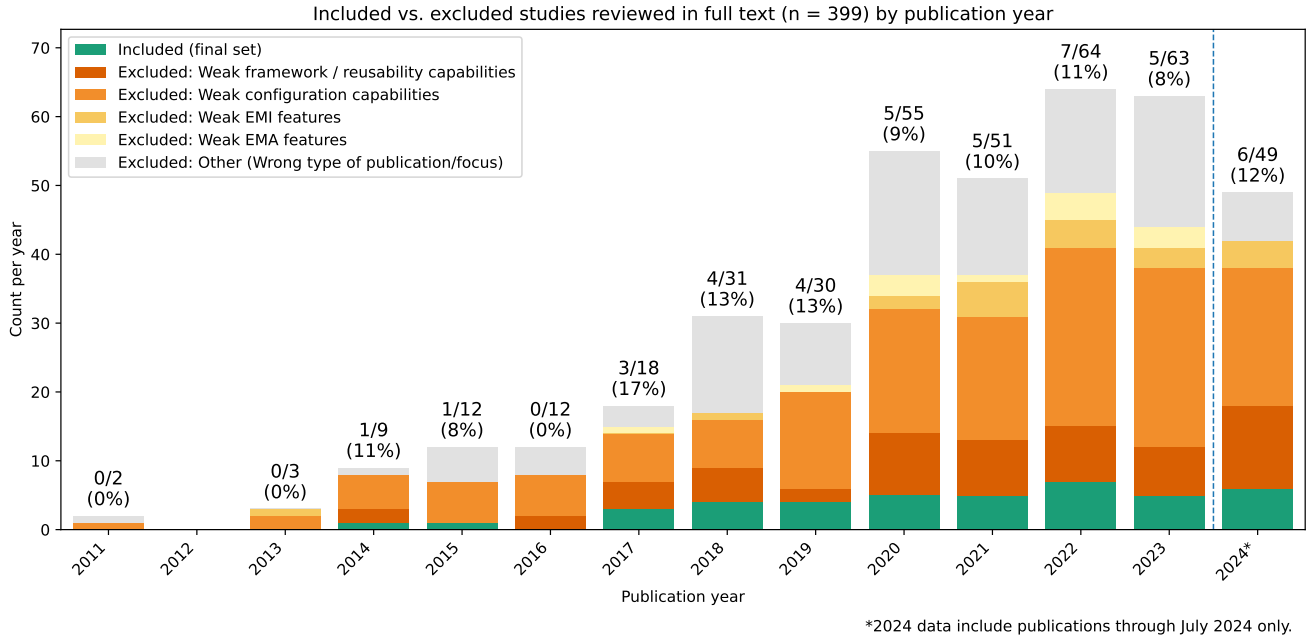


Figure 2. Distribution of included publications by year from 2011 to search conduction on July 11, 2024.

Table 1. Overview of included records.

Ref.	Year	Publication Title	System Name	Apps
[20]	2022	Enabling Patient Adherence via Personalised, Just-in Time Adaptive Interventions in ADLIFE Architecture	ADLIFE	Patient Empowerment Platform (PEP); Personalized Care Plan Management Platform (PCPMP); onFHIR Repository
[21]	2023	AwarNS: A framework for developing context-aware reactive mobile applications for health and mental health	AwarNS	SyMptOMS; TUG Test; SyMptOMS-ET
[22]	2022	Evaluating the Impact of Adaptive Personalized Goal Setting on Engagement Levels of Government Staff With a Gamified mHealth Tool: Results From a 2-Month Randomized Controlled Trial	GameBus	
[23]	2019	HealthSense: Software-defined Mobile-based Clinical Trials	HealthSense	HealthSense app
[24]	2024	Development and pre-pilot testing of STAMP + CBT: an mHealth app combining pain cognitive behavioral therapy and opioid support for patients with advanced cancer and pain	Insight™	STAMP + CBT; STAMP
[25]	2022	Development and preliminary effectiveness of a smartphonebased, just-in-time adaptive intervention for adults with alcohol misuse who are experiencing homelessness	Insight™	Smart-T Alcohol
[26]	2022	A Proof of Concept Pilot Examining Feasibility and Acceptability of the Positively Healthy Just-in-Time Adaptive, Ecological Momentary, Intervention Among a Sample of Sexual Minority Men Living with HIV	LifeData	RealLifeExp
[27]	2018	Smartphone-based ecological momentary assessment and intervention in a blended coping-focused therapy for distressing voices: Development and case illustration	LifeData	RealLife™Exp app
[28]	2021	Assessing the Impact of Engagement and Real-Time Feedback in a Mobile Behavior Change Intervention for Supporting Critical Thinking in Engineering Research Projects	LifeGuide Toolbox	CriticalThinking app

Ref.	Year	Publication Title	System Name	Apps
[29]	2020	A Smartphone App With a Digital Care Pathway for Patients Undergoing Spine Surgery: Development and Feasibility Study	ManageMySurgery	
[30]	2021	Remote Digital Psychiatry for Mobile Mental Health Assessment and Therapy: MindLogger Platform Development Study	MindLogger	MindLogger app
[31]	2019	Exploring the State-of-Receptivity for mHealth Interventions	MobileCoach	Ally app
[32]	2017	Enriching Mental Health Mobile Assessment and Intervention with Situation Awareness	MoodBuster	SituMan
[33]	2024	A Physical Activity and Diet Just-in-Time Adaptive Intervention to Reduce Blood Pressure: The myBPmyLife Study Rationale and Design	MyDataHelps	myBPmyLife app
[34]	2024	Is a brief mindfulness ecological momentary intervention more efficacious than a self-monitoring app for social anxiety disorder? A randomized controlled trial	PACO	
[35]	2022	A "No-Code" App Design Platform for Mobile Health Research: Development and Usability Study	Pathverse	Pathverse app
[36]	2020	An ecological momentary intervention incorporating personalised feedback to improve symptoms and social functioning in schizophrenia spectrum disorders	PsyMate™	SMARTapp (Schizophrenia Mobile Assessment and RealTime feedback application)
[37]	2014	Purple: A Modular System for Developing and Deploying Behavioral Intervention Technologies	Purple	Mobilyze; PC CHIP; Purple Robot; Onward; FOCUS
[38]	2023	Imager—A mobile health mental imagery-based ecological momentary intervention targeting reward sensitivity: A randomized controlled trial	SEMA3	Imager
[39]	2024	SEMA3: A free smartphone platform for daily life surveys	SEMA3	
[40]	2020	Personalized support for well-being at work: an overview of the SWELL project	SWELL	mBeats; NiceWork; Brightr
[41]	2021	Sense2Stop: A micro-randomized trial using wearable sensors to optimize a just-in-time-adaptive stress management intervention for smoking relapse prevention	Sense2Stop	???
[42]	2022	StudyU: A Platform for Designing and Conducting Innovative Digital N-of-1 Trials	StudyU	StudyU app
[43]	2018	A Novel Approach for Fully Automated, Personalized Health Coaching for Adults with Prediabetes: Pilot Clinical Trial	Sweetch	Sweetch app
[44]	2019	Technical Challenges to Deliver Sensor-based Psychological Interventions Using Smartphones	SyMptOMS	
[45]	2024	Challenges and opportunities in mobile e-coaching	TIIM	TIIM app
[46]	2017	ULTEMAT: A mobile framework for smart ecological momentary assessments and interventions	ULTEMAT	MoodBuster
[47]	2020	Low-Burden Mobile Monitoring, Intervention, and Real-Time Analysis Using the Wear-IT Framework: Example and Usability Study	Wear-IT	Wear-IT app
[48]	2023	A Highly Configurable EMA and JITAI Mobile App Framework Utilized in a Large-Scale German Study on Breast Cancer Aftercare	eSano	ACTonCancer app
[49]	2021	eSano – An eHealth Platform for Internet- and Mobile-based Interventions*	eSano	eSano app
[50]	2018	Development of a Just-in-Time Adaptive mHealth Intervention for Insomnia: Usability Study	iREST	iREST app
[51]	2024	An Ecological Mobile Momentary Intervention to Support Dynamic Goal Pursuit: Feasibility and Acceptability Study	m-Path	
[52]	2023	m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice	m-Path	m-Path app; m-Path Sense app

Ref.	Year	Publication Title	System Name	Apps
[53]	2017	mCerebrum: A Mobile Sensing Software Platform for Development and Validation of Digital Biomarkers and Interventions	mCerebrum	
[54]	2021	Project CHOICE: Choosing healthy options in coping with emotions, an EMA/EMI plus in-person intervention for alcohol use	mEMA	mEMA app
[55]	2018	Pilot randomized trial of MOMENT, a motivational counseling-plus-ecological momentary intervention to reduce marijuana use in youth	mEMA	mEMA app
[56]	2022	Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study	mindLAMP	mindLAMP smartphone app
[57]	2023	Digital Phenotyping Data to Predict Symptom Improvement and Mental Health App Personalization in College Students: Prospective Validation of a Predictive Model	mindLAMP	mindLAMP app
[58]	2019	Client experiences of blending a coping-focused therapy for auditory verbal hallucinations with smartphone-based ecological momentary assessment and intervention	movisensXS	movisensXS app
[59]	2020	schema: an open-source, distributed mobile platform for deploying mHealth research tools and interventions	schema	schema app
[60]	2015	A platform for user empowerment through Self Ecological Momentary Assessment / Intervention		

Note. **Apps**: Individual mobile health apps or systems named in the paper created with the corresponding framework..

Study designs, trials, and domains. Of the 41 included publications, 25 (60.97 %) reported original empirical evaluation results involving human participants, whereas 16 (39.02 %) were classified as framework or platform descriptions without presenting empirical outcomes (Table 2). On system level 21/34 (61.76 %) were empirically evaluated. Seven (17.07 %) of the framework-oriented publications described planned evaluations, including randomized controlled trials (4/41; 9.75 %), single-arm feasibility/pilot studies (1/41; 2.44 %), micro-randomized trials (1/41; 2.44 %) and large-scale observational deployments (1/41; 2.44 %), but no empirical results were reported within those papers.

Among the publications with empirical evaluation (n=25; 60.97 %), single-arm feasibility or pilot studies were most common (n=12; 29.27 %), followed by non-randomized controlled studies (n=7; 17.07 %) and randomized controlled trials (n=5; 12.19 %). One (2.44 %) publication reported a large-scale observational deployment, and two (4.88 %) studies required classification decisions due to mixed or partially unclear reporting. Single-arm feasibility and pilot studies represented the largest design category, whereas randomized and non-randomized controlled trials were less frequently reported.

Reported sample sizes varied substantially. Small exploratory deployments included as few as one to five participants, while several studies enrolled between 41 and 200 participants. A small number of large-scale implementations exceeded several thousand users[20,30,49,52,53], with the largest reported deployment involving more than 50,000 participants[52]. Most empirical studies investigated intervention groups only, whereas randomized and non-randomized controlled designs included explicit comparison groups.

Target populations were heterogeneous and spanned, non-exclusive, adolescents (3/41; 7.32 %), young adults (18/41; 43.9 %), adults (25/41; 60.97 %), and older adults (10/41; 24.39 %), with multiple publications explicitly targeting unrestricted populations (“everyone”; 10/41; 24.39 %) (Table 3). One (2.44 %) publication explicitly mentioned a family setting. Additional inclusion or exclusion criteria frequently involved health status (21/34; 61.76 %), language proficiency (10/41; 24.39 %), technical skills (11/41; 26.83 %), occupation (7/41; 17.07 %), education level (5/41; 12.19 %), gender (4/41; 9.75 %), ethnicity (1/41; 2.44 %), or socioeconomic characteristics (2/41; 4.88 %).

