International Journal of Medical Informatics 174 (2023) 105061

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Review article

Using digital phenotyping to understand health-related outcomes: A scoping review

Kyungmi Lee[a,*], Tim Cheongho Lee[b,*], Maria Yefimova[c], Sidharth Kumar[d], Frank Puga[e], Andres Azuero[e], Arif Kamal[f], Marie A. Bakitas[e,g,h], Alexi A. Wright[i], George Demiris[j], Christine S. Ritchie[k], Carolyn E.Z. Pickering[e], J. Nicholas Dionne-Odom[e,g,h]

[a] Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, United States
[b] College of Gyedang General Education, Sangmyung University, Seoul, Republic of Korea
[c] Health Department of Nursing, University of California San Francisco, San Francisco, CA, United States
[d] Department of Computer Science, University of Alabama at Birmingham, Birmingham, AL, United States
[e] School of Nursing, University of Alabama at Birmingham, Birmingham, AL, United States
[f] Department of Medicine, Duke University School of Medicine, Durham, NC, United States
[g] Division of Geriatrics, Gerontology, and Palliative Care, University of Alabama at Birmingham, Department of Medicine, Birmingham, AL, United States
[h] University of Alabama at Birmingham, Center for Palliative and Supportive Care, Birmingham, AL, United States
[i] Harvard Medical School, Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, United States
[j] Department of Biobehavioral and Health Sciences, School of Nursing & Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
[k] Division of Palliative Care and Geriatric Medicine and Mongan Institute Center for Aging and Serious Illness, Massachusetts General Hospital, Boston, MA, United States

ARTICLE INFO

Keywords:
Digital phenotyping
Health outcomes
Scoping review
Smartphones

ABSTRACT

Background: Digital phenotyping may detect changes in health outcomes and potentially lead to proactive measures to mitigate health declines and avoid major medical events. While health-related outcomes have traditionally been acquired through self-report measures, those approaches have numerous limitations, such as recall bias, and social desirability bias. Digital phenotyping may offer a potential solution to these limitations.
Objectives: The purpose of this scoping review was to identify and summarize how passive smartphone data are processed and evaluated analytically, including the relationship between these data and health-related outcomes.
Methods: A search of PubMed, Scopus, Compendex, and HTA databases was conducted for all articles in April 2021 using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Review (PRISMA-ScR) guidelines.
Results: A total of 40 articles were included and went through an analysis based on data collection approaches, feature extraction, data analytics, behavioral markers, and health-related outcomes. This review demonstrated a layer of features derived from raw sensor data that can then be integrated to estimate and predict behaviors, emotions, and health-related outcomes. Most studies collected data from a combination of sensors. GPS was the most used digital phenotyping data. Feature types included physical activity, location, mobility, social activity, sleep, and in-phone activity. Studies involved a broad range of the features used: data preprocessing, analysis approaches, analytic techniques, and algorithms tested. 55% of the studies (n = 22) focused on mental health-related outcomes.
Conclusion: This scoping review catalogued in detail the research to date regarding the approaches to using passive smartphone sensor data to derive behavioral markers to correlate with or predict health-related outcomes. Findings will serve as a central resource for researchers to survey the field of research designs and approaches performed to date and move this emerging domain of research forward towards ultimately providing clinical utility in patient care.

* Corresponding authors at: Frances Payne Bolton School of Nursing, Case Western Reserve University, 9501 Euclid Ave, Cleveland, OH 44106, United States (K. Lee). College of Gyedang General Education, Sangmyung University, 20 Hongjimun 2-gil, Jongno-gu, Seoul 03016, Republic of Korea (T. C. Lee).
E-mail addresses: kxl916@case.edu (K. Lee), tcleethephilosopher@smu.ac.kr (T.C. Lee).

https://doi.org/10.1016/j.ijmedinf.2023.105061
Received 3 August 2022; Received in revised form 10 February 2023; Accepted 24 March 2023
Available online 30 March 2023
1386-5056/

1. Introduction

Measurements of health-related outcomes over time, such as mobility, mood, and physical distress, can help detect signs of decline and trigger steps to intervene [1] and ultimately enhance health outcomes [2]. Measuring these health-related outcomes is typically accomplished by the participant actively self-reporting measures and answering validated questionnaires [3]. However, acquiring these data using a self-report approach has several limitations, including recall and social desirability bias [4]. One potential solution to these limitations is digital phenotyping, or the passive modeling of smartphone data, such as Global Positioning System (GPS) and accelerometer data. These data do not require deliberate participant input yet still detect moment-to-moment user actions, potentially signaling current health states.

Digital phenotyping has been defined as the "moment-by-moment quantification of the individual-level human phenotype *in situ* using data from personal digital devices," specifically from mobile devices such as smartphones [5]. Smartphones are now nearly ubiquitous in North America and provide a rich source of data about human behavior [6]. Using passively collected smartphones data, digital phenotyping can be used to model variations in an individual's behavior over time, including their mobility (via GPS data), sociability (via text message and telephone logs), and sleep (via accelerometer data and screen activity logs) [7]. Digital phenotyping can model changes in behavior at the individual level while accounting for between-individual heterogeneity because each individual tends to interact with their smartphones uniquely [8–10]. Furthermore, passive smartphone data can be linked with health-related self-report measures and clinic physical examination findings [7]. Emerging digital phenotyping approaches could lead to clinically useful digital "biomarkers" that enhance diagnostic assessment, tailor treatment choices, improve disease monitoring and form the basis of new interventions and care delivery models [11].

Although several reviews have examined the use of smartphones generally for health and well-being [12–15], to the best of our knowledge, how passive smartphone data are processed and used in digital phenotyping, including their relationship to health-related outcomes, has not been systematically reviewed. Thus, the purpose of this review is to identify and summarize how passive smartphone data are processed and evaluated as behavioral markers, defined as behaviors, thoughts, feelings, traits, or states identified using personal sensing [16].

2. Methods

We used the Joanna Briggs Institute (JBI) approach, Arskey and O'Malley framework [17], and the PRISMA extension of scoping reviews (PRISMA-ScR) [18] to guide this scoping review [19].

2.1. Search strategy

Keywords were identified by reviewing relevant studies and MeSH terms. As shown in Table 1, search strings were derived using key words such as digital-phenotyp*, digital-trac*, Digital-sens*, Smartphone*, sensor data*, "Health Outcome Assessment", "Quality of life", "Health", Well-being, and "health outcomes". Databases searched included PubMed (MEDLINE), Scopus, Compendex, and Health Technology Assessment (HTA) with no specified date limits. The search was performed in April 2021. Two university research librarians helped in the literature search (Table 1).

2.2. Inclusion and exclusion criteria

Studies included were: (a) peer-reviewed reports of original research, (b) conference proceedings, (c) published in English, and (d) focused on collecting passive smartphone data for digital phenotyping. Studies excluded were: (a) methodologically focused and did not collect data (e.g., design overviews, conceptual frameworks), (b) posters, meeting abstracts, protocols, or literature reviews, (c) non-smartphone (e.g., clinical lab equipment) studies, and (d) intervention or app development studies.

Table 1
Search Terms for the Literature Search.

Database	Search Terms	Results
PubMed	(digital-phenotyp*[tiab] OR digital-trac*[tiab] OR Digital-sens*[tiab] OR Smartphone*[tiab] OR sensor data*[tiab]) AND ("Health Outcome Assessment"[Mesh] OR "Quality of life"[Mesh] OR "Health"[Mesh:NoExp] OR Wellness[tiab] OR Wellbeing[tiab] OR Well-being[tiab] OR "quality of life"[tiab] OR "health outcomes"[tiab] OR "Health"[tiab]) AND ("Mental Health"[Mesh] OR " Physical Health"[Mesh] OR "Social Health"[Mesh])	821
Scopus	(digital-phenotyp* OR digital-trac* OR Digital-sens* OR Smartphone* OR sensor data*) AND ("Health Outcome Assessment" OR "Quality of life" OR "Health" OR Wellness OR Wellbeing OR Well-being OR "health outcomes") AND ("Mental Health" OR " Physical Health" OR " Social Health") AND NOT INDEX(medline)	1891
Compendex	(digital-phenotyp* OR digital-trac* OR Digital-sens* OR Smartphone* OR sensor data*) WN KY) AND (("Health Outcome Assessment" OR "Quality of life" OR Health OR Wellness OR Wellbeing OR Well-being OR "health outcomes") WN KY)) AND (("Mental Health" OR "Physical Health" OR "Social Health") WN KY))) AND ([70] WN LA))	380
Health Technology Assessment	(digital-phenotyp* OR digital-trac* OR Digital-sens* OR Smartphone* OR sensor data*) AND ("Outcome Assessment, Health Care" OR "Outcome and Process Assessment, Health Care" OR "Health Services Research" OR "Quality of life" OR "Health" OR Wellness OR Wellbeing OR Well-being OR "quality of life" OR "health outcomes" OR "Health") AND ("Mental Health" OR " Physical Health" OR " Social Health" OR "social health" OR "Mental health" OR "physical fitness")	78

2.3. Review process

Articles were imported from EndNote© into Rayyan (https://rayyan.qcri.org/) for the screening and selection process [20] (See Fig. 1). Titles and abstracts of all articles were screened by two assessors (KL and TCL) for eligibility. Screened-in full text articles were obtained and inspected for inclusion by the assessors, independently.