Regarding application domains, the majority of systems were applied in mental health (21/34; 61.76 %) and/or physical health (17/34; 50 %) contexts, with several platforms supporting both. Mental health applications commonly addressed depression (8/34; 23.53 %), anxiety (7/34; 20.59 %), mood (10/34; 29.41 %), stress (11/34; 32.35 %), addiction (5/34; 14.7 %), and schizophrenia (2/34; 5.88 %). Physical health applications included cardiovascular health (4/34; 11.76 %), diabetes (1/34; 2.94 %), sleep (4/34; 11.76 %), diet and nutrition (7/34; 20.59 %), physical activity (10/34; 29.41 %), rehabilitation (1/34; 2.94 %), and chronic disease management (1/34; 2.94 %). A subset of systems (15/34; 44.12 %) was presented as generic or multi-domain frameworks without restriction to a single health or research area, enabling deployment across diverse intervention contexts.

Table 2. Study Design, Focus and User Numbers

Meta Info		Study Design and Focus				User numbers				
	System Name	Empirical	Design	Focus	Planned/ ongoing	Total	Intervention	Control	Non Patient	Other
[20]	ADLIFE	No	Framework description only	System architecture / platform design	Single-arm feasibility/pilot	2702	882		1243 + 577	
[21]	AwarNS	No	Framework description only	System architecture / platform design						
[22]	GameBus	Yes	Randomized controlled trial	Intervention effectiveness		176	84	82 + 10		
[23]	HealthSense	Yes	Controlled study (non-randomized)	Both		58	26 + 8 + 11	13		
[24]	Insight™	Yes	Single-arm feasibility/pilot	Both		43	14 + 14 + 15			
[25]	Insight™	Yes	Single-arm feasibility/pilot	Both		41	41			
[26]	LifeData	Yes	Single-arm feasibility/pilot	Both		22	22			
[27]	LifeData	Yes	Single-arm feasibility/pilot	Both		1	1			
[28]	LifeGuide Toolbox	Yes	Randomized controlled trial	Both		60	30	30		
[29]	ManageMySurgery	Yes	Single-arm feasibility/pilot	Both		47	47			
[30]	MindLogger	Yes	Large-scale observational deployment	System architecture / platform design		4315	4315			
[31]	MobileCoach	Yes	Controlled study (non-randomized)	Both		189	189			
[32]	MoodBuster	Yes	Single-arm feasibility/pilot	Both		21	9 + 12			
[33]	MyDataHelps	No	Framework description only	System architecture / platform design	Randomized controlled trial	484	242	242		
[34]	PACO	Yes	Randomized controlled trial	Intervention effectiveness		191	96	95		
[35]	Pathverse	Yes	Single-arm feasibility/pilot	Both		5	5			
[36]	PsyMate™	Yes	Controlled study (non-randomized)	Intervention effectiveness		50	23 + 27			
[37]	Purple	No	Framework description only	System architecture / platform design						
[38]	SEMA3	Yes	Randomized controlled trial	Intervention effectiveness		99	52	47		
[39]	SEMA3	No	Framework description only	System architecture / platform design						
[40]	SWELL	Yes	Controlled study (non-randomized)	Both		98	35			15 + 6 + 22 + 20
[41]	Sense2Stop	No	Framework description only	System architecture / platform design	Micro-randomized trial	75	75			
[42]	StudyU	No	Framework description only	System architecture / platform design						
[43]	Sweetch	Yes	Single-arm feasibility/pilot	Both		55	55			
[44]	SyMptOMS	No	Framework description only	System architecture / platform design						
[45]	TIIM	No	Framework description only	System architecture / platform design						
[46]	ULTEMAT	No	Framework description only	System architecture / platform design	Randomized controlled trial					

Meta Info		Study Design and Focus				User numbers				
	System Name	Em-pir-ical	Design	Focus	Planned/ ongoing	Total	Interven-tion	Control	Non Patient	Other
[47]	Wear-IT	Yes	Controlled study (non-randomized)	System architecture / platform design		5			5	5
[48]	eSano	No	Framework description only	System architecture / platform design						
[49]	eSano	No	Framework description only	System architecture / platform design	Randomized controlled trial	975	900		75	
[50]	iREST	Yes	Single-arm feasibility/pilot	Both		19	19			
[51]	m-Path	Yes	Single-arm feasibility/pilot	Both		73	73			
[52]	m-Path	Yes	Single-arm feasibility/pilot	System architecture / platform design		53500	50000		3500	
[53]	mCerebrum	No	Framework description only	System architecture / platform design	Large-scale observational deployment	2251	2251			
[54]	mEMA	Yes	Single-arm feasibility/pilot	Both		20	20			
[55]	mEMA	Yes	Randomized controlled trial	Both	Randomized controlled trial	70	28 + 27	15		
[56]	mindLAMP	No	Framework description only	System architecture / platform design						
[57]	mindLAMP	Yes	Controlled study (non-randomized)	Both		66	22 + 22	22		
[58]	movisensXS	Yes	Controlled study (non-randomized)	Intervention effectiveness		17	17			
[59]	schema	No	Framework description only	System architecture / platform design						
[60]	nan	No	Framework description only	System architecture / platform design						

Note. **Empirical**: Indicates whether the publication reports original empirical evaluation results involving human participants within the paper. Studies referencing prior evaluations without presenting original methods or results, as well as protocol, planned, or ongoing studies without outcomes, were classified as framework descriptions without empirical evaluation.; **Design**: Primary empirical evaluation design reported in the publication. Categories distinguish between single-arm feasibility/pilot studies, non-randomized controlled studies, randomized controlled trials (RCTs), micro-randomized trials, and large-scale observational deployments. Framework descriptions without empirical evaluation were coded separately.; **Focus**: Primary focus of the publication, categorized as system architecture/platform design, intervention effectiveness, or both when substantial emphasis was placed on technical development and outcome evaluation.; **Planned/ ongoing**: Planned or protocol-based study design reported in the publication when no empirical results were presented, reflecting the intended evaluation methodology described by the authors.; **Total**: Total number of participants across all groups; **Interven-tion**: Number of participants assigned to an intervention group; **Control**: Number of participants assigned to a control group; **Non Patient**: Number of non-patient participants (e.g. healthcare professionals, eCoaches, content managers and other system users) in an intervention group; **Other**: Number of participants in other groups (e.g., focus groups or interviews).

Table 3. Domain, Target Population and Additional Factors

Meta Info		Domain			
	System Name	Specific	Overall	Target Popula-tion	Additional Factors
[20]	ADLIFE	Cardiovascular health, Chronic disease management (general), Diet & nutrition, Physical activity & exercise	generic, Physical health	elderly, young adults	health status
[21]	AwarNS		generic		
[22]	GameBus	Physical activity & exercise	Physical health	adults, elderly, young adults	location, occupation

Meta Info		Domain			
	System Name	Specific	Overall	Target Population	Additional Factors
[23]	HealthSense	Addiction cessation, Chronic disease management (general), Depression, Diet & nutrition, Emotion, Mood	generic, Mental health, Physical health	adolescen adults, other, young adults	health status
[24]	Insight™	Cancer care, Emotion, Physical health, Sleep, Stress	Mental health, Physical health	adults, elderly, young adults	health status, location, technical skills
[25]	Insight™	Addiction cessation, Alcohol Use Disorder (AUD)	Mental health, Physical health	adults, elderly, young adults	other, socioeconomic status
[26]	LifeData	Sexual & reproductive health	Physical health	adults	gender, health status, location, other
[27]	LifeData	Chronic disease management (general), Emotion	Mental health	adults	health status, location, technical skills
[28]	LifeGuide Toolbox	Behavior Change	Education	young adults	education, location, occupation, technical skills
[29]	ManageMySurgery	Anxiety, Physical health, Rehabilitation / physiotherapy	Mental health	adults, elderly, young adults	language, technical skills
[30]	MindLogger	Cognitive/neurological disorders	Mental health	everyone	health status
[31]	MobileCoach	Physical activity & exercise	Physical health	adults	health status, language
[32]	MoodBuster	Depression, Mood	Mental health		health status, language, occupation, technical skills
[33]	MyDataHelps	Cardiovascular health, Diet & nutrition, Physical activity & exercise	Physical health	adults	health status, language, location
[34]	PACO	Anxiety, Depression; Mood, Stress	Mental health	adults, young adults	education, health status, occupation, socioeconomic status
[35]	Pathverse		generic	adults	health status, technical skills
[36]	PsyMate™	Schizophrenia	Mental health	adults, elderly, young adults	health status, other
[37]	Purple	Anxiety, Cancer care, Depression, Emotion, Mood, Schizophrenia, Stress	Mental health, Physical health	adults	
[38]	SEMA3	Depression, Mood, Stress	Mental health	young adults	education, health status, language, location
[39]	SEMA3	Anxiety, Diet & nutrition, Emotion, Physical activity & exercise, Sexual & reproductive health, Stress	generic, Mental health, Physical health	adolescen adults, elderly, every-one, family setting, young adults	gender, health status, occupation
[40]	SWELL	Cardiovascular health, Physical activity & exercise, Stress	Mental health, Physical health	adults	occupation
[41]	Sense2Stop	Addiction cessation, Emotion, Stress	Mental health	adults	health status, language, location

Meta Info		Domain			
	System Name	Specific	Overall	Target Population	Additional Factors
[42]	StudyU		generic	everyone	
[43]	Sweetch	Diabetes, Diet & nutrition, Physical activity & exercise	Physical health	adults, elderly, young adults	education, ethnicity, gender, health status, language, technical skills
[44]	SyMptOMS		generic, Mental health	everyone	
[45]	TIIM	Emotion, Mood, Physical activity & exercise, Stress	Mental health, Physical health	everyone	
[46]	ULTEMAT	Depression, Mood	generic, Mental health	adults	health status, location
[47]	Wear-IT		generic	adults	
[48]	eSano	Cancer care, Emotion, Mood, Stress	generic, Mental health	adults	health status
[49]	eSano	Anxiety, Chronic disease management (general), Depression	generic, Mental health	adults	
[50]	iREST	Sleep	Physical health	adults, elderly, young adults	health status, occupation, technical skills
[51]	m-Path	Emotion, Goal Pursuit, Mood	Mental health	adults, young adults	gender, health status, language, location
[52]	m-Path	Emotion, Mood, Physical activity & exercise	generic, Mental health, Physical health	everyone	
[53]	mCerebrum	Addiction cessation, Cardiovascular health, Diet & nutrition Physical activity & exercise, Stress	generic, Mental health, Physical health	everyone	
[54]	mEMA	Addiction cessation, Alcohol Use Disorder (AUD), Anxiety, Depression, Emotion, Mood, Stress	Mental health	young adults	health status, technical skills
[55]	mEMA	Addiction cessation, Emotion	Mental health	adolescen young adults	health status, language, location
[56]	mindLAMP		generic	everyone	
[57]	mindLAMP	Anxiety, Depression, Sleep, Stress	Mental health, Physical health	adults, young adults	education, health status, technical skills
[58]	movisensXS	Chronic Disease Management (General):	Mental health	adults, elderly, young adults	health status, language, technical skills
[59]	schema		generic	everyone	
[60]	nan	Diet & nutrition, Mood, Sleep	generic, Physical health	everyone	

Note. **Specific**: Application domain(s) in which the app is used; multiple selections possible including `generic framework` and `specific domain`; **Target Population**: Reported target age group or population as described by the study authors; reference ranges: `infants` (0–1), `children` (2–12), `adolescents` (10–19), `young adults` (18–24), `adults` (25–64), `older adults` (65+); ranges may overlap; `family setting` and `healthcare staff setting` coded as separate binary variables; `everyone` indicates no specific age restriction; multiple selections or custom age ranges possible; **Additional Factors**: Additional inclusion or exclusion characteristics beyond age and intervention condition.