2.4. Data extraction

Data extracted included citation (study site), sample (study length), topical focus, measures, smartphone sensors (application), analytical methods, and main results (Table 2).

2.5. Data analysis

The results were reported as a narrative summary. Consistent with directive content analysis [21], study team members discussed the findings listed in the summary table to ensure that the categories and the contents underneath were expansive enough to capture the findings of each article.

3. Results

The initial search yielded 3,170 articles. After removing duplicates and assessing each for eligibility, 138 article full text articles were reviewed, resulting in 40 articles that met the inclusion and exclusion

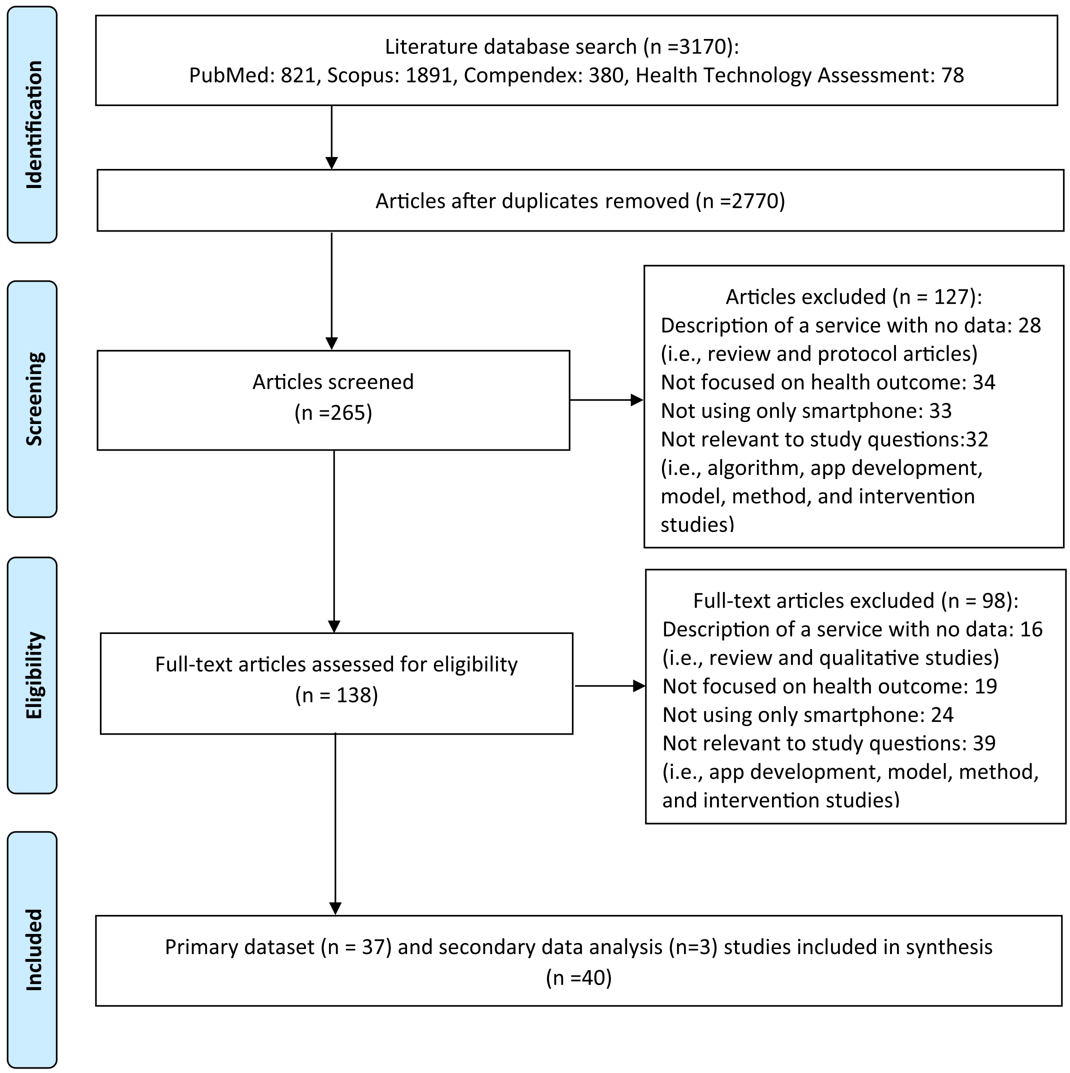

Fig. 1. Flow diagram of study selection process.

criteria (Fig. 1).

Findings were organized under six topics: 1) characteristics of the included studies, 2) data collection, 3) feature extraction (see appendix), 4) data analytics, 5) behavioral markers, and 6) health-related outcomes.

3.1. *Characteristics of the included studies*

Among the 40 studies, 26 were conducted in the USA, six in Europe, two in Australia, and five in Asia. See Table 2 for study characteristics. Of those studies that reported a study timeframe, half (n = 20) lasted 1–2 weeks (n = 14) or 4 weeks (n = 6). Six (15%) studies lasted 5–8 weeks and another 15% (n = 6) lasted > 10 weeks. Study sample size ranged from 7 to 816. Sixty-five percent of the reviewed studies (n = 26) had a sample size of less than 53 participants. 12 (30%) of the selected studies targeted college students. A number of different smartphone apps were used for passive data collection; these included study team-developed software (n = 6), Purple Robot (n = 5), Sensus (n = 4), Beiwe (n = 3), StudentLife (n = 3), CrossCheck (n = 2), mindLAMP (n = 2), and myriad others in single studies.

3.2. *Data collection*

Data collection in all studies included both passive raw smartphone sensor data and participant-reported data, which are considered the "ground truth" thereby serving as validation of passive data [22].

3.2.1. *Smartphone-based passive data collection*

Table 3 shows the smartphone passive data used in studies. GPS sensor data were the most frequently collected type. Over half of studies (n = 25, 62.5%) collected GPS data about users' locations and movements [6,23–46]. For 3 studies, GPS was the only data type collected or gathered in conjunction with patterns of Wi-Fi connections [25,39,46], or GPS in conjunction with accelerometer data (n = 11) [6,23,25,27,32,33,36,37,39,43,45]. Regarding the users' location, Bluetooth (n = 3) which indicates the distance among people was also used in an attempt to examine levels of sociability [47–49]. In fact, of the three studies that used Bluetooth, two aimed to detect users' physical encounters [47,48]. Microphone data were used in 5 studies (12.5%) and gyroscope data in one study. Microphone data were used to infer social activity [25], location [23,34,50], and physical activities [39], while gyroscope data were used to detect daily physical activities (i.e., resting, exercising, running, walking) [37].

Other smartphone data collected included the frequency of using text messages (n = 17) [23,24,27,29,31–34,39,44,49,51–56], the number of calls (n = 21), timings of screen on/off [29,31,33,38–40,56,57], ambient light values [25,29,39,43,58], and time spent on the phone [41,52,59]. Data in terms of battery level and status [31], browser usage

Table 2
Studies included in the review reporting digital phenotyping to understand health-related quality of life using smartphone.

Citation (Study Site)	Sample (Study Length)	Topical Focus	Measures	Smartphone Sensors (Application)	Analytical Methods	Main Results
[23] Abdullah et al., 2016 (USA)	7 bipolar disorders (4 weeks)	Social rhythm	Social rhythm metrics −5	GPS, Accelerometer, Call/text logs, Microphone (Mood Rhythm)	Logistic regression Classification (SVM)	Root-mean-square error (RMSE) 0.92 for personalized, 1.40 for the generalized. 85% (precision, binary), 86% recall
[24] Barnett et al., 2018 (USA)	15 schizophrenias (Up to 3 months)	Schizophrenia relapse symptoms	Depression Sleep quality Warning symptoms	GPS, Call/text logs (Beiwe)	Multivariate time series linear classifiers	The rate of behavioral anomalies detected in the 2 weeks prior to relapse was 71% higher than the rate of anomalies during other time periods.
[25] Ben-Zeev et al., 2015 (USA)	47 students (10 weeks)	Depression symptom Stress Loneliness	PHQ-9 Perceived Stress Scale (PSS) UCLA Loneliness Scale (UCLA-LS)	GPS, Wi-Fi, Accelerometer, Sound, Light sensor, Microphone (Developed software for study)	Mixed-effects linear Penalized functional regression	Stress: geospatial activity and sleep duration ($p < .05$, respectively). Depression: speech duration, geospatial activity, and sleep duration ($p < .05$, respectively) Loneliness: kinesthetic activity ($p < .01$)
[6] Berry et al., 2019 (USA)	23 amyotrophic lateral sclerosis (24 weeks)	Amyotrophic lateral sclerosis	Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R)	Accelerometer, GPS, Call logs (Beiwe)	Linear regression	The mean pause time in speech had increased by 0.02 sec per month across the sample.
[47] Boonstra et al., 2015 (Australia)	14 (1 week)	Social networks	Non-negative matrix factorization (NMF)	Bluetooth (Purple Robot)	Correlation	Correlation between network metrics and the number of times a smartphone scanned.
[48] Boonstra et al., 2017 (Australia)	63 (4 weeks)	Social network	PHQ-9 Generalized Anxiety Disorder (GAD)-7	Bluetooth (BluetoothManager)	Fruchterman-Reingold	Social networks of proximity were estimated from Bluetooth data and 95% of the edges were scanned at least every 30 min.
[26] Boukhechba et al., 2018a (USA)	228 students (2 weeks)	Social anxiety	Social Interaction Anxiety Scale (SIAS)	GPS (Sensus)	Classification (Neural networks) Regression	85% (accuracy, binary), 85% F1 score Social anxiety score (on a scale of 0–80) with RMSE 7.06
[27] Boukhechba et al., 2018b (USA)	72 students (2 weeks)	Social Anxiety Depression symptom Positive/ Negative Affect	SIAS Depression Anxiety Stress Scales (DASS) Positive and Negative Affect Schedule (PANAS)	GPS, Accelerometer, Calls/texts logs (Sensus)	Correlation	Students' social anxiety, depression and affect levels are associated with their mobility, activity levels, and communication patterns.
[28] Canzian & Musolesi, 2015 (UK)	28 (14 days)	Depressive mood	PHQ-9	GPS (MoodTraces)	Classification (SVM)	96% (precision, binary), 94% recall
[29] Cao et al., 2020 (USA)	13 families of adolescent with major depressive disorder (8 weeks)	Depressive symptom	PHQ-9 Hamilton Rating Scale for Depression (HAM-D) Hamilton Anxiety Rating Scale (HAM-A)	GPS, Calls/text logs, Light sensor, Screen usage (MobileSens)	Linear regressor A support vector regressor	88% (accuracy, binary) in teens' PHQ-9 score 90% accuracy in teen and parents' PHQ-9 score RMSE 3.38 in teens vs 3.47 in parents.
[59] DeMasi et al., 2017 (USA)	53 (8 weeks)	Well-being	Mood and energy	Accelerometer, Phone usages (Funf Open Sensing Framework)	Classification (Logistic regressions)	95% (accuracy, binary)
[51] Dissing et al., 2020 (Denmark)	816 (4 weeks)	Well-being	UCLA-LS Major Depression Inventory (MDI) Knowledge, Skills and Abilities (KSQ)	Call/text logs (Developed software for study)	Multiple linear regression	In baseline, a higher number of smartphone interactions was associated with lower levels of loneliness as well as lower levels of disturbed sleep for men. In follow-up analyses, a high vs low level of smartphone interaction was associated with an increase in loneliness and depressive symptoms over time for women.
[30] Fraccaro et al., 2019 (UK)	21 (10 days)	Social function	Activity diaries	GPS (GPSLogger)	Clustering algorithms and semantic enrichment	75% precision, 60% recall All thresholds: F1 score = 0.65. RMSE = 1.1 (SD 0.3)
[31] Gao et al., 2016 (China)	127 (30 days)	Social anxiety Loneliness	SIAS UCLA-LS	GPS, Call/text logs, App usage, Screen on/off, Charging (SOLVD)	Wilcoxon-Mann-Whitney test	Correlation among smartphone usage and social anxiety and loneliness
[32] Gong et al., 2019 (USA)	52 (2 weeks)	Social anxiety	SIAS	GPS, Accelerometer, Call/text logs (Sensus)	Correlation Effect size	Behavioral metrics observed in temporal proximity to phone calls have stronger associations with social anxiety scores than

(continued on next page)

Table 2 (*continued*)

Citation (Study Site)	Sample (Study Length)	Topical Focus	Measures	Smartphone Sensors (Application)	Analytical Methods	Main Results
						metrics observed in temporal proximity to text message events.
[50] Harari et al., 2017 (USA)	48 students (10 weeks)	Well-being	Socio-demographic characteristics	Accelerometer, Microphone (StudentLife)	Latent growth curve	Stability estimates were moderate to high for activity (r mean = 0.66) and sociability (r mean = 0.72)
[33] Henson et al., 2020 (USA)	88 (45 schizophrenias, 43 healthy control) (3 months)	Social rhythm	PHQ-9 GAD-7 Psychosis symptoms Sleep Sociability	GPS, Accelerometer, Screen on/off, Call/text logs (mindLAMP/Beiwe)	Correlation	Social rhythms were negatively associated with symptoms of anxiety, depression, psychosis, and poor sleep (Spearman $\rho = -0.23$ to -0.30, $p < 0.001$) in schizophrenias. In healthy control, more stable social rhythms were positively correlated with symptomatology (Spearman $\rho = 0.20$ to 0.44, $p < 0.05$).
[57] Henson et al., 2021 (USA)	88 (54 schizophrenias, 34 healthy control) (>5 days)	Schizophrenia relapse symptoms	PHQ-9 GAD-7 Sleep Sociability	Screen on/off (mindLAMP/Beiwe)	Linear regression, Specification curve analysis	An association between smartphone screen time metrics and cognition (adjusted R2 = 0.107, P <.001) in patients with schizophrenia was found. Specification curve analysis revealed a wide range of heterogenous associations with screen time from very negative to very positive.
[34] He-Yueya et al., 2020 (USA)	61 schizophrenias (14 days)	Behavioral stability	10 items: how you've been doing over the last few days.	GPS, Microphone, Call/text logs (CrossCheck)	Classification (Gradient boosted regression trees)	2.468 Mean absolute error (MAE) (accuracy, binary) Stability Index (MAE = 2.556) Greater stability in social activity (e.g., calls and messages) were associated with lower symptoms, and greater stability in physical activity (e.g., being still) appeared associated with elevated symptoms.
[35] Huang et al., 2016 (USA)	18 (10 days)	Social anxiety	SIAS	GPS (Developed software for study)	Linear regression	Least Absolute Shrinkage and Selection Operator (LASSO) ($\lambda = 4.4$) demonstrates the negative relationship between visiting religious locations and social anxiety level.
[53] Huang et al., 2017 (USA)	52 students (2 weeks)	Social anxiety	SIAS	Accelerometer, Call/text logs (Sensus)	Effect-size analysis	The results show substantially different behavioral markers prior to outgoing phone calls when comparing individuals with high and low social anxiety.
[60] Kelly et al., 2016 (Ireland)	541 (>72 h)	Health status	Short form (SF)-36	Accelerometer (Health-U)	Correlation	While results have shown a statistically significant correlation between duration of activity and health status (r = 0.042), these correlations are not strong enough to make accurate predictions about a persons' health status.
[54] King et al., 2020 (Netherlands)	27 students (6 weeks)	Mood	Center for Epidemiological Studies Depression Scale (CES-D)	Call/SMS logs, App usage (iYouVU)	Classification (SVM)	65% (Accuracy, 4 classes), 61.8% F-score * Dataset provided by Asselbergs et al. (2016) from Netherlands
[36] Masud et al., 2020a (Bangladesh)	33 (11 weeks)	Depression severity level	PHQ-9	GPS, Accelerometer (Data Collector)	Linear regression Classification (SVM)	Root-mean-square deviation (RMSD) 3.256 87.2% (accuracy, binary)
[37] Masud et al., 2020b (Bangladesh)	33 (11 weeks)	Depression symptom	Quick Inventory of Depressive Symptomatology–Self-report (QIDSSR)16	GPS, Gyroscope, Accelerometer (Purple Robot)	Linear regression Classification (Quadratic discriminant)	RMSD 3.117 95% (accuracy, binary)

(*continued on next page*)

Table 2 (*continued*)