EMA sampling characteristics. EMA functionality was predominantly implemented through `signal-based prompting` and reported in 40/41 (97.56 %) publications with 33/34 (97.06 %) individual systems supporting scheduled fixed or random assessments (Table 4). Event-contingent sampling was reported in 34/41 (82.93 %) publications for 29/34 (85.29 %) individual systems, while 26/41 (63.41 %) publications with 24/34 (70.59 %) were reported to support interval-based data collection. Context-aware triggering mechanisms were described in 18/41 (43.9 %) publications for 17/34 (50 %) systems. Passive sensor-based data collection without active user input was reported in 24/41

(58.54 %) publications and implemented in 23/34 (67.65 %) systems. Overall, most platforms combined multiple sampling strategies rather than relying on a single approach, indicating support for hybrid EMA configurations.

Table 4. EMA methods supported by the system (x = present).

Meta Info		EMA Sampling Strategies				
Ref.	System Name	Context	Signal	Event	Interval	Passive
[20]	ADLIFE	x	x	x		x
[21]	AwarNS	x	x	x	x	x
[22]	GameBus			x	x	x
[23]	HealthSense	x	x	x	x	x
[24]	Insight™		x	x	x	
[25]	Insight™		x	x		
[26]	LifeData		x	x		
[27]	LifeData		x	x	x	x
[28]	LifeGuide Toolbox		x	x		x
[29]	ManageMySurgery	x	x	x	x	
[30]	MindLogger		x	x		
[31]	MobileCoach	x	x		x	x
[32]	MoodBuster	x	x			
[33]	MyDataHelps	x	x	x	x	x
[34]	PACO		x	x		
[35]	Pathverse		x	x	x	x
[36]	PsyMate™		x			
[37]	Purple	x	x	x		x
[38]	SEMA3		x	x		
[39]	SEMA3		x	x	x	
[40]	SWELL	x	x	x	x	x
[41]	Sense2Stop	x	x	x		x
[42]	StudyU	x	x	x	x	x
[43]	Sweetch	x	x		x	x
[44]	SyMptOMS		x	x		x
[45]	TIIM		x	x	x	x
[46]	ULTEMAT	x	x	x	x	x
[47]	Wear-IT	x	x	x	x	x
[48]	eSano		x	x	x	
[49]	eSano		x	x		
[50]	iREST		x	x	x	x
[51]	m-Path	x	x			
[52]	m-Path	x	x	x	x	x
[53]	mCerebrum	x	x	x	x	x
[54]	mEMA		x	x	x	
[55]	mEMA		x	x	x	
[56]	mindLAMP		x	x	x	x
[57]	mindLAMP		x		x	x
[58]	movisensXS		x	x		
[59]	schema		x		x	
[60]		x	x	x	x	x
Total (Record Level)		18/41 (43.9 %)	40/41 (97.56 %)	34/41 (82.93 %)	26/41 (63.41 %)	24/41 (58.54 %)
Total (System Level)		17/34 (50 %)	33/34 (97.06 %)	29/34 (85.29 %)	24/34 (70.59 %)	23/34 (67.65 %)

Note. **Context**: Entry triggered by a context-aware notification tailored to the participant's situation; **Signal**: Entry prompted at scheduled fixed or random times; **Event**: Entry triggered by a predefined event or self-initiated by the participant; **Interval**: Data collected at fixed time intervals (manual or automated); **Passive**: Passive continuous (sensor-based) data collection without user interaction.

Intervention characteristics. Static intervention content was widely supported across platforms (Table 5). Advice was the most frequently reported static component in 35/41 (85.37 %) publications for 28/34 (82.35 %) systems), followed by reminders reported in 33/41 (80.49 %) publications for 28/34 (82.35 %) systems and educational material reported in 24/41 (58.54 %) publications for 21/34 (70.59 %) systems.

Dynamic intervention functionality was common. Automated intervention delivery without provider involvement was reported in 39/41 (95.12 %) publications for 32/34 (94.12 %), whereas manual provider-triggered intervention features were described in 12/41 (29.27 %) publications for 12/34 (35.29 %) systems. Context-triggered dynamic intervention content was described in 21/41 (51.22 %) publications and implemented in 19/34 (55.88 %) systems. Personalized messages based on predefined user data (e.g., profile information or baseline assessments) were described in 34/41 (82.93 %) publications for 27/34 (79.41 %) systems.

Table 5. Static and dynamic intervention methods (x = present).

Meta Info		Static Intervention Content			Dynamic Intervention Content			
Ref.	System Name	Education	Advice	Reminders	Automated	Manual	Context	Pers. Msg.
[20]	ADLIFE	x	x	x	x			x
[21]	AwarNS	x	x	x	x		x	x
[22]	GameBus				x			x
[23]	HealthSense		x	x	x		x	x
[24]	Insight™	x	x	x	x			x
[25]	Insight™	x	x	x	x			x
[26]	LifeData	x	x		x			x
[27]	LifeData		x	x	x	x		x
[28]	LifeGuide Toolbox	x	x	x	x	x		x
[29]	ManageMySurgery	x	x	x	x	x		x
[30]	MindLogger	x	x	x	x	x		x
[31]	MobileCoach	x	x	x	x		x	x
[32]	MoodBuster		x		x		x	
[33]	MyDataHelps	x	x	x	x		x	x
[34]	PACO	x	x		x			x
[35]	Pathverse	x		x	x	x		
[36]	PsyMate™		x	x				x
[37]	Purple	x	x	x	x	x	x	x
[38]	SEMA3	x	x	x	x			x
[39]	SEMA3		x	x	x			x
[40]	SWELL	x	x	x	x		x	x
[41]	Sense2Stop			x	x		x	
[42]	StudyU	x	x	x	x	x	x	x
[43]	Sweetch		x	x	x		x	x
[44]	SyMptOMS				x		x	
[45]	TIIM	x	x	x	x			x
[46]	ULTEMAT		x	x	x	x	x	
[47]	Wear-IT		x		x		x	x
[48]	eSano	x	x	x	x		x	x
[49]	eSano	x	x	x	x	x		x
[50]	iREST	x	x	x	x	x		x
[51]	m-Path		x	x	x		x	x
[52]	m-Path	x	x	x	x	x	x	x
[53]	mCerebrum		x	x	x		x	
[54]	mEMA		x		x			x
[55]	mEMA		x		x		x	x
[56]	mindLAMP	x	x	x	x		x	x
[57]	mindLAMP	x	x	x	x		x	x
[58]	movisensXS			x		x		

Meta Info		Static Intervention Content			Dynamic Intervention Content			
Ref.	System Name	Education	Advice	Reminders	Automated	Manual	Context	Pers. Msg.
[59]	schema	x	x	x	x			x
[60]				x	x		x	x
Total (Record Level)		24/41 (58.54 %)	35/41 (85.37 %)	33/41 (80.49 %)	39/41 (95.12 %)	12/41 (29.27 %)	21/41 (51.22 %)	34/41 (82.93 %)
Total (System Level)		21/34 (61.76 %)	28/34 (82.35 %)	28/34 (82.35 %)	32/34 (94.12 %)	12/34 (35.29 %)	19/34 (55.88 %)	27/34 (79.41 %)

Note. **Education**: Static intervention content providing educational material (e.g., articles); **Advice**: Static intervention content providing general advice or feedback; **Reminders**: Static intervention content delivering fixed or timed reminders; **Automated**: Dynamic intervention features delivered automatically without involvement of a health care provider; **Manual**: Dynamic intervention features reviewed or triggered manually by a health care provider; **Context**: Dynamic intervention content triggered by contextual data (e.g., location or sensor information); **Pers. Msg.**: Dynamic intervention messages tailored using predefined personal user data (e.g., profile or baseline questionnaire).

JITAI and AI components. Defined triggers for just-in-time adaptive intervention (JITAI) delivery were reported in 29/41 (70.73 %) publications for 26/34 (76.47 %) systems (Table 6). Automated JITAI modes were described in 16/41 (39.02 %) publications for 16/34 (47.06 %) systems, while manual JITAI modes involving provider review were reported in 5/41 (12.19 %) publications for 5/34 (14.7 %) systems. Adaptable intervention content was reported in 27/41 (65.85 %) publications implemented in 24/34 (70.59 %) systems, and 21/41 (51.22 %) publications reported context-aware intervention delivery for 20/34 (58.82 %) systems. Explicit just-in-time delivery at the moment of an identified need was described in 19/41 (46.34 %) publications for 19/34 (55.88 %) systems.

Artificial intelligence components were reported less frequently. AI-generated static intervention content was described in 10/41 (24.39 %) publications for 10/34 (29.41 %) systems, dynamic AI-generated content in 7/41 (17.07 %) publications for 7/34 (20.59 %) systems, and AI-supported intervention decision logic in 4/41 (9.75 %) publications for 4/34 (11.76 %) systems.

Table 6. JITAI methods (column 4-8) and AI in intervention design (column 9-11); x = present.

Meta Info		JITAI	JITAI Modes		JITAI Content			AI Generated Content		AI Logic
Ref.	System Name	Triggers	Manual	Automated	Adaptable	Context Aware	JIT	Static	Dynamic	Decisions
[20]	ADLIFE	x		x	x	x	x	x		
[21]	AwarNS	x			x	x	x	x	x	
[22]	GameBus									
[23]	HealthSense	x		x	x	x	x			
[24]	Insight™	x		x	x	x	x			
[25]	Insight™	x			x					
[26]	LifeData	x			x		x	x		
[27]	LifeData									
[28]	LifeGuide Toolbox	x			x		x	x		
[29]	ManageMySurgery									
[30]	MindLogger	x			x		x			
[31]	MobileCoach	x		x	x	x	x		x	
[32]	MoodBuster	x		x		x	x			
[33]	MyDataHelps	x		x	x	x	x			
[34]	PACO									
[35]	Pathverse									
[36]	PsyMate™					x		x		
[37]	Purple	x	x	x	x	x	x	x		x
[38]	SEMA3	x			x					
[39]	SEMA3							x		
[40]	SWELL	x			x	x				
[41]	Sense2Stop	x		x	x	x	x		x	x
[42]	StudyU	x	x	x	x	x	x			

Meta Info		JITAI	JITAI Modes		JITAI Content			AI Generated Content		AI Logic
Ref.	System Name	Triggers	Manual	Automated	Adaptable	Context Aware	JIT	Static	Dynamic	Decisions
[43]	Sweetch	x		x	x	x	x		x	
[44]	SyMptOMS									
[45]	TIIM	x		x	x					
[46]	ULTEMAT									
[47]	Wear-IT	x		x	x	x	x		x	x
[48]	eSano	x	x	x	x					
[49]	eSano									
[50]	iREST	x	x		x					
[51]	m-Path	x		x	x	x	x			
[52]	m-Path	x			x	x				
[53]	mCerebrum	x	x	x	x	x	x	x	x	x
[54]	mEMA	x			x		x			
[55]	mEMA	x				x				
[56]	mindLAMP	x			x	x	x	x		
[57]	mindLAMP			x	x				x	
[58]	movisensXS									
[59]	schema	x			x	x				
[60]		x				x		x		
Total (Record Level)		29/41 (70.73 %)	5/41 (12.19 %)	16/41 (39.02 %)	27/41 (65.85 %)	21/41 (51.22 %)	19/41 (46.34 %)	10/41 (24.39 %)	7/41 (17.07 %)	4/41 (9.75 %)
Total (System Level)		26/34 (76.47 %)	5/34 (14.7 %)	16/34 (47.06 %)	24/34 (70.59 %)	20/34 (58.82 %)	19/34 (55.88 %)	10/34 (29.41 %)	7/34 (20.59 %)	4/34 (11.76 %)

Note. **Triggers**: Defined triggers used to initiate interventions; **Manual**: JITAI features manually reviewed or initiated by a health care provider; **Automated**: JITAI features delivered automatically without provider involvement; **Adaptable**: Intervention content adapts based on user state or interaction; **Context Aware**: Intervention content adapts based on contextual data (e.g., location, sensors, or user input); **JIT**: Intervention content delivered at the moment of identified need using real-time data; **Static**: Intervention content generated during development using AI but not adapted during use; **Dynamic**: Intervention content dynamically generated or adapted in the app using AI based on user interaction; **Decisions**: AI used to support intervention logic or decision-making within the intervention.

Engagement and adherence features. Engagement support was primarily providced through push notifications reported in all publications for all systems, (Table 7). Local device-scheduled notifications were reported in 32/41 (78.05 %) publications for 29/34 (85.29 %) systems, while 25/41 (60.97 %) publications for 23/34 (67.65 %) systems reported support for remotely triggered notifications via a server component. In 12/41 (29.27 %) publications for 12/34 (35.29 %) systems, push notifications were explicitly mentioned, but the delivery mechanism was not specified clearly enough to allow definite classification for one or both aforementioned delivery methods. Categories were not mutually exclusive, as some systems supported multiple notification mechanisms or provided insufficient technical detail for unambiguous categorization.

In-app reminder functionality was described in 22/41 (53.66 %) publications for 20/34 (58.82 %) systems, while 9/41 (21.95 %) publications with 8/34 (23.53 %) systems included task lists or to-do overviews without active alerts. External notifications (e.g., SMS or email) were reported in 6/41 (14.63 %) publictions for 6/34 (17.65 %) systems, and only 1/41 (2.44 %) publications for 1/34 (2.94 %) systems explicitly stated support for alerts directed to other persons such as healthcare providers[24].

Gamification features were comparatively uncommon. Partial game elements were reported in 4/41 (9.75 %) publications for 4/34 (11.76 %) systems, competitive mechanics in 1/41 (2.44 %) publications for 1/34 (2.94 %) system, milestone-based reward structures in 6/41 (14.63 %) publications for 5/34 (14.7 %) systems, and non-competitive progress tracking in 10/41 (24.39 %) publications for 10/34 (29.41 %) systems. No purely game-based interventions were reported.

Visualization features were more prevalent. User-facing visual dashboards or progress indicators implementations were reported in 26/41 (63.41 %) publications for 24/34 (70.59 %) systems, and 27/41 (65.85 %) publications with 25/34 (73.53 %) individual systems reported dashboards or monitoring widget support intended for researchers or system administrators.

Table 7. Engagement and adherence methods (x = present).

	Meta Info	Push Notification			Reminders		Not In-App Notification		Gamification				Visual. Widgets	
	System Name	?	Local	Remote	In-App	Todos	External	To Others	Part. Game	Compe- tition	Mile- stones	Pro- gress	User	Researchers
[20]	ADLIFE			x	x							x	x	x
[21]	AwarNS		x											
[22]	GameBus			x	x	x			x	x	x		x	x
[23]	HealthSense		x	x	x		x							x
[24]	Insight™		x	x				x	x				x	x
[25]	Insight™		x											
[26]	LifeData			x										
[27]	LifeData		x		x									x
[28]	LifeGuide Toolbox		x	x								x		x
[29]	ManageMySurgery	x		x	x	x							x	x
[30]	MindLogger		x	x					x				x	x
[31]	MobileCoach	x	x		x							x	x	
[32]	MoodBuster		x		x								x	x
[33]	MyDataHelps		x	x			x				x	x	x	
[34]	PACO		x	x										x
[35]	Pathverse		x	x	x						x		x	x
[36]	PsyMate™		x	x										
[37]	Purple		x	x	x	x	x					x	x	x
[38]	SEMA3		x	x	x								x	x
[39]	SEMA3		x	x									x	x
[40]	SWELL	x	x		x						x	x	x	x
[41]	Sense2Stop	x	x		x									x
[42]	StudyU		x		x	x						x	x	x
[43]	Sweetch	x	x		x	x						x	x	
[44]	SyMptOMS			x										x
[45]	TIIM	x	x									x	x	
[46]	ULTEMAT		x	x	x	x			x				x	x
[47]	Wear-IT		x	x									x	x
[48]	eSano		x	x	x								x	x
[49]	eSano	x			x		x							x
[50]	iREST	x	x	x	x	x							x	x
[51]	m-Path		x		x						x		x	
[52]	m-Path	x		x	x						x	x	x	x
[53]	mCerebrum	x	x		x								x	x
[54]	mEMA				x									
[55]	mEMA	x	x				x							x
[56]	mindLAMP		x	x		x							x	
[57]	mindLAMP		x	x		x	x							
[58]	movisensXS	x	x											
[59]	schema			x									x	x
[60]	nan		x	x									x	
	Total (Record Level)	12/41 29.27 %	32/41 78.05 %	25/41 60.97 %	22/41 53.66 %	9/41 21.95 %	6/41 14.63 %	1/41 2.44 %	4/41 9.75 %	1/41 2.44 %	6/41 14.63 %	10/41 24.39 %	26/41 63.41 %	27/41 65.85 %
	Total (System Level)	12/34 35.29 %	29/34 85.29 %	23/34 67.65 %	20/34 58.82 %	8/34 23.53 %	6/34 17.65 %	1/34 2.94 %	4/34 11.76 %	1/34 2.94 %	5/34 14.7 %	10/34 29.41 %	24/34 70.59 %	25/34 73.53 %

Note. **?**: Push notifications are mentioned but delivery mechanism not specified; **Local**: Notifications scheduled locally by the app on the user's device; **Remote**: Notifications triggered remotely via a server component; **In-App**: Notifications displayed only within the app interface during use; **Todos**: Task lists or overviews visible within the app without active alerts; **External**: Notifications sent outside the app (e.g., SMS, email, or calls); **To Others**: Alerts or warnings sent to other persons such as healthcare providers; **Part. Game**: Intervention includes game elements but is not entirely game-based (no fully game-based interventions found for this list); **Competition**: Gamification using competition or point-based mechanics (e.g., leaderboards or scores); **Milestones**: Gamification using progress rewards (e.g., badges, trophies, or milestones); **Progress**: Non-competitive progress tracking without reward mechanics; **User**: Visualizations shown to users within the app (e.g., charts or progress indicators); **Researchers**: Dashboards or widgets intended for researchers or system maintainers.

Software development characteristics. Interdisciplinary collaboration between technical and clinical or research domains was frequently addressed. Interdisciplinary approaches were explicitly described in detail in 11/41 (26.83 %) publications and mentioned more briefly in 27/41 (65.85 %) publications. Two (4.88 %) publications implied interdisciplinary collaboration without detailed description, and only one (2.44 %) publication did not refer to interdisciplinary aspects.

With regard to the development approach, 12/41 (29.27 %) publications described the systems as in-house developed full systems and 4/41 (9.75 %) as in-house developed prototypes. Thirteen (31.71 %) publications referred to open-source systems, whereas 10/41 (24.39 %) described paid or licensed software solutions. One (2.44 %) publications referred to a publicly available third-party application, and in 1/41 (2.44 %) cases the development approach remained unclear.

Software licenses were explicitly reported in 9/41 (21.95 %) publications, where they included permissive and copyleft licenses such as the Apache License 2.0, MIT License, BSD 2- and 3-Clause Licenses, GNU General Public License v3, and other attribution-based open-source licenses; however, most publications did not explicitly state licensing terms.

Interoperability or integration with other systems was mentioned in 25/41 (60.97 %) publications, although specific standards or technical mechanisms were mostly not detailed. Only 3/41 (7.32 %) publications explicitly referred to established interoperability standards such as `HL7 FHIR`.

All systems supported smartphone-based deployment by inclusion criteria. Beyond smartphones, additional explicitly included platforms were tablets in 16/34 (47.06 %) and web browser-based interfaces in 20/34 (58.82 %) systems, as well as smartwatches in 7/34 (20.59 %), other wearables 12/34 (35.29 %), desktop access 5/34 (14.7 %), and, in several cases custom hardware or stationary sensor components 7/34 (20.59 %), see Table 8. In most cases, it is not specified whether the web browser-based interface supports only admins and healthcare providers or also participants. The incorporation of glasses and similar technologies was not reported.

Table 8. Supported devices (x = present).

	Meta Info							
Ref.	System Name	other wearables	smartwatch	glasses	desktop	custom hardware	tablet	web browser
[20]	ADLIFE					x	x	x
[21]	AwarNS						x	
[22]	GameBus	x						x
[23]	HealthSense	x	x			x	x	x
[24]	Insight™							
[25]	Insight™							
[26]	LifeData						x	
[27]	LifeData							
[28]	LifeGuide Toolbox							x
[29]	ManageMySurgery						x	x
[30]	MindLogger						x	x
[31]	MobileCoach							
[32]	MoodBuster						x	
[33]	MyDataHelps	x	x					
[34]	PACO							
[35]	Pathverse	x					x	x
[36]	PsyMate™							
[37]	Purple	x			x	x	x	x
[38]	SEMA3							
[39]	SEMA3						x	x

Meta Info								
Ref.	System Name	other wearables	smartwatch	glasses	desktop	custom hardware	tablet	web browser
[40]	SWELL	x	x		x	x	x	x
[41]	Sense2Stop	x				x		
[42]	StudyU				x		x	x
[43]	Sweetch					x		
[44]	SyMptOMS	x	x					x
[45]	TIIM	x						x
[46]	ULTEMAT						x	x
[47]	Wear-IT	x	x					
[48]	eSano						x	x
[49]	eSano				x		x	x
[50]	iREST	x			x			x
[51]	m-Path		x					x
[52]	m-Path							x
[53]	mCerebrum	x	x			x		
[54]	mEMA							x
[55]	mEMA							
[56]	mindLAMP						x	x
[57]	mindLAMP							
[58]	movisensXS							
[59]	schema						x	
[60]								x
Total (Record Level)		12/41 (29.27 %)	7/41 (17.07 %)	0/41 (0 %)	5/41 (12.19 %)	7/41 (17.07 %)	17/41 (41.46 %)	22/41 (53.66 %)
Total (System Level)		12/34 (35.29 %)	7/34 (20.59 %)	0/34 (0 %)	5/34 (14.7 %)	7/34 (20.59 %)	16/34 (47.06 %)	20/34 (58.82 %)

Most systems supported Android (33/34; 97.06 %) and iOS (26/34; 76.47 %) platforms, with Android slightly more frequently specified. Twenty (58.82 %) systems additionally supported browser-based web interfaces or web applications. In 3/41 (7.32 %) publications platform support was not clearly, while only for one (2.94 %) of the systems this could not be resolved via other included publications. Note that while some publications reported support for wearables, none reported explicitly that the system is running on the wearable device itself. Several publications provided only partial information and lacked clear communication regarding which systems or platforms are supported for users, administrators, and other stakeholders.

External data sources were variably integrated. Non mutually exclusive 23/34 (67.65 %) systems supported access to external data from smartwatch or wearable sensors, 18/34 (52.94 %) locational information, 11/34 (32.35 %) microphone and 6/34 (17.65 %) camera data, 13/34 (38.23 %) external software systems, and 6/34 (17.65 %) stationary sensors, whereas the remaining 11/34 (32.35 %) systems relied primarily on self-reported input without external sensor integration.

Technological characteristics. Reporting of technical implementation details varied substantially across publications.

Regarding frontend development, 18/41 (43.9 %) publications reported the use of native software development kits (SDKs), and 8/41 (19.51 %) described cross-platform native frameworks. One (2.44 %) publication employed a hybrid or WebView-based framework. Web user interface frameworks and web technologies were mentioned in 8/41 (19.51 %) publications, sometimes in combination with native or cross-platform approaches and no clear separation. In 16/41 (39.02 %) publications, no information about frontend technologies was provided.

Backend implementation details were sparsely reported. Twenty-seven (65.85 %) publications did not specify backend frameworks or technologies at all. The remaining publications referred to technologies such as Laravel, Node.js, Parse, or general-purpose programming languages including Java, although reporting was often limited to naming a language rather than describing architectural structure.