Citation (Study Site)	Sample (Study Length)	Topical Focus	Measures	Smartphone Sensors (Application)	Analytical Methods	Main Results
[52] Messner et al., 2019 (Germany)	157 students (8 weeks)	Well-being	Stress Drive Mood	Call/SMS logs, Phone usage (Insights)	Multivariate multilevel models	Stress: number of SMS (–3.539, SE = 0.937) Mood: total usage time (–0.019, SE = 0.004) and call duration (–0.016, SE = 0.007) Drive: Facebook usage time (–0.127, SE = 0.041)
[49] Pulekar & Agu, 2016 (UAS)	9 students (2 Weeks)	Social loneliness	Big-Five personality traits UCLA-LS	Call/text logs, App usage, Bluetooth, Wi-Fi, Browser usage (Socialoscope)	Classification (J48)	98% (accuracy, binary), while factoring in user personality types.
[38] Rhim et al., 2020 (South Korea)	78 students (4 months)	Subjective well-being	Concise Measure of Subjective Well-Being (COMOSWB)	GPS, App usage, Screen on/off (Developed software for study)	Hierarchical regression Classification (High performance models)	The significance of user attributes (e.g., personality, self-esteem) on subjective well-being and salient factors derived from smartphone data (e.g., time spent on campus, ratio of standing/sitting stationary, expenses) that significantly account for subjective well-being. 71% (the average F1- macro score for all users was 38%)
[55] Ryan et al., 2020 (USA)	26 bipolar disorder & 12 healthy control (28 days)	Mood Energy Intellect (thoughts)	mood	Call/text logs (Developed software for study)	Multivariate mixed effect	An increase in rapid thoughts over time was associated with a decrease in outgoing text messages (β = –0.02; P =.04), and an increase in impulsivity self-ratings was related to a decrease in total call duration (β = –0.29; P =.02).
[39] Saeb et al., 2017 (USA)	208 (6 weeks)	Depression symptom Anxiety	PHQ-9 GAD-7	GPS, Light sensor, Microphone, Screen on/off, Accelerometer, Call/text logs, Wi-Fi (Purple Robot)	Classification (Decision trees with the gradient boosting optimization)	Foursquare: 0.62 (Average AUC) phone sensor: 0.84 (Average AUC) Foursquare + phone sensor: 0.88 (Average AUC, binary) Significant relationships between the time spent in certain locations and depression and anxiety were identified, although these relationships were not consistent.
[40] Saeb et al., 2016 (USA)	48 students (10 weeks)	Depression symptom	PHQ-9	GPS, Screen on/off (StudentLife)	Correlation	Correlation between GPS features and clinical PHQ-9 *Dataset provided by Wang et al. (2014) from USA
[41] Saeb et al., 2015a (USA)	28 (2 weeks)	Depression symptom	PHQ-9	GPS, Phone usage (Purple Robot)	Classification (Logistic regression)	86.5% (accuracy, binary)
[42] Saeb et al., 2015b (USA)	18 (2 weeks)	Depression symptom	PHQ-9	GPS (Purple Robot)	Correlation	Correlation between location features and clinical PHQ-9
[43] Sarda et al., 2019 (India)	47 diabetes (2 weeks)	Diabetes base on depression	PHQ-9	Accelerometer, GPS, Light sensor, Call logs (Developed software for study)	Classification (XGBoost)	79.07% (accuracy, binary [95% CI: 74%, 84%])
[44] Singh et al., 2018 (USA)	50 (10 weeks)	Individual's risk propensity	Risk Propensity Personality traits	Call/text logs, GPS (Funf platform)	Classification (SVM)	78% (precision, binary), 90% recall
[56] Stanislaus et al., 2020 (Denmark)	75 bipolar disorder, 15 healthy control, & 32 unaffected relatives (>1 month)	Activity Depression symptom Mania	International Physical Activity Questionnaire (IPAQ) HAM-D FAST	Accelerometer, Screen on/off, Call/text logs (Monsenso)	Linear mixed model	Patients with bipolar disorder had decreased physical (number of steps) and social activity (more missed calls) but a longer call duration compared with healthy control. Unaffected relatives also had decreased physical activity compared with healthy control but did not differ on daily self-reported activity or social activity.

(*continued on next page*)

Table 2 (*continued*)

Citation (Study Site)	Sample (Study Length)	Topical Focus	Measures	Smartphone Sensors (Application)	Analytical Methods	Main Results
[61] Staples et al., 2017 (USA)	17 schizophrenias (6 weeks)	Sleep	Pittsburgh Sleep Quality Index (PSQI)	Accelerometer (Beiwe)	Multiple linear regression	0.75 MAE accuracy
[45] Thakur & Roy, 2020 (India)	45 students (10 weeks)	Subjective feeling of psychological well-being	GT-PHQ-9 PSS PSQI UCLA-LS	GPS, Accelerometer, Call logs. (StudentLife)	Classification (Logistic regression)	Stress: 82.6% (AUC, binary) Depression: 74% (AUC, binary) *Dataset provided by Wang et al. (2014) from USA
[58] Wang et al., 2017 (USA)	36 schizophrenias (30 days)	Schizophrenia relapse symptoms	7-item Brief Psychiatric Rating Scale (BPRS)	Audio amplitude, Accelerometer, Light Sensor, App usages, Call logs (CrossCheck)	Classification (Gradient Boosted Regression Trees)	1.45 MAE (accuracy, 3 classes) with r = 0.70, p < 0.0001.
[46] Yue et al., 2020 (USA)	79 students (15 days)	Depression symptom	PHQ-9	GPS, Wi-Fi (LifeRhythm)	Classification (SVM)	83% (Precision, Binary) F1 scores (up to 0.76 compared to 0.5 before data fusion)

Table 3
Summary of the smartphone-sensor data for behavioral markers or health-related outcomes.

Domain	Health related outcomes	Behavioral markers	Type of feature	Smartphone Sensors
Mental health	Early symptoms of relapse in schizophrenia		Mobility [24], Sociability [24,58], Screen [57], Physical activities [58], Sleep [58]	GPS [24], Accelerometer [58], Call/text logs [24], Call logs [58], Screen on/off [57], Audio amplitude [58], Light sensor [58], App usages [58]
	Depression symptoms		Location [25,37,39,40,42,46], Sleep [25,43], Speech duration [25], Mobility [28,36,43,59], Social activity [43,56,59], Physical activity [36,37,43,56], Daily-life behavior [41]	GPS [25,28,36,37,39–43,46,59], Wi-Fi [25,39,46], Accelerometers [25,36,37,39,43,56], Gyroscope [37], Sound [25], Microphone [25,39], Call/text logs [39,56,59], Call logs [43], Light sensor [25,39,43,59], Screen [39,40,56,59], Phone usage [41]
		Stress	Sleep duration [25], Location [25]	GPS [25], Wi-Fi [25], Accelerometers [25], Sound [25], Light sensor [25], Microphone [25]
		Loneliness	Kinesthetic activity [25], Social activity [25,49], Phone usage [29], Behavioral dynamics [29]	GPS [29], Wi-Fi [49], Accelerometers [25], Sound [25], Call/text logs [29,49] App usage [29,49] Bluetooth [49], Browser usage [49], Screen [29], Light sensor [25]
		Social anxiety	Mobility [26,27,35], Physical activity [27], Social activity [27,53], Phone usage [29], Behavioral dynamics [29,32,53]	GPS [26,27,29,32,35], Accelerometers [27,32,53], Call/text logs [27,29,32,53], Screen [29], App usage [29] Battery level [29]
		Mood	Levels of valence (mood), and arousal (drive) [54], Daily behaviors [55]	Call/text logs [54,55], App usage [54]
Physical health	Monitoring of Amyotrophic lateral sclerosis		Speech [6]	Speech [6]
		Social rhythms	Physical activity [23], Location [23] Mobility [23], Social activity [23], Circadian Routine and Weekend Day Routine [33]	GPS [23,33], Accelerometer [23,33], Call/text logs [23,33], Microphone [23], Screen [33]
		Behavioral stability	Physical activity [34], Speech [34], Phone usage [34]	GPS [34], Call/text logs [34], Microphone [34]
		Sleep disturbances	Sleep duration [61]	Accelerometer [61]
Social health		Social networks	A proxy for social ties [47,48]	Bluetooth [47,48]
		Social function	Daily activities [31]	GPS [31]
	Well-being		Sleep [51], Physical activity [38,45,50,51], Social activity [30,45,50], Phone usage [38,52], Location [38], Mobility [45]	Accelerometer [45,50,51], Phone usages [51,52], Call/text logs [30,52], Call logs [45], Microphone [50], GPS [38,45], App usage [38], Screen [38]
		Health status	Mobility [60]	Accelerometer [60]
		Individual's risk propensity	Location [44], Social activity [44]	Call/text logs [44], GPS [44]

[49], and app usage [31,38,49,54,58] were also collected in 15% (n = 7) of the selected studies.

Most studies (n = 28, 70%) collected data from a combination of sensors. Eight studies collected data from ≥ 5 sensors [23,25,29,31,33,39,49,58]. Studies with > 3 sensors (n = 20, 50%) commonly relied on machine learning prediction models to process and interpret data. Twelve studies (30%) collected data from one sensor.