Database technologies were similarly underreported. Twenty-six (65.85 %) publications did not provide any information regarding database solutions at all. Fourteen (34.15 %) vaguely mentioned database-related concepts such as NoSQL systems, cloud-based storage, or database-as-a-service (DBaaS), typically without detailed architectural description.

Cloud infrastructure was rarely specified. Thirty-one (75.61 %) publications did not report any cloud service usage. Among those that did, six (14.63 %) referred to Google Cloud Platform and three (7.32 %) to Amazon Web Services, with one (2.44 %) publication indicating the use of both providers. Three times (7.32 %) a publication mentioned other unspecified or proprietary cloud services.

Thirteen (31.71 %) publications referenced Internet of Things (IoT)-related concepts, typically in connection with wearable devices, sensor networks, or distributed sensing components.

The use of data analytics and machine learning tools was reported in a subset of publications. Thirteen (31.71 %) publications mentioned machine learning without specifying tools or frameworks. Others referred to custom in-house machine learning toolsets (6/41; 14.63 %), TensorFlow (1/41; 2.44 %), SciPy (1/41; 2.44 %), Apache Spark (1/41; 2.44 %), or advanced statistical modeling approaches such as multilevel vector autoregression (1/41; 2.44 %) and network modeling (1/41; 2.44 %). In multiple cases, no dedicated machine learning tools were reported.

Data security measures were inconsistently described. Eleven (26.83 %) publications provided some details, like the use of encryption libraries in general (7/41; 17.07 %), OAuth 2.0 (2/41; 4.88 %), AES-256e (1/41; 2.44 %). Penetration tests, JWT, Multi-factor authentication, and password-based access control each were mentioned once. However, in most cases (26/41; 63.41 %) security measures were mentioned only in general terms without further technical detail, and 5/41 (12.19 %) publications did not report security aspects at all.

Across the reviewed publications, authentication and authorization (13/41, 31.71 %), data encryption (11/41, 26.83 %), secure data storage (15/41, 36.58 %), and compliance with privacy regulations (15/41, 36.58 %) are the most frequently reported privacy measures. A substantial number of studies also highlight permissions management (12/41, 29.27 %), secure communication (10/41, 24.39 %), and privacy by design (6/41, 14.63 %), indicating a strong emphasis on access control and system-level safeguards. Furthermore, anonymization and pseudonymization (5/41, 12.19 %) and data minimization and purpose limitation (6/41, 14.63 %) are reported in a smaller subset of publications. More specialized measures, such as logging and monitoring (2/41, 4.88 %), secure coding practices (3/41, 7.32 %), third-party/SDK management (1/41, 2.44 %), and regular updates and patch management (1/41, 2.44 %), are only rarely mentioned. Similarly, user-focused approaches—including user control and transparency (9/41, 21.95 %) and user education and awareness (8/41, 19.51 %)—appear inconsistently across the literature. Compliance with privacy regulations like GDPR, CCPA, or HIPAA was mentioned in 15/41 (36.58 %) publications. Notably, a considerable number of publications provide only unspecified references to privacy measures (27/41, 65.85 %), while no explicit measures are reported in 3/41 (7.32 %) of the studies. This indicates that, although core technical and regulatory safeguards are commonly acknowledged, the overall reporting of data privacy practices remains inconsistent and often lacks sufficient detail.

Software development process characteristics. Aspects of the software development process were variably reported. Thirty (73.17 %) publications referred to a defined software engineering process, whereas 11/41 (26.83 %) did not describe any structured development approach. Literature or market analyses were mentioned in 29/41 (70.73 %), and quality assurance measures in 18/41 (43.9 %) publications. Cost considerations were addressed in 9/41 (21.95 %) publications and time or project duration aspects in 5/41 (12.19 %). In several cases, development process elements were only partially described.

The potential influence of user interface (UI) or user experience (UX) design on intervention effectiveness was mentioned in 32/41 (78.05 %) publications, discussed in detail in 1/41 (2.44 %), and not addressed in 8/41 (19.51 %).

Testing practices were rarely specified. Only 1/41 (2.44 %) publication explicitly reported automated testing procedures, 15/41 (36.58 %) mentioned testing in general or manual testing approaches without technical detail, and 24/41 (58.54 %) did not report any testing or CI/CD considerations. One (2.44 %) publication mentioned security tests.

Regulatory and policy considerations. Regulatory or policy aspects were mentioned in 28/41 (68.29 %), discussed in detail in 2/41 (4.88 %), and not addressed in 11/41 (26.83 %) publications. Explicit references to regulatory compliance included the General Data Protection Regulation (GDPR) in 13/41 (31.71 %) and the Health Insurance Portability and Accountability Act (HIPAA) in 8/41 (19.51 %) publications. Certification as a medical device was reported in 2/41 (4.88 %) publications, and 3/41 (7.32 %) referred to compliance with medical device regulations. One (2.44 %) publication mentioned other data protection frameworks or certifications.

Visualization in reporting. Out of 41 publications, 13 (31.71 %) did not include any screenshots, resulting in 9 out of 34 systems (26.47 %) lacking visual representation in the papers. Two of these publications (4.88 %) provided screenshots exclusively in supplementary materials. Overall, 27 out of 41 publications (65.85 %), covering 25 out of 34 systems (73.53 %), included at least one user interface screenshot. Screenshots of dashboards or overview interfaces were presented in 10 publications (24.39 %), corresponding to 10 systems (29.41 %). Additionally, 8 publications

(19.51 %) showcased backend interfaces, content management systems, or admin panels for 6 systems (17.65 %). Nine publications (21.95 %) further included additional visualizations in their supplementary materials.

Out of 41 publications, 16 (39.02 %) did not include any technical figures, resulting in 11 out of 34 systems (32.35 %) lacking visualization of technical aspects in the papers. Two of these publications (4.88 %) provided technical figures exclusively in supplementary materials. Overall, 21 out of 41 publications (51.22 %), covering 20 out of 34 systems (58.82 %), included figures of the system architecture. Figures of the application itself were presented in 17 publications (41.46 %), corresponding to 15 systems (44.12 %). Furthermore, 15 publications (36.58 %), covering 14 systems (41.18 %), included framework-related figures, while 10 publications (24.39 %), covering 9 systems (26.47 %), presented code architecture.

All publications included some form of data visualization. The most common types were app data tables (24/41; 58.54 %), non-app data tables (21/41; 51.22 %), non-app data charts (16/41; 39.02 %), and app data charts (15/41; 36.58 %). Additionally, 22 out of 41 publications (53.66 %) included a study flow chart or similar visualizations, while 25 out of 41 (60.97 %) presented other study-related visualizations.

3.3 Critical appraisal within sources of evidence

Given the nature of this review, no formal risk-of-bias assessment was conducted. However, substantial variability in methodological rigor and reporting transparency was observed across publications. Empirical evaluations were frequently limited to single-arm feasibility or pilot studies, with fewer randomized controlled trials. Technical reporting was heterogeneous, with many publications omitting detailed descriptions of backend architecture, database technologies, cloud infrastructure, security mechanisms, or interoperability standards. Similarly, regulatory compliance and certification aspects were inconsistently described. These differences may reflect varying levels of system maturity and documentation depth across the included sources. However, all included systems were required to demonstrate evidence of implementation and operational deployment.

3.4 Results of individual sources of evidence

Several publications described comprehensive, reusable platforms designed to support multiple EMA or EMI studies across domains. These systems typically provided configurable assessment modules, rule-based intervention engines, dashboard-based study management, and support for remote data synchronization. Examples included open-source platforms as well as commercially licensed systems, reflecting different sustainability and deployment models.

A subset of systems demonstrated large-scale deployment beyond pilot contexts. One platform reported more than 50,000 users across multiple applications[52], while others documented several thousand participants in observational or clinical contexts[20,30,49,53]. These large-scale implementations were typically associated with broader infrastructure components, including backend services and researcher dashboards.

Machine learning and adaptive logic were incorporated in several systems, although the degree of transparency regarding analytic pipelines varied. Some platforms explicitly described in-house machine learning toolsets or integration of established frameworks such as TensorFlow, whereas others referred to adaptive or AI-based components without technical specification.

Integration with external data sources such as wearable devices, operating system APIs, and sensor streams was a defining feature of multiple systems. Frameworks including mCerebrum[53], Sense2Stop[41], and AwarNS[21] implemented multimodal sensing combined with on-device Sense–Analyze–Act pipelines. These platforms supported real-time processing of physiological and contextual signals and enabled sensor-triggered just-in-time interventions. Adaptive logic was frequently linked to algorithmically derived states such as stress, activity, or geolocation, reducing reliance on continuous server-side processing.

A smaller subset of publications explicitly reported interoperability standards (e.g., HL7 FHIR) or integration with healthcare information systems.

Finally, one system – eSano[48,49] – covered the topic of certification as a medical device and several described compliance with data protection regulations, indicating progression toward regulated deployment contexts.

Illustrative system exemplars. To illustrate the diversity of architectural approaches, technological emphases, and methodological strategies identified in this review, selected systems are highlighted below. These exemplars represent distinct focal dimensions, including platform architecture, sensor-based infrastructures, artificial intelligence and digital phenotyping, healthcare integration and interoperability, and experimental optimization of adaptive decision rules.

m-Path[52] is presented as an easy-to-use and highly tailorable online platform that provides a configurable framework for implementing smartphone-based ecological momentary assessment (EMA) and ecological momentary intervention (EMI) in both research and clinical settings. The publication primarily focuses on system architecture and platform

design, emphasizing modular configurability and the reduction of development effort for researchers seeking to deploy customized EMA/EMI protocols. The framework is explicitly positioned for blended care scenarios, aiming to integrate digital assessment and intervention components into routine clinical practice. In addition to the architectural description, the paper reports empirical results from a single-arm feasibility study involving dashboard users (researchers and clinicians), including structured usability assessments and qualitative feedback procedures. User-reported outcomes indicated high acceptance and willingness to recommend the platform within the evaluated sample.

mCerebrum[53] is an open-source mobile sensing platform designed for developing multisensor biomarkers and sensor-triggered just-in-time adaptive interventions in real-world settings. The system emphasizes high-frequency raw data acquisition from smartphones and external wearable sensors, including physiological and inertial streams. Its modular architecture implements an on-device Sense–Analyze–Act pipeline with real-time biomarker computation and optimized resource management. A context-aware scheduler enables sensor-triggered prompts while managing participant burden. The platform is publicly available under the BSD 2-Clause license and has been deployed in large-scale field studies.

mindLAMP[56,57] is an open-source digital phenotyping platform designed to integrate patient-generated health data from smartphones and wearables into research and clinical care. Its architecture combines a configurable mobile and web frontend, a HIPAA-compliant backend, and the Cortex analytics toolkit supporting feature extraction, machine learning, and just-in-time adaptive interventions. The system supports interoperability with FHIR-based health systems and enables continuous push-based data synchronization for real-time monitoring. In a prospective validation study (n=67), machine learning models predicted changes in depressive symptoms with moderate accuracy (AUC=0.71), demonstrating the feasibility of algorithm-driven personalization based on passive and active smartphone data.

ADLIFE[20] presents a healthcare-integrated architecture for delivering personalized just-in-time adaptive interventions to patients with chronic conditions such as heart failure and COPD. The system is built around an HL7 FHIR-based interoperability framework, integrating patient and clinician platforms through standardized care plan resources and real-time event subscriptions. A rule-based JTAI engine continuously monitors patient adherence and performance metrics to trigger tailored behavioral interventions grounded in established behavior change techniques. The architecture is designed for large-scale pilot validation across multiple clinical sites, positioning the platform within formal healthcare infrastructures and regulatory contexts.