3.2.2. *Smartphone-based self-reported data & questionnaires*

Nine studies indicated the type of self-report data used as ground truth [23,28,30,31,38,41,45,46,58], while the remaining studies did not report ground truth data. Most studies (n = 38) administered participant-reported questionnaires on participants' smartphones.

3.3. *Feature extraction*

Feature extraction means transforming the raw data acquired by sensors into features that allow for more meaningful information [16].

3.3.1. *Physical activity*

Accelerometer data were used to model users' physical activity [23,27,34,36,37,38,43,45,50,51,56,58]. These studies aimed to recognize basic daily activities such as walking [34,36,37,39,45,50,58,59], standing or sitting [34,37,38,39], running [37,39,45,50,58,59], exercising [36,37], lying down [59], riding a bike [34,39,58,59] being in a vehicle [34,39,58], or being stationary [23,25,34,36,50,58,60]. In addition, physical activities were explored by detecting steps taken

[36,37,56] and distinguishing physical activity from activities of daily living [30,33,61].

3.3.2. Location and mobility

Derivation of users' mobility and location was computed in 24 studies using data from GPS, Wi-Fi, and/or Bluetooth sensors, most (n = 19) of which aimed to identify basic daily mobility trajectories such as users' daily range of movement as reported by GPS data. These studies assessed the number of location clusters (n = 11), total distance (n = 9), location entropy (n = 8), time spent at home (n = 5), normalized entropy (n = 4), and transition time (n = 3) (see appendix). One study that included 15 patients with schizophrenia showed how passive data—including radius of gyration, maximum diameter, the standard deviation of flight length, and average flight duration—were potentially able to detect early warning signs of relapse [24].

3.3.3. Social activity

Social activity was examined in 12 studies using call and text logs to detect social interaction [23,24,27,30,35,43–45,49,56,58,59], the number of outgoing and incoming texts (n = 4), total outgoing and incoming text length (n = 4), number of outgoing and incoming calls (n = 6), and the total duration of outgoing and incoming calls (n = 4). Bluetooth was used to examine social ties [47,48]. The employment of microphones was to assess silence, noise, and voices of sociability behavior [50]. Six studies estimated communication patterns, such as the number of text messages sent and received. Speech was detected with a microphone [6,23,27,34]; analyses evaluated conversation time [23,34], speaking rate [6], pitch variations [6,23], and speech duration [27,34].

3.3.4. Sleep

Users' sleep was examined [25,33,43,58,59] by detecting sleep duration [25,58], bedtime and rise time [58], and sleep patterns [33,43,59] using accelerometers [25,59], sound [25], light sensors [25,43,58], screen on/off [33], and smartphone usage [59].

3.3.5. In-phone activity

Smartphone usage patterns were collected [31,38,41,49,52,54], including screen on/off [31], social media activity [31,38,49,52,54], browser usage [31,49], and application usage [31,38,41,49,52]. These data were correlated with participant-reported measures of well-being [38,52], social anxiety [31], loneliness [31,49], mood [54], and depressive symptoms [41]. For example, Saeb et al. (2015a) reported that phone usage duration and usage frequency were correlated with the severity of depressive symptoms from participant-reported PHQ-9 in 28 adult participants (r = 0.54, p =.011, and r = 0.52, p =.015, respectively) [41].

3.4. Data analytics

The main analytical methods used for passive data analysis were regression models and predictive modeling with machine learning techniques [62].

3.4.1. Type of machine learning method

Most studies (n = 33) reported on supervised machine learning (i.e., classification and regression) to classify data or predict outcomes accurately using self-reported data as "ground truth". Unsupervised learning (i.e., clustering) was used only in a few studies (n = 4), while 10 studies implemented both supervised and unsupervised machine learning approaches. A very broad range of classification of algorithms of machine learning techniques were applied (n = 19): support vector machines [23,28,36,44,46,54], logistic regression [41,45,59], linear classifiers [6,24], gradient boosted regression trees [34,58], neural networks [26], support vector regressor [29], quadratic discriminant [37], high-performance models [38], decision trees with the gradient boosting optimization [39], XGBoost [43], J48 [49], linear dynamic model [53], linear regressor [57], and multiple linear regression [61].

Among the included studies, unsupervised learning algorithms were often used to pre-process sensor data before using supervised methods for further processing. In this review, 14 studies chose to use clustering algorithms, including *k*-means [26,27,29,33,38–42] and density-based spatial clustering of applications with noise (DBSCAN) [23,30,36,37,46].

3.4.2. Cross validation strategies

The most used types of cross-validation were leave-one-out [26,28,36,37,41,43,44,46,58,59,61], 10-fold [23,26,34,35,39], and *k*-fold [45]. Four studies did not report cross-validation. In addition, the most reported evaluation metrics were accuracy [26,29,34,36,37,41,43,49,54,58,59,61], sensitivity [23,28,30,44], area under the curve (AUC) [39,45], precision [23,28,30,44,46], and F1 score [26,30,38,46]. The rates of metrics used included: accuracy from 65 to 98%, sensitivity (i.e., recall) from 60 to 94%, precision from 75 to 96%, AUC from 62 to 84%, F1 from 65 to 85%.

3.5. Behavioral markers

Behavioral markers are higher-level features, which reflect behaviors, cognitions, and emotions translated from low-level features (e.g., social activity) and sensor data (e.g., call/text logs) [16].

3.5.1. Mood and Stress

Measurements of participants' moods included positive or negative mood valence and arousal (i.e., the intensity an emotion or mood state) [54] and impulsivity self-ratings [55], which have been detected using app usage and call/text log patterns. For instance, a study of 47 students showed an association of daily stress levels (e.g., depression and loneliness) with a smartphone's sensor-derived geospatial activity (using GPS and WiFi; $p < .05$) and sleep duration (using device use data, accelerometer inferences, ambient sound features, and ambient light levels; $p < .05$) by mixed-effects linear modeling [25].

3.5.2. Sleep disturbance

One study of 17 patients with schizophrenia found that patients' sleep duration could be predicted with approximately 75% accuracy using passively collected accelerometer data and actively collected Pittsburgh Sleep Questionnaire Inventory (PSQI) score [61].

3.5.3. Loneliness

Passive data were used to measure the relationship of loneliness with phone usage behaviors [31], kinesthetic activity [25], and users' daily interactions and communications [49]. For example, a study of 9 students over 2 weeks indicated that of the big 5 personality traits, extraversion and emotional stability could be predicted with 98% accuracy through smartphone-sensed loneliness from their communication and interaction patterns (e.g., calls and social media usage) [49].

3.5.4. Social rhythms

Social rhythms refer to the day-to-day variability of daily, habitual behaviors (e.g., mealtimes, bedtimes, and patterns of social interaction) [63]. Abdullah et al. (2016) reported, with 85% accuracy, the automatic assessment of social rhythm metrics for location, mobility, and conversation frequency in the daily lives of 7 patients with bipolar disorder [23]. In another study of 88 patients with schizophrenia that investigated patients' circadian rhythms, social rhythms were negatively associated with symptoms of anxiety, depression, and poor sleep, while more stable social rhythms were positively correlated with symptomatology [33].

3.5.5. Social context

Social anxiety was measured using Bluetooth Scans to quantify the proximity of social networks [47,48] or communication patterns or phone use patterns [31,32]. In one study of 18 college students, social interaction among students was associated with improved academic performance when they were surrounded by peers [35]. Similarly, one study with 21 participants reported that social functioning (e.g., employment, shopping, and social activities) could be predicted with 75% precision with daily life based on GPS data. [30].

3.6. Health-related outcomes

Health-related outcomes are health-related issues and disorders that can be potentially identified with the entire set of features and behavioral markers.

3.6.1. Depression symptoms

Depressive symptoms (n = 12, 30%) were the most investigated health-related measure among self-reported outcomes [25,28,29,36,37,39–43,46,56], most measured by the PHQ-9 [28,29,36,39–43,46]. Researchers used smartphone-sensor data to correlate depressive symptoms with features such as geographical location and mobility [25,28,29,36,37,39,40,42,43,46], physical activity [25,36,37,43,56], social activity [29,43,56], sleep duration [25,43], daily behavior [41], and speech duration [25]. The results of the studies showed an accuracy of 79%-96% in classifying symptoms of depression as seen in Table 2.

3.6.2. Early symptoms of relapse in schizophrenia

Passive smartphone-sensor data were used to investigate symptoms of schizophrenia [57,58], using features such as physical activities [57,58], mobility [57], sleep [58], and sociability [57,58]. The main intent of the studies was to predict early warning signs for relapse (relapse detection rate was 71% higher than in other time periods) [57] or schizophrenia symptoms (predicting a user's Brief Psychiatric Rating Scale score within 1.59 error using only passive sensing) [58].

3.6.3. Monitoring of Amyotrophic lateral sclerosis

One clinical marker reported for physical health using smartphone-sensor data was amyotrophic lateral sclerosis (ALS) [6]. Berry et al. (2019) reported the timing of speech-based call logs and found that the mean pause time during speech increased by 0.02 s/month across 23 ALS patients followed for up to 24 weeks [6].

3.6.4. Well-being

Well-being was associated in five studies using sleep [59], physical activity [50,59], social interactions [50,51], location [38], mobility [60], and phone usage behavior [38,52]. Kelly et al. (2016) conducted a study that tracked 541 participants' movement duration and average stationary time using an accelerometer for at least 74 h, and findings reported a correlation between the duration of activity and health (r = 0.042) [60].

4. Discussion

In this scoping review, we aimed to synthesize the literature, focusing on how researchers have used passive smartphone-sensor data to derive behavioral markers and correlate or predict health-related outcomes. Among the 40 health-related publications using digital phenotyping to analyze passively collected data, 55% of the studies (n = 22) focused on mental health-related outcomes. The findings show an area of imminent opportunity for advancing research and clinical insights because the data and methods used in schizophrenia relapse symptoms, depression symptoms, and mood are highly translatable and relevant to a range of other health related outcomes.

Fig. 2 shows a layered, hierarchical approach to converting raw sensor data into knowledge, in which sensor data are converted into features that are integrated to estimate behaviors, emotions, and health-related outcomes. Automated sensing enables ubiquitous and unobtrusive sensing of daily life activities and behaviors with the help of smartphones that generate raw sensor data [64]. The layers of sensors represent inputs to the sensing platform in the form of raw smartphone-sensed data. The feature layer stands as a construct that contains information from a reliable measurement by sensor data, such as mobility and location. Behavioral markers are higher-level features that are measured using low-level features and sensor data [16]. For instance, Abdullah et al. (2016) reported that a behavioral marker for circadian sleep rhythm could include features such as bedtime and waketime, sleep duration, and phone usage. Markers of sleep quality might also include ambient sound features but may also include bedtime and wake

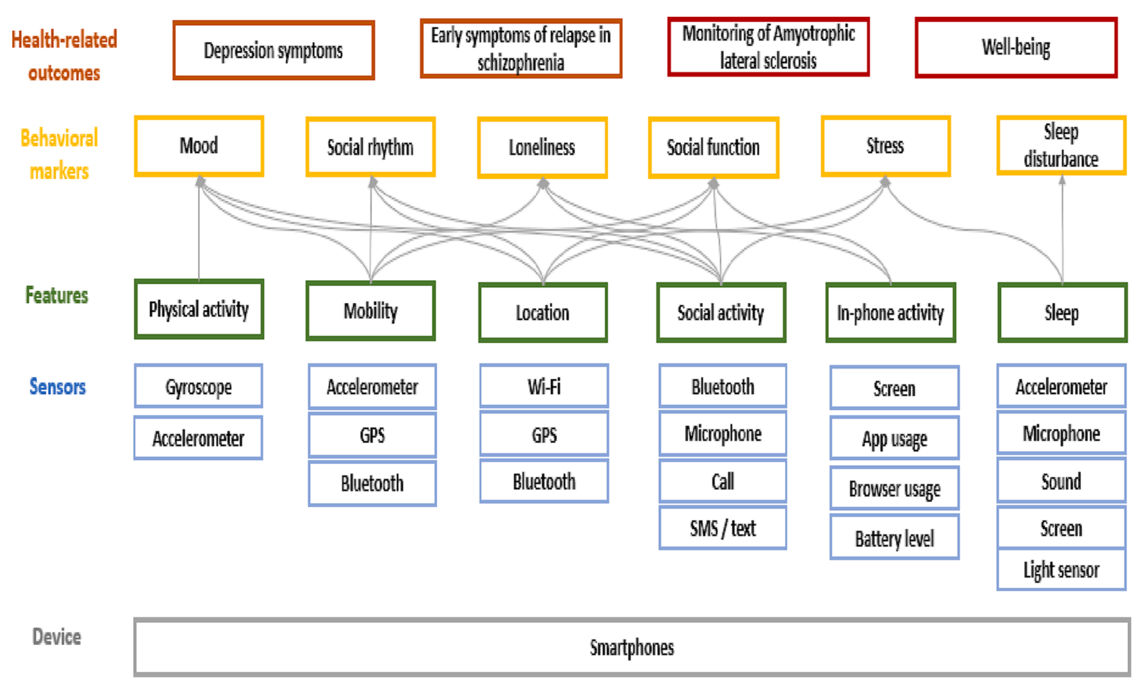

Fig. 2. Hierarchical stream of digital phenotyping for health-related outcomes using a smartphone.

time [23]. A limited set of features have been used in the health-related outcomes layer to predict symptoms of early warning signs of relapse.

This review demonstrated that the study length of passive smartphone-sensor data collection varied considerably across studies, ranging from one week to three years. In general, longer-term data collection is considered more advantageous due to the potential for producing richer information about patterns that may not be evident in shorter-term studies [65]. In this review, however, findings by Henson and colleagues demonstrated that assessing passive data for a minimum of 5 days was sufficient to infer over 3 months of screen time [57], indicating that it may not always be necessary to collect smartphone data over a long-time frame. Nonetheless, further study is needed on the appropriate time length of data collection as it is likely that the length of data collection for smartphone usage patterns is dependent on the functionalities and variables and outcomes being assessed. Hence, the duration of the study should be carefully considered within the specific context to ensure optimization of passive data monitoring.

The wide range of data analytic approaches reported in this review makes it difficult to compare results across studies (see Table 2). While a limitation to synthesis efforts at this time, it has been a lesser priority in this emerging area compared to the need to experiment and explore different analysis approaches to digital phenotyping; hence, it is not surprising that a variety of different approaches have been used [66]. Machine learning approaches were commonly used in studies in order to handle the large amounts of data yet they often lacked appropriate procedures of validation. As the area matures, it will be important to shift the priority of exploring different analysis approaches to identifying and standardizing the most efficient, accurate, and clinically pragmatic approaches. To aid in this effort going forward, it will be important to have standardized checklists for reporting digital phenotyping research, such as machine learning features, algorithm definitions, missing data, and validation approaches [11]. Further, the use of techniques for Functional Data Analysis (FDA) [67], a subfield of statistics with focus on the analysis of data that can be naturally viewed as smooth curves, was not observed in this review. FDA approaches can be useful for frequent longitudinal data such as passive data collected from smartphones. FDA, however, is a relatively new subfield and therefore unfamiliar to practitioners, which might be the reason it did not appear in this review.

30% of the studies ($n = 12$) in this review were conducted with a focus on college students. College students are particularly well suited to digital phenotyping as smartphone ownership among college-aged adults is higher than any other age group [68]. Students' response rates to active data prompts show that students are accepting of and adherent to digital phenotyping apps on their mobile devices [69]. Future research on the topic should consider adding a monitoring system that may need to be improved to attract a more diverse type of users and meet their expectations. Since the need for consistent monitoring can sometimes be found in older populations with various functional or cognitive abilities, populations may differ from college students in terms of daily living and health care needs. In addition, most of the studies had a sample size of less than 53 participants in this review. Due to the amount of variability resulting from differences in device-usage patterns, lifestyle, and the environment, personal sensing platforms will likely require a large user base to be widely applicable. Moreover, even though passive data has the potential to revolutionize healthcare, only five of the included studies [30,44,45,52,60] mentioned the ethical and privacy issues that stimulate participant protection while fostering innovation based on passive data. To utilize passive data ethically and comprehend the broad ramifications of new technology, clinicians, researchers, and other healthcare professionals need a clear ethical framework.

There may be limitations to this review. Because this review was focused on smartphone phones, diverse health-related studies (e.g., physical health-related) may not have been captured in our search, despite the use of broad terminology. Some may raise an issue regarding the additional limitation inherent to scoping review methodology that quality assessment was not performed on the included studies.

5. Conclusion

The implementation of digital phenotyping may enable the ubiquitous and continuous identification and prediction of individuals' health-related behaviors within the context of their social, mental, and physical functioning, reflecting the lived experiences of people in their natural environments. In this review, we have catalogued the research to date and detailed the approaches of using passive smartphone sensor data to derive behavioral markers to correlate with or predict health-related outcomes. This review has also included the types of phenotypes (e.g., physical activity, social activity and sleep disturbance) that can be captured using smartphone-sensor data. Findings will serve as a central resource for researchers to survey the field of research designs and approaches performed to date and move this emerging area forward towards ultimately providing clinical utility in patient care.