Sense2Stop[41] describes a micro-randomized trial designed to optimize decision rules for a sensor-triggered just-in-time adaptive intervention aimed at preventing smoking relapse. The system uses wearable chest and wrist sensors to collect physiological and movement data, which are processed in real time on the smartphone using machine-learned algorithms to detect stress and smoking episodes. At each eligible decision point, participants are randomly assigned to receive or not receive a stress-management prompt, enabling the estimation of context-specific intervention effects. Randomization is constrained by burden- and safety-related availability rules, positioning the study as a methodologically rigorous approach to optimizing adaptive intervention logic.

3.5 Synthesis of results

Across the included publications, substantial heterogeneity was observed in architectural design, technological sophistication, evaluation maturity, and reporting depth. While the systems aimed to provide reusable infrastructures for ecological momentary assessment and intervention, they differed markedly in their emphasis on configurability, sensor integration, adaptive logic, interoperability, and clinical integration. A spectrum of adaptivity emerged, ranging from rule-based reminder systems to sensor-driven just-in-time adaptive interventions and machine-learning–supported digital phenotyping platforms. At the same time, reporting practices varied considerably, particularly regarding backend technologies, database architectures, security mechanisms, interoperability standards, and regulatory positioning. Overall, the evidence base reflects a rapidly evolving research field characterized by architectural innovation and methodological experimentation alongside uneven empirical validation and technical disclosure.

Architectural design patterns and implementation strategies

Across the included systems, reusable EMA/EMI platforms predominantly followed client–server smartphone architectures, often complemented by web-based dashboards for study configuration, monitoring, and data export. Where frontend technologies were mentioned, Native mobile development was most frequently reported, while cross-platform or hybrid solutions were less common and often insufficiently specified. Backend technologies, database

infrastructures, and cloud deployments were inconsistently described, with the majority of publications omitting detailed architectural information. Open-source solutions were represented alongside in-house and proprietary systems; however, licensing transparency and technical disclosure varied substantially across publications. Interoperability was frequently mentioned, yet only a small subset of systems specified concrete standards such as `HL7 FHIR`. Overall, architectural patterns ranged from relatively lightweight survey delivery systems to modular, extensible frameworks designed for large-scale or clinically integrated deployments.

Spectrum of configurability and adaptivity

A continuum of adaptivity emerged across systems. At the basic level, platforms implemented configurable scheduling and rule-based reminders. More advanced systems incorporated contextual triggers derived from self-reports or passive sensing. Sensor-driven just-in-time adaptive interventions (JITAIs) were implemented in several architectures, enabling interventions to be triggered by physiological, behavioral, or environmental signals. High-frequency multimodal sensing infrastructures were limited to a small subset of systems, reflecting substantial variation in sensing depth and computational complexity. A smaller subset employed micro-randomized trial designs to experimentally optimize decision rules for intervention delivery. Machine learning and digital phenotyping approaches were increasingly reported, supporting symptom prediction, behavioral feature extraction, and adaptive intervention logic. Engagement support was predominantly implemented through push notifications and reminder systems, whereas gamification elements were comparatively rare. These architectural differences directly influence how interventions are personalized, timed, and integrated into research and clinical workflows.

Evaluation maturity and empirical alignment

The empirical landscape was characterized by a predominance of feasibility and pilot studies, with fewer randomized controlled trials and controlled comparative evaluations. A substantial proportion of publications focused on framework presentation without reporting empirical outcomes, although some described planned, ongoing or related trials. Sample sizes varied considerably, ranging from small exploratory deployments to large-scale implementations. Architectural development often co-occurred with early-stage evaluation rather than large-scale controlled trials.

Technical transparency and engineering rigor

Reporting depth varied markedly across publications. Frontend technologies were more frequently described than backend frameworks, database technologies, or cloud infrastructures. Machine learning extensions of the system were more often discussed in general rather than actually implemented. Yet, when integrated, specific model architectures, training procedures, or deployment pipelines were mostly described without detail. Security and privacy measures were commonly referenced, but frequently lacked technical specificity. Continuous integration, automated testing, and formal software engineering processes were rarely reported. These patterns indicate heterogeneous engineering maturity and limited reproducibility based solely on published descriptions.

Interoperability, regulation, and healthcare integration

Healthcare integration emerged as an important, though unevenly realized, dimension. While many systems referenced regulatory frameworks such as GDPR or HIPAA, detailed descriptions of compliance mechanisms were uncommon. Explicit medical device certification was rare. Interoperability standards were infrequently implemented in a technically specified manner, although some architectures demonstrated integration with established healthcare information systems. Overall, the field shows increasing awareness of regulatory and clinical integration requirements, yet implementation depth varies substantially.

Visualization in reporting

The results show variability in how visual information is reported across publications. While a majority of publications include user interface screenshots and system architecture figures, a substantial proportion does not provide such materials, and some information is only available in supplementary files. Reporting of technical aspects is less consistent than user-facing elements, with fewer publications including framework or code architecture figures. All publications include some form of data visualization, most commonly tables, while charts and study-related visualizations such as flow diagrams are present in a smaller subset. Differences between publication-level and system-level counts further indicate that not all systems are equally represented in the visual reporting.

Temporal evolution of architectural and design patterns

Within the temporal scope of the included publications (2014–2024), both the number of identified systems and the publication volume increased, with a marked rise after 2018 and a concentration between 2020 and 2024, indicating growing research activity in configurable EMA/EMI systems.

Across this period, both framework-oriented publications and empirically evaluated systems were represented. However, no consistent temporal progression was observed in the reporting depth of backend technologies, database infrastructures, cloud services, or security mechanisms, as incomplete technical disclosure occurred across the full time range, and architectural approaches, levels of adaptivity, and evaluation designs remained heterogeneous.

Artificial intelligence or machine learning components were reported in several systems—primarily from 2020 onward—including AI-generated content and machine-learning–supported decision logic; however, these features remained limited to a subset of systems, and their technical detail and operational definitions varied across publications.

Explicit references to JITAI terminology, micro-randomized trial designs, and formalized adaptive decision rules appeared more frequently after 2019, suggesting increasing methodological formalization rather than a structural shift in system architectures. Interoperability standards and healthcare-integrated architectures, including explicit references to HL7 FHIR and regulatory frameworks, were predominantly reported in publications from 2022–2023.

Overall, the evidence reflects continued diversification and experimentation rather than a uniform progression in architectural sophistication or evaluation rigor.

4 Discussion

4.1 Summary of evidence

The findings of this review map 41 publications describing 34 configurable native mHealth systems for EMA, EMI, and JITAI delivery, addressing the primary objective of identifying and characterizing reusable platforms in this domain. Across these systems, substantial heterogeneity was observed in architectural design, implementation strategies, configurability mechanisms, and evaluation approaches, reflecting the scope of SRQ1 and SRQ2. While publication activity increased over time, particularly after 2018, no consistent temporal progression in architectural sophistication, adaptivity, or reporting depth was evident (SRQ3).

The dominant pattern was a smartphone-centered platform with configurable study logic and frequent use of web-based dashboards for configuration, monitoring, and data export. Where implementation details were reported, native SDK development was most common (n=14), followed by cross-platform native frameworks (n=8), with hybrid/WebView approaches rarely reported (n=1). However, core infrastructure details were often missing, including backend frameworks (no information in n=28), database technologies (no information in n=26), and cloud services (no information in n=32).

Regarding how these systems supported interventions through adaptability and deployment features, most platforms combined multiple EMA sampling strategies rather than relying on a single approach. Signal-based prompting was nearly universal (39/41), with frequent support for event-contingent sampling (n=33) and interval-based collection (n=25). Passive sensing without user input was present in over half of the systems (n=23), while context-aware triggering was reported less often (n=17), indicating that true context-dependent assessment logic was implemented in a minority.

Intervention support was characterized by broad availability of both static and dynamic components. Static content most commonly included advice (n=34), reminders (n=32), and educational material (n=24). Dynamic intervention delivery was predominantly automated (n=38), while provider-triggered or provider-reviewed mechanisms were less common (n=11). Personalization based on predefined user data was widely supported (n=34), and context-triggered dynamic content was reported in about half of systems (n=20).

JITAI-related capabilities were frequently claimed but operationalized at varying levels of specificity. Defined JITAI triggers were reported in 29 systems, adaptable intervention content in 27 systems, context-aware delivery in 21 systems, and explicit just-in-time delivery in 19 systems. AI-related components were comparatively uncommon and concentrated in a small subset. AI-generated static content was reported in 10 systems, dynamic AI-generated content in 7, and AI-supported decision logic in 4, suggesting that most "adaptivity" was implemented via configuration, rules, or context triggers rather than AI-driven decision-making.

Engagement support focused on operational delivery mechanisms rather than behavioral design features. Push notifications were the dominant approach (local scheduling: n=31; remotely triggered: n=24), with many publications mentioning notifications without sufficient detail to classify the delivery mechanism (n=12; categories not mutually exclusive). In-app reminders (n=21) and to-do/task views (n=8) were common, whereas external notifications (n=6) and alerts to third parties (n=1) were rare. Gamification was uncommon (partial game elements: n=3; competitive mechanics: n=1), while visualization was more prevalent for both participants (n=25) and researchers/administrators (n=26).

The evaluation base was mixed and often early-stage. Just over half of publications reported original empirical results (23/41), while 17 focused on framework or platform description without outcomes, sometimes including

planned trials. Among empirically evaluated systems (n=23), feasibility/pilot studies were most common (n=9), followed by non-randomized controlled studies (n=6) and RCTs (n=5), with sample sizes ranging from very small pilots to deployments exceeding several thousand users and one report exceeding 50,000 participants. Overall, the evidence indicates a broad and active design space with substantial heterogeneity in architectural patterns, adaptivity mechanisms, and evaluation maturity, alongside persistently uneven technical transparency.

For temporal evolution (SRQ3), the dataset supports a clear increase in publication volume from 2018 onward and concentration between 2020 and 2024, with a peak in 2022. Beyond publication frequency, the available data do not justify strong claims of a linear shift in sophistication, adaptivity, or evaluation rigor across years, because architectural reporting, adaptivity claims, and empirical designs remain heterogeneous throughout the period observed.

Taken together, these findings suggest that the field has progressed from developing isolated intervention applications toward configurable infrastructures, but that this transition remains incomplete. While reusable platforms are increasingly available, they are often described at the level of features rather than as fully specified software systems, limiting reproducibility and comparability. In practice, this means that many platforms are reusable in principle, but difficult to evaluate, extend, or integrate due to missing technical detail. As a result, the current evidence base reflects a design space that is functionally rich but structurally fragmented, with limited convergence toward shared architectural standards or reporting practices.

Relation to prior reviews. The present findings are broadly consistent with earlier reviews showing that EMA- and EMI-based mobile interventions are heterogeneous, often evaluated in small or early-stage studies, and described with inconsistent terminology and reporting detail. Earlier EMI reviews emphasized the diversity of intervention designs and the predominance of emerging smartphone-based approaches, especially in mental health contexts[61,62]. Likewise, a systematic review of smartphone-delivered EMIs based on EMA found that most studies were small, quasi-experimental, relied mainly on self-reported EMA, and rarely used sensors, while also highlighting poor standardization in the reporting of EMA and EMI components[63]. Our review aligns with these observations but extends them by shifting the unit of analysis from interventions to reusable technical systems. From that perspective, heterogeneity is not limited to intervention content or study design; it also applies to platform architecture, configurability mechanisms, sensing pipelines, and deployment infrastructures.