What was already known on the topic

- Digital phenotyping enables the continuous identification and prediction of individuals' health-related behaviors.
- Most studies focused on mental health indicators such as depression symptoms and mood.

What this study added to our knowledge

- We demonstrated a layer of features derived from raw sensor data that can then be integrated to estimate and predict behaviors, emotions, and health-related outcomes.
- Most studies collected data from a combination of sensors. GPS was the most used digital phenotyping data.
- Digital phenotypes have the potential to be used to measure early relapse symptoms and monitor health-related disorders.

CRediT authorship contribution statement

Kyungmi Lee: Conceptualization, Methodology, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Tim Cheongho Lee:** Data curation, Formal analysis, Writing – review & editing, Supervision. **Maria Yefimova:** Writing – review & editing. **Sidharth Kumar:** Writing – review & editing. **Frank Puga:** Writing – review & editing. **Andres Azuero:** Writing – review & editing. **Arif Kamal:** Writing – review & editing. **Marie A. Bakitas:** Writing – review & editing. **Alexi A. Wright:** Writing – review & editing. **George Demiris:** Writing – review & editing. **Christine S. Ritchie:** Writing – review & editing. **Carolyn E.Z. Pickering:** Writing – review & editing. **J. Nicholas Dionne-Odom:** Funding acquisition, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Effort for this review was funded in part by the Cambia Health Foundation (no grant number, 2020; PI: Dionne-Odom). The authors would like to thank Rebecca Billings and Jennifer M. Long for searching for literature at the University of Alabama at Birmingham.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.

org/10.1016/j.ijmedinf.2023.105061.

References

[1] C. Amenomori, T. Mizumoto, H. Suwa, Y. Arakawa, K. Yasumoto, A method for simplified HRQOL measurement by smart devices, Lecture Notes of the Institute for Computer Sciences, Social-Inform. Telecommun. Eng. 279 (2018) 91–98.
[2] S.K. Raina, State of The Globe: Health-related quality of life as health status measure: Time to move on, J. Global Infectious Dis. 11 (2019) 89–90.
[3] C. Kingsley, S. Patel, Patient-reported outcome measures and patient-reported experience measures, Bja Educ. 17 (2017) 137–144.
[4] A. Althubaiti, Information bias in health research: definition, pitfalls, and adjustment methods, J. Multidiscip. Healthc. 9 (2016) 211–217.
[5] J.P. Onnela, S.L. Rauch, Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health, Neuropsychopharmacology 41 (2016) 1691–1696.
[6] J.D. Berry, S. Paganoni, K. Carlson, K. Burke, H. Weber, P. Staples, J. Salinas, J. Chan, J.R. Green, K. Connaghan, J. Barback, J.P. Onnela, Design and results of a smartphone-based digital phenotyping study to quantify ALS progression, Ann. Clin. Transl. Neurol. 6 (2019) 873–881.
[7] D.J. Cote, I. Barnett, J.-P. Onnela, T.R. Smith, Digital phenotyping in patients with spine disease: a novel approach to quantifying mobility and quality of life, World Neurosurg. 126 (2019) e241–e249.
[8] S.H. Jain, B.W. Powers, J.B. Hawkins, J.S. Brownstein, The digital phenotype, Nat. Biotechnol. 33 (2015) 462–463.
[9] I. Nahum-Shani, E.B. Hekler, D. Spruijt-Metz, Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework, Health Psychol. 34 (2015) 1209–1219.
[10] T.R. Insel, Digital phenotyping: A global tool for psychiatry, World Psychiatry 17 (2018) 276–277.
[11] K. Huckvale, J. Torous, M.E. Larsen, Assessment of the Data Sharing and Privacy Practices of Smartphone Apps for Depression and Smoking Cessation, JAMA Netw Open 2 (2019) e192542.
[12] T. Donker, K. Petrie, J. Proudfoot, J. Clarke, M.-R. Birch, H. Christensen, Smartphones for smarter delivery of mental health programs: A systematic review, J. Med. Internet Res. 15 (2013) 239–251.
[13] E. Reinertsen, G.D. Clifford, A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses, Physiol. Meas. 39 (2018) 05TR01.
[14] V.P. Cornet, R.J. Holden, Systematic review of smartphone-based passive sensing for health and wellbeing, J. Biomed. Inform. 77 (2018) 120–132.
[15] F. Gravenhorst, A. Muaremi, J. Bardram, A. Grünerbl, O. Mayora, G. Wurzer, M. Frost, V. Osmani, B. Arnrich, P. Lukowicz, Mobile phones as medical devices in mental disorder treatment: an overview, Pers Ubiquitous Comput. 19 (2015) 335–353.
[16] D.C. Mohr, M. Zhang, S.M. Schueller, Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning, Annu. Rev. Clin. Psychol. 13 (2017) 23–47.
[17] H. Arksey, L. O'Malley, Scoping studies: towards a methodological framework, Int. J. Soc. Res. 8 (2005) 19–32.
[18] PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation, Ann. Intern. Med. 169 (2018) 467-473.
[19] M.D. Peters, C. Godfrey, P. McInerney, C. Baldini Soares, H. Khalil, D. Parker, Scoping reviews, Joanna Briggs Institute reviewer's manual 2015 (2017) 1–24.
[20] M. Ouzzani, H. Hammady, Z. Fedorowicz, A. Elmagarmid, Rayyan—a web and mobile app for systematic reviews, Syst. Rev. 5 (2016) 1–10.
[21] W. Potter, D. Levine-Donnerstein, Rethinking validity and reliability in content analysis, J. Appl. Commun. Res. 27 (1999) 258-284.
[22] A. Trifan, M. Oliveira, J.L. Oliveira, Passive Sensing of Health Outcomes Through Smartphones: Systematic Review of Current Solutions and Possible Limitations, JMIR Mhealth Uhealth 7 (2019) e12649.
[23] S. Abdullah, M. Matthews, E. Frank, G. Doherty, G. Gay, T. Choudhury, Automatic detection of social rhythms in bipolar disorder, J. Am. Med. Inform. Assoc. 23 (2016) 538–543.
[24] I. Barnett, J. Torous, P. Staples, L. Sandoval, M. Keshavan, J.-P. Onnela, Relapse prediction in schizophrenia through digital phenotyping: A pilot study, Neuropsychopharmacology 43 (2018) 1660–1666.
[25] D. Ben-Zeev, E.A. Scherer, R. Wang, H. Xie, A.T. Campbell, Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health, Psychiatr. Rehabil. J. 38 (2015) 218–226.
[26] M. Boukhechba, P. Chow, K. Fua, B.A. Teachman, L.E. Barnes, Predicting social anxiety from global positioning system traces of college students: feasibility study, JMIR mental health 5 (2018) e10101.
[27] M. Boukhechba, A.R. Daros, K. Fua, P.I. Chow, B.A. Teachman, L.E. Barnes, DemonicSalmon: Monitoring mental health and social interactions of college students using smartphones, Smart Health 9–10 (2018) 192–203.
[28] L. Canzian, M. Musolesi, Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis, in: 3rd ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015, 1293-1304.
[29] J. Cao, A.L. Truong, S. Banu, A.A. Shah, A. Sabharwal, N. Moukaddam, Tracking and Predicting Depressive Symptoms of Adolescents Using Smartphone-Based Self-Reports, Parental Evaluations, and Passive Phone Sensor Data: Development and Usability Study, JMIR Ment Health 7 (2020) e14045.
[30] P. Fraccaro, S. Lavery-Blackie, S.N. Van der Veer, N. Peek, Behavioural Phenotyping of Daily Activities Relevant to Social Functioning Based on Smartphone-Collected Geolocation Data, Stud. Health Technol. Inform. 264 (2019) 945–949.
[31] Y. Gao, A. Li, T. Zhu, X. Liu, X. Liu, How smartphone usage correlates with social anxiety and loneliness, PeerJ 4 (2016) e2197.
[32] J. Gong, Y. Huang, P.I. Chow, K. Fua, M.S. Gerber, B.A. Teachman, L.E. Barnes, Understanding behavioral dynamics of social anxiety among college students through smartphone sensors, Inf. Fusion 49 (2019) 57–68.
[33] P. Henson, I. Barnett, M. Keshavan, J. Torous, Towards clinically actionable digital phenotyping targets in schizophrenia, npj Schizophrenia 6 (1) (2020) 13.
[34] J. He-Yueya, B. Buck, A. Campbell, T. Choudhury, J.M. Kane, D. Ben-Zeev, T. Althoff, Assessing the relationship between routine and schizophrenia symptoms with passively sensed measures of behavioral stability, npj Schizophrenia 6 (35) (2020) 35.
[35] Y. Huang, H. Xiong, K. Leach, Y. Zhang, P. Chow, K. Fua, B.A. Teachman, L.E. Barnes, Assessing social anxiety using GPS trajectories and point-of-interest data, in: ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2016, 898-903.
[36] M.T. Masud, M.A. Mamun, K. Thapa, D.H. Lee, M.D. Griffiths, S.H. Yang, Unobtrusive monitoring of behavior and movement patterns to detect clinical depression severity level via smartphone, J. Biomed. Inform. 103 (2020), 103371.
[37] M.T. Masud, N. Rahman, A. Alam, M.D. Griffiths, M. Alamin, Non-Pervasive Monitoring of Daily-Life Behavior to Access Depressive Symptom Severity Via Smartphone Technology, 2020 IEEE Region 10 Symposium (TENSYMP) (2020) 602-607.
[38] S. Rhim, U. Lee, K. Han, Tracking and modeling subjective well-being using smartphone-based digital phenotype, Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (2020) 211-220.
[39] S. Saeb, E.G. Lattie, K.P. Kording, D.C. Mohr, Mobile phone detection of semantic location and its relationship to depression and anxiety, JMIR mHealth and uHealth 5 (8) (2017) e7297.
[40] S. Saeb, E.G. Lattie, S.M. Schueller, K.P. Kording, D.C. Mohr, The relationship between mobile phone location sensor data and depressive symptom severity, PeerJ 4 (2016) e2537.
[41] S. Saeb, M. Zhang, C.J. Karr, S.M. Schueller, M.E. Corden, K.P. Kording, D.C. Mohr, Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study, J. Med. Internet Res. 17 (7) (2015) e4273.
[42] S. Saeb, M. Zhang, M. Kwasny, C.J. Karr, K. Kording, D.C. Mohr, The relationship between clinical, momentary, and sensor-based assessment of depression, in: 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), IEEE (2015) 229-232.
[43] A. Sarda, S. Munuswamy, S. Sarda, V. Subramanian, Using Passive Smartphone Sensing for Improved Risk Stratification of Patients With Depression and Diabetes: Cross-Sectional Observational Study, JMIR Mhealth Uhealth 7 (1) (2019) e11041.
[44] V.K. Singh, R. Goyal, S. Wu, Riskalyzer: Inferring Individual Risk-Taking Propensity Using Phone Metadata, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2 (2018) 1-21.
[45] S.S. Thakur, R.B. Roy, Predicting mental health using smart-phone usage and sensor data, J. Ambient Intell. Human Comput. 12 (2021) 9145–9161.
[46] C. Yue, S. Ware, R. Morillo, J. Lu, C. Shang, J. Bi, J. Kamath, A. Russell, A. Bamis, B. Wang, Automatic Depression Prediction Using Internet Traffic Characteristics on Smartphones, Smart Health 18 (2020), 100137.
[47] T.W. Boonstra, M.E. Larsen, H. Christensen, Mapping dynamic social networks in real life using participants' own smartphones, Heliyon 1 (3) (2015) e00037.
[48] T.W. Boonstra, A. Werner-Seidler, B. O'Dea, M.E. Larsen, H. Christensen, Smartphone app to investigate the relationship between social connectivity and mental health, Annu Int Conf IEEE Eng Med Biol Soc. (2017) 287-290.
[49] G. Pulekar, E. Agu, Autonomously sensing loneliness and its interactions with personality traits using smartphones, 2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT), IEEE (2016) 134-137.
[50] G.M. Harari, S.D. Gosling, R. Wang, F. Chen, Z. Chen, A.T. Campbell, Patterns of behavior change in students over an academic term: A preliminary study of activity and sociability behaviors using smartphone sensing methods, Comput. Hum. Behav. 67 (2017) 129–138.
[51] A.S. Dissing, N. Hulvej Rod, T.A. Gerds, R. Lund, Smartphone interactions and mental well-being in young adults: A longitudinal study based on objective high-resolution smartphone data, Scand. J. Public Health. 49 (2020) 325–332.
[52] E.M. Messner, R. Sariyska, B. Mayer, C. Montag, C. Kannen, A. Schwerdtfeger, H. Baumeister, Insights - Future Implications of Passive Smartphone Sensing in the Therapeutic Context, Verhaltenstherapie (2019) 1–10.
[53] Y. Huang, J. Gong, M. Rucker, P. Chow, K. Fua, M.S. Gerber, B. Teachman, L.E. Barnes, Discovery of behavioral markers of social anxiety from smartphone sensor data, 1st ACM Workshop on Digital Biomarkers (2017) 9-14.
[54] S. King, M. Ebraheem, K. Zanna, T. Neal, Learning a Privacy-Preserving Global Feature Set for Mood Classification Using Smartphone Activity and Sensor Data, 15th IEEE International Conference on Automatic Face and Gesture Recognition (2020) 582-586.
[55] K.A. Ryan, P. Babu, R. Easter, E. Saunders, A.J. Lee, P. Klasnja, L. Verchinina, V. Micol, B. Doil, M.G. McInnis, A.M. Kilbourne, A smartphone app to monitor mood symptoms in bipolar disorder: Development and usability study, JMIR Mental Health 7 (9) (2020) e19476.
[56] S. Stanislaus, M. Vinberg, S. Melbye, M. Frost, J. Busk, J.E. Bardram, L.V. Kessing, M. Faurholt-Jepsen, Smartphone-based activity measurements in patients with newly diagnosed bipolar disorder, unaffected relatives and control individuals, Int. J. Bipolar Disord. 8 (2020) 1–14.