Our findings are also consistent with prior JITAI-focused reviews that reported incomplete reporting of core adaptive-intervention components and limited automation in practice[64]. In that sense, the current review reinforces the broader conclusion that the field often uses adaptation-oriented language more readily than it documents the underlying decision logic in a reproducible manner. At the same time, our dataset suggests that the technical design space may be somewhat more advanced than intervention-centered reviews alone imply: several systems incorporated passive sensing, multimodal data streams, dashboard-based configuration, and, in some cases, interoperability with clinical systems. However, these more advanced capabilities remained concentrated in a subset of platforms and were still reported unevenly. Taken together, prior reviews and the present findings point to the same core tension: rapid conceptual growth in EMA/EMI/JITAI research has not yet been matched by equally mature standards for technical reporting, architectural transparency, and reusable system documentation.

4.2 Strengths and contribution

A key strength of this review is that it focused specifically on implemented, configurable, and reusable native mHealth systems rather than on isolated intervention studies alone. This made it possible to distinguish between publication units and underlying platforms and to characterize the field not only in terms of clinical application areas, but also in terms of architectural patterns, configurability mechanisms, runtime adaptation, interoperability, and software engineering characteristics. By combining a broad mapping perspective with system-level technical extraction, the review extends prior EMA/EMI/JITAI reviews that primarily emphasized intervention content, outcomes, or user engagement, and provides a structured overview of reusable infrastructures as a distinct object of study.

4.3 Limitations

This review was limited to peer-reviewed publications and did not include gray literature (e.g., preprints, technical reports, or repository-first systems), which may contain relevant reusable EMA/EMI infrastructures. Despite searching five databases, commercially deployed or proprietary platforms available via the web or in app stores and not described in academic publications may have been missed.

Eligibility criteria required evidence of implemented software, strengthening practical relevance but potentially underrepresenting emerging architectures and conceptual systems. The requirement for both EMA and EMI features may have excluded EMA-only platforms, and the focus on native mobile app frameworks may have omitted

web-first systems or hybrid service compositions based on multiple third-party components. Additionally, some intervention-focused clinical studies may have used configurable frameworks without explicitly describing them, leading to underrepresentation due to reporting practices. Furthermore, some systems may be described across multiple publications, each of which individually did not meet the inclusion criteria, but collectively would have fulfilled them. These criteria likely biased the sample toward academically published research prototypes and intervention-focused systems with formal architectural descriptions, while underrepresenting commercial platforms, public health deployments (e.g., SMS-based systems), and industry-developed frameworks not described in peer-reviewed literature. If the aim is to map the broader ecosystem of configurable EMA/EMI systems, this may result in a restricted perspective. Only publications in English were considered, which may have excluded relevant systems described in other languages.

Technical reporting was frequently incomplete, limiting reliable comparison of backend architectures, database technologies, cloud infrastructures, interoperability mechanisms, and security controls. In several cases, "not reported" should not be interpreted as "not implemented". Despite predefined inclusion criteria, several criteria—particularly configurability, framework characteristics, and EMA/EMI support—required interpretative judgment during screening and full-text assessment. Because these features were often incompletely or inconsistently reported, classification decisions depended on reviewer interpretation. This may have led to the exclusion of systems that partially met the criteria but lacked sufficiently explicit description.

Overall, the findings indicate that visual reporting is common but not consistently comprehensive. Although most publications provide at least some insight into system interfaces or architecture, gaps remain in both the availability and depth of visual documentation. Technical details, particularly at the level of code architecture, are less frequently illustrated, and some relevant materials are only accessible through supplementary content. At the same time, data visualizations are consistently included, primarily in the form of tables. These patterns suggest heterogeneity in reporting practices across publications.

Terminology related to adaptive system design was used inconsistently across publications and often lacked conceptual precision or detail. Concepts such as personalization, context-awareness, adaptivity, and artificial intelligence were defined and operationalized heterogeneously. In particular, "AI" ranged from rule-based or statistical approaches to supervised machine learning and digital phenotyping pipelines, often without sufficient detail to distinguish deterministic logic from adaptive data-driven models.

Finally, no formal risk-of-bias assessment was conducted, and empirical outcomes were not synthesized for effectiveness, as the review focused on characterizing technological systems rather than evaluating clinical or behavioral outcomes.

4.4 Conclusions

Reusable native mHealth systems for EMA, EMI, and JITAI delivery show wide variation in architecture, adaptivity mechanisms, sensing depth, and evaluation maturity. Most platforms support flexible EMA schedules and combine multiple sampling strategies, while intervention delivery commonly includes both static content and automated dynamic features, with personalization frequently implemented via predefined user data and contextual triggers. Explicit JITAI capabilities are often reported, but AI-driven decision logic remains rare, and many adaptivity claims are implemented through configuration rules rather than data-driven approaches.

A central cross-cutting limitation of the current evidence base is inconsistent technical transparency, particularly regarding backend architectures, database technologies, cloud infrastructures, interoperability mechanisms, and security implementations. This limits reproducibility, comparability, and the ability to select or reuse existing platforms in both research and applied settings.

These findings have several implications. For researchers and system developers, more standardized and complete reporting of technical architectures and implementation details is necessary to enable reproducibility and informed reuse of existing platforms. For the research community, clearer and more consistent operational definitions of adaptivity, personalization, and AI-related components are needed to improve comparability across studies. The development of domain-specific reporting guidelines for configurable EMA/EMI systems may help improve consistency, comparability, and transparency across publications. Clearer conceptual distinctions between rule-based, context-aware, and data-driven adaptive mechanisms may improve interpretability and reduce ambiguity in the characterization of system capabilities. In addition, increased availability of open-source implementations and publicly accessible codebases would substantially improve transparency, reproducibility, and reuse of existing systems, and support cumulative development across research groups.

For system design and deployment, greater emphasis on modular and interoperable architectures may facilitate reuse of components and reduce duplication of development effort. More consistent adoption and explicit implementation of interoperability standards may support integration with healthcare systems and improve scalability beyond

research settings. Stronger alignment between claimed adaptive capabilities and empirical validation, particularly through controlled or optimization-focused study designs, is needed to substantiate system effectiveness.

Future work should focus on improving technical reporting practices, strengthening interoperability and standardization efforts, and better integrating software engineering practices with rigorous empirical evaluation. Together, these steps may support the transition from heterogeneous research prototypes toward more standardized, interoperable, and reusable mHealth infrastructures.

Ethics approval and consent to participate

This review used only published data and did not involve human participants.

Funding

This work was partly supported by the ePro-Schools project (DE Funding Code `01EA2405`) under the ERA4Health partnership (`grant agreement No. 101095426`) of the EU Horizon Europe Research and Innovation Programme.

Acknowledgements

We thank an external contributor for assistance with title and abstract screening. We thank Lilly Sophia Brandstetter (Institute for medical Data Science, University Hospital Würzburg, Würzburg, Germany) for providing feedback on and proofreading the manuscript.

Author contributions

C.V. and R.P. developed the research questions and conceived the study. C.V. conducted the literature search, performed primary study selection, analyzed the data, synthesized the findings, developed analysis scripts, and drafted and revised the manuscript. Title and abstract screening was performed by C.V., M.S., and an additional external reviewer who preferred to remain anonymous. C.V. performed full-text review and data extraction for all included studies. M.S., A.R.I., and M.W. independently reviewed and extracted data from subsets of studies. M.S. and A.R.I. resolved disagreements in study selection and data extraction. R.P. supervised the project and critically revised the manuscript. All authors have read and approved the final manuscript.

Competing interests

The authors report no conflict of interest.

Additional information

Correspondence and requests for materials should be addressed to Carsten Vogel.

Reprints and permissions information is available at http://www.nature.com/reprints.

References

1. Shiffman, S., Stone, A. A. & Hufford, M. R. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* **4**, 1–32, DOI: 10.1146/annurev.clinpsy.3.022806.091415 (2008).

2. Heron, K. E. & Smyth, J. M. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br. J. Heal. Psychol.* **15**, 1–39, DOI: 10.1348/135910709X466063 (2010).

3. Nahum-Shani, I. *et al.* Just-in-time adaptive interventions (jitais) in mobile health: Key components and design principles for ongoing health behavior support. *Annals Behav. Medicine* **52**, 446–462, DOI: 10.1007/s12160-016-9830-8 (2018).

4. de Vries, L. P., Baselmans, B. M. L. & Bartels, M. Smartphone-based ecological momentary assessment of well-being: A systematic review and recommendations for future studies. *Appl. Res. Qual. Life* **16**, 2361–2408, DOI: 10.1007/s10902-020-00324-7 (2021).

5. Walz, L. C., Nauta, M. H. & Aan Het Rot, M. Experience sampling and ecological momentary assessment studies in anxiety disorders: A systematic review. *Clin. Psychol. Rev.* **34**, 594–606, DOI: 10.1016/j.cpr.2014.06.008 (2014).

6. Balaskas, A., Schueller, S. M., Cox, A. L. & Doherty, G. Ecological momentary interventions for mental health: A scoping review. *PLOS ONE* **16**, e0248152, DOI: 10.1371/journal.pone.0248152 (2021).

7. Dao, K. P., de Vries, H., Meijer, E. & Hoving, C. Smartphone-delivered ecological momentary interventions based on ecological momentary assessments to promote health behaviors: Systematic review and adapted checklist for reporting ema/emi studies. *JMIR mHealth uHealth* **9**, e22890, DOI: 10.2196/22890 (2021).

8. Colombo, D. *et al.* Current state and future directions of technology-based ecological momentary assessment and intervention for major depressive disorder: A systematic review. *J. Clin. Medicine* **8**, 465, DOI: 10.3390/jcm8040465 (2019).

9. Gross, T. & Malzhacker, T. The experience sampling method and its tools: A review for developers, study administrators, and participants. *Proc. ACM on Human-Computer Interact.* **7**, 1–29, DOI: 10.1145/3593234 (2023).

10. Henry, L. M., Scherer, E. A. & Ridenour, T. A. Selecting an ecological momentary assessment platform: Considerations and recommendations for researchers. *J. Med. Internet Res.* **26**, e51125, DOI: 10.2196/51125 (2024).

11. Thai, S. & Page-Gould, E. Experiencesampler: An open-source scaffold for building smartphone apps for experience sampling. *Psychol. Methods* **23**, 639–654, DOI: 10.1037/met0000151 (2018).

12. Van De Ven, P. *et al.* Ultemat: A mobile framework for smart ecological momentary assessments and interventions. *Internet Interv.* **9**, 74–81, DOI: 10.1016/j.invent.2017.07.001 (2017).

13. Mestdagh, M., Dejonckheere, E. & Kuppens, P. m-path: An easy-to-use and highly tailorable platform for ecological momentary assessment and intervention. *Front. Digit. Heal.* **5**, 1182175, DOI: 10.3389/fdgth.2023.1182175 (2023).

14. Page, M. J. *et al.* The prisma 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, DOI: 10.1136/bmj.n71 (2021). https://www.bmj.com/content/372/bmj.n71.full.pdf.

15. Tricco, A. C. *et al.* Prisma extension for scoping reviews (prisma-scr): Checklist and explanation. *Annals Intern. Medicine* **169**, 467–473, DOI: 10.7326/M18-0850 (2018). PMID: 30178033, https://doi.org/10.7326/M18-0850.

16. Vogel, C., Schlüter, M., Idrees, A. R., Winter, M. & Pryss, R. a. Configurable ema, emi, and jitai mhealth platforms: A systematic scoping review and technical implications for reusable no-code frameworks. OSF, DOI: 10.17605/OSF.IO/43256 (2025).

17. Foster, E. D. & Deardorff, A. Open science framework (osf). *J. Med. Libr. Assoc. JMLA* **105**, 203–206, DOI: 10.5195/jmla.2017.88 (2017).

18. McGowan, J. *et al.* Press peer review of electronic search strategies: 2015 guideline statement. *J. Clin. Epidemiol.* **75**, 40–46, DOI: https://doi.org/10.1016/j.jclinepi.2016.01.021 (2016).

19. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan—a web and mobile app for systematic reviews. *Syst. Rev.* **5**, 210, DOI: 10.1186/s13643-016-0384-4 (2016).