[57] P. Henson, E. Rodriguez-Villa, J. Torous, Investigating Associations Between Screen Time and Symptomatology in Individuals With Serious Mental Illness: Longitudinal Observational Study, J Med Internet Res. 23 (2021) e23144.
[58] R. Wang, W. Wang, M.S. Aung, D. Ben-Zeev, R. Brian, A. Campbell, T. Choudhury, M. Hauser, J. Kane, E. Scherer, M. Walsh, Predicting symptom trajectories of schizophrenia using mobile sensing, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1(3) (2017) 1-24.
[59] O. Demasi, S. Feygin, A. Dembo, A. Aguilera, B. Recht, Well-being tracking via smartphone-measured activity and sleep: Cohort study, JMIR mHealth and uHealth 5 (10) (2017) e7820.
[60] D. Kelly, B. Caulfield, K. Curran, Crowdsourced data collection of physical activity and health status, in: An App Solution, International Conference on Wireless Mobile Communication and Healthcare, Springer, 2016, pp. 151–159.
[61] P. Staples, J. Torous, I. Barnett, K. Carlson, L. Sandoval, M. Keshavan, J.P. Onnela, A comparison of passive and active estimates of sleep in a cohort with schizophrenia, npj Schizophrenia 3 (1) (2017) 37.
[62] C.M. Bishop, N.M. Nasrabadi, Pattern recognition and machine learning, Technometrics 49 (3) (2007) 366.
[63] P.L. Haynes, D. Gengler, M. Kelly, Social rhythm therapies for mood disorders: an update, Curr. Psychiatry Rep. 18 (2016) 1–8.
[64] G.M. Harari, N.D. Lane, R. Wang, B.S. Crosier, A.T. Campbell, S.D. Gosling, Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges, Perspect Psychol Sci. 11 (2016) 838–854.
[65] C. Tossell, P. Kortum, C. Shepard, A. Rahmati, L. Zhong, Exploring smartphone addiction: Insights from long-term telemetric behavioral measures, Int. J. Interactive Mobile Technologies 9 (2015) 37–43.
[66] J. Torous, P. Staples, J.P. Onnela, Realizing the potential of mobile mental health: new methods for new data in psychiatry, Curr Psychiatry Rep. 17 (2015) 1–7.
[67] P. Kokoszka, M. Reimherr, Introduction to functional data analysis, Chapman and Hall/CRC, 2017.
[68] S.T. Pew Research Center, Demographics of mobile device ownership and adoption in the United States, 2022.
[69] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, A.T. Campbell, Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones, in: ACM International Joint Conference on Pervasive and Ubiquitous Computing (2014) 3-14.
[70] N. English, C. Zhao, K.L. Brown, C. Catlett, K. Cagney, Making Sense of Sensor Data: How Local Environmental Conditions Add Value to Social Science Research, Soc. Sci. Comput. Rev. 40 (1) (2022) 179–194.
[71] C.E. Shannon, The mathematical theory of communication. 1963, MD computing: computers in medical practice 14(4) (1997) 306-317.
[72] R.M. Merchant, D.A. Asch, P. Crutchley, L.H. Ungar, S.C. Guntuku, J.C. Eichstaedt, S. Hill, K. Padrez, R.J. Smith, H.A. Schwartz, Evaluating the predictability of medical conditions from social media posts, PloS One 14 (6) (2019) e0215476.
[73] B. Mahesh, Machine learning algorithms-a review, Int. j. sci. res. 9 (2020) 381–386.