20. Laleci Erturkmen, G. B. *et al.* Enabling patient adherence via personalised, just-in time adaptive interventions in adlife architecture. In *Proceedings of the 10th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*, DSAI 2022, 44–49, DOI: 10.1145/3563137.3563151 (ACM, 2022).

21. González-Pérez, A. *et al.* Awarns: A framework for developing context-aware reactive mobile applications for health and mental health. *J. Biomed. Informatics* **141**, 104359, DOI: 10.1016/j.jbi.2023.104359 (2023).

22. Nuijten, R. *et al.* Evaluating the impact of adaptive personalized goal setting on engagement levels of government staff with a gamified mhealth tool: Results from a 2-month randomized controlled trial. *JMIR mHealth uHealth* **10**, e28801, DOI: 10.2196/28801 (2022).

23. Curtis, A., Pai, A., Cao, J., Moukaddam, N. & Sabharwal, A. Healthsense: Software-defined mobile-based clinical trials. In *The 25th Annual International Conference on Mobile Computing and Networking*, MobiCom '19, 1–15, DOI: 10.1145/3300061.3345433 (ACM, 2019).

24. Azizoddin, D. R. *et al.* Development and pre-pilot testing of stamp + cbt: an mhealth app combining pain cognitive behavioral therapy and opioid support for patients with advanced cancer and pain. *Support. Care Cancer* **32**, DOI: 10.1007/s00520-024-08307-7 (2024).

25. Walters, S. T. *et al.* Development and preliminary effectiveness of a smartphone-based, just-in-time adaptive intervention for adults with alcohol misuse who are experiencing homelessness. *Alcohol. Clin. Exp. Res.* **46**, 1732–1741, DOI: 10.1111/acer.14908 (2022).

26. Sizemore, K. M. *et al.* A proof of concept pilot examining feasibility and acceptability of the positively healthy just-in-time adaptive, ecological momentary, intervention among a sample of sexual minority men living with hiv. *J. Happiness Stud.* **23**, 4091–4118, DOI: 10.1007/s10902-022-00587-2 (2022).

27. Bell, I. H. *et al.* Smartphone-based ecological momentary assessment and intervention in a blended coping-focused therapy for distressing voices: Development and case illustration. *Internet Interv.* **14**, 18–25, DOI: 10.1016/j.invent.2018.11.001 (2018).

28. Asiri, Y. A., Millard, D. E. & Weal, M. J. Assessing the impact of engagement and real-time feedback in a mobile behavior change intervention for supporting critical thinking in engineering research projects. *IEEE Transactions on Learn. Technol.* **14**, 445–459, DOI: 10.1109/tlt.2021.3104817 (2021).

29. Ponder, M. *et al.* A smartphone app with a digital care pathway for patients undergoing spine surgery: Development and feasibility study. *JMIR Perioper. Medicine* **3**, e21138, DOI: 10.2196/21138 (2020).

30. Klein, A. *et al.* Remote digital psychiatry for mobile mental health assessment and therapy: Mindlogger platform development study. *J. Med. Internet Res.* **23**, e22369, DOI: 10.2196/22369 (2021).

31. Künzler, F. *et al.* Exploring the state-of-receptivity for mhealth interventions. *Proc. ACM on Interactive, Mobile, Wearable Ubiquitous Technol.* **3**, 1–27, DOI: 10.1145/3369805 (2019).

32. Soares Teles, A. *et al.* Enriching mental health mobile assessment and intervention with situation awareness. *Sensors* **17**, 127, DOI: 10.3390/s17010127 (2017).

33. Golbus, J. R. *et al.* A physical activity and diet just-in-time adaptive intervention to reduce blood pressure: The mybpmylife study rationale and design. *J. Am. Hear. Assoc.* **13**, DOI: 10.1161/jaha.123.031234 (2024).

34. Zainal, N. H., Tan, H. H., Hong, R. Y. & Newman, M. G. Is a brief mindfulness ecological momentary intervention more efficacious than a self-monitoring app for social anxiety disorder? a randomized controlled trial. *J. Anxiety Disord.* **104**, 102858, DOI: 10.1016/j.janxdis.2024.102858 (2024).

35. Liu, S., La, H., Willms, A. & Rhodes, R. E. A "no-code" app design platform for mobile health research: Development and usability study. *JMIR Form. Res.* **6**, e38737, DOI: 10.2196/38737 (2022).

36. Hanssen, E. *et al.* An ecological momentary intervention incorporating personalised feedback to improve symptoms and social functioning in schizophrenia spectrum disorders. *Psychiatry Res.* **284**, 112695, DOI: 10.1016/j.psychres.2019.112695 (2020).

37. Schueller, S. M., Begale, M., Penedo, F. J. & Mohr, D. C. Purple: A modular system for developing and deploying behavioral intervention technologies. *J. Med. Internet Res.* **16**, e181, DOI: 10.2196/jmir.3376 (2014).

38. Marciniak, M. A. *et al.* Imager—a mobile health mental imagery-based ecological momentary intervention targeting reward sensitivity: A randomized controlled trial. *Appl. Psychol. Heal. Well-Being* **16**, 576–596, DOI: 10.1111/aphw.12505 (2023).

39. O'Brien, S. T. *et al.* Sema3: A free smartphone platform for daily life surveys. *Behav. Res. Methods* **56**, 7691–7706, DOI: 10.3758/s13428-024-02445-w (2024).

40. Kraaij, W. *et al.* Personalized support for well-being at work: an overview of the swell project. *User Model. User-Adapted Interact.* DOI: 10.1007/s11257-019-09238-3 (2019).

41. Battalio, S. L. *et al.* Sense2stop: A micro-randomized trial using wearable sensors to optimize a just-in-time-adaptive stress management intervention for smoking relapse prevention. *Contemp. Clin. Trials* **109**, 106534, DOI: 10.1016/j.cct.2021.106534 (2021).

42. Konigorski, S. *et al.* Studyu: A platform for designing and conducting innovative digital n-of-1 trials. *J. Med. Internet Res.* **24**, e35884, DOI: 10.2196/35884 (2022).

43. Everett, E., Kane, B., Yoo, A., Dobs, A. & Mathioudakis, N. A novel approach for fully automated, personalized health coaching for adults with prediabetes: Pilot clinical trial. *J. Med. Internet Res.* **20**, e72, DOI: 10.2196/jmir.9723 (2018).

44. González-Pérez, A., Miralles, I., Granell, C. & Casteleyn, S. Technical challenges to deliver sensor-based psychological interventions using smartphones. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, UbiComp '19, 915–920, DOI: 10.1145/3341162.3346271 (ACM, 2019).

45. van 't Klooster, J.-W. J. R., Rabago Mayer, L. M., Klaassen, B. & Kelders, S. M. Challenges and opportunities in mobile e-coaching. *Front. Digit. Heal.* **5**, DOI: 10.3389/fdgth.2023.1304089 (2024).

46. van de Ven, P. *et al.* Ultemat: A mobile framework for smart ecological momentary assessments and interventions. *Internet Interv.* **9**, 74–81, DOI: 10.1016/j.invent.2017.07.001 (2017).

47. Brick, T. R., Mundie, J., Weaver, J., Fraleigh, R. & Oravecz, Z. Low-burden mobile monitoring, intervention, and real-time analysis using the wear-it framework: Example and usability study. *JMIR Form. Res.* **4**, e16072, DOI: 10.2196/16072 (2020).

48. Vogel, C. *et al.* A highly configurable ema and jitai mobile app framework utilized in a large-scale german study on breast cancer aftercare. In *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, 103–110, DOI: 10.1109/cbms58004.2023.00200 (IEEE, 2023).

49. Kraft, R. *et al.* esano – an ehealth platform for internet- and mobile-based interventions. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1997–2002, DOI: 10.1109/embc46164.2021.9629534 (IEEE, 2021).

50. Pulantara, I. W., Parmanto, B. & Germain, A. Development of a just-in-time adaptive mhealth intervention for insomnia: Usability study. *JMIR Hum. Factors* **5**, e21, DOI: 10.2196/humanfactors.8905 (2018).

51. O'Driscoll, C. *et al.* An ecological mobile momentary intervention to support dynamic goal pursuit: Feasibility and acceptability study. *JMIR Form. Res.* **8**, e49857, DOI: 10.2196/49857 (2024).

52. Mestdagh, M. *et al.* m-path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. *Front. Digit. Heal.* **5**, DOI: 10.3389/fdgth.2023.1182175 (2023).

53. Hossain, S. M. *et al.* mcerebrum: A mobile sensing software platform for development and validation of digital biomarkers and interventions. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, SenSys '17, 1–14, DOI: 10.1145/3131672.3131694 (ACM, 2017).

54. Blevins, C. E., Marsh, E. L., Stein, M. D., Schatten, H. T. & Abrantes, A. M. Project choice: Choosing healthy options in coping with emotions, an ema/emi plus in-person intervention for alcohol use. *Subst. Abus.* **42**, 569–576, DOI: 10.1080/08897077.2020.1806182 (2021).

55. Shrier, L. A. *et al.* Pilot randomized trial of moment, a motivational counseling-plus-ecological momentary intervention to reduce marijuana use in youth. *mHealth* **4**, 29–29, DOI: 10.21037/mhealth.2018.07.04 (2018).

56. Vaidyam, A., Halamka, J. & Torous, J. Enabling research and clinical use of patient-generated health data (the mindlamp platform): Digital phenotyping study. *JMIR mHealth uHealth* **10**, e30557, DOI: 10.2196/30557 (2022).

57. Currey, D. & Torous, J. Digital phenotyping data to predict symptom improvement and mental health app personalization in college students: Prospective validation of a predictive model. *J. Med. Internet Res.* **25**, e39258, DOI: 10.2196/39258 (2023).

58. Moore, E., Williams, A., Bell, I. & Thomas, N. Client experiences of blending a coping-focused therapy for auditory verbal hallucinations with smartphone-based ecological momentary assessment and intervention. *Internet Interv.* **19**, 100299, DOI: 10.1016/j.invent.2019.100299 (2020).

59. Shatte, A. B. R. & Teague, S. J. schema: an open-source, distributed mobile platform for deploying mhealth research tools and interventions. *BMC Med. Res. Methodol.* **20**, DOI: 10.1186/s12874-020-00973-5 (2020).

60. Álvarez, R., Murua, A., Artetxe, A., Epelde, G. & Beristain, A. A platform for user empowerment through self ecological momentary assessment / intervention. In *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare - "Transforming healthcare through innovations in mobile and wireless technologies"*, MOBIHEALTH, DOI: 10.4108/eai.14-10-2015.2261615 (ICST, 2015).

61. Versluis, A., Verkuil, B., Spinhoven, P., van der Ploeg, M. M. & Brosschot, J. F. Changing mental health and positive psychological well-being using ecological momentary interventions: A systematic review and meta-analysis. *J. Med. Internet Res.* **18**, e152, DOI: 10.2196/jmir.5642 (2016).

62. Balaskas, A., Schueller, S. M., Cox, A. L. & Doherty, G. Ecological momentary interventions for mental health: A scoping review. *PLOS ONE* **16**, e0248152, DOI: 10.1371/journal.pone.0248152 (2021).

63. Dao, K. P. *et al.* Smartphone-delivered ecological momentary interventions based on ecological momentary assessments to promote health behaviors: Systematic review and adapted checklist for reporting ecological momentary assessment and intervention studies. *JMIR Mhealth Uhealth* **9**, e22890, DOI: 10.2196/22890 (2021).

64. Oikonomidi, T., Ravaud, P., LeBeau, J. & Tran, V.-T. A systematic scoping review of just-in-time, adaptive interventions finds limited automation and incomplete reporting. *J. Clin. Epidemiol.* **154**, 108–116, DOI: 10.1016/j.jclinepi.2022.12.006 (2023).