International Journal of Scientific Interdisciplinary Research

Volume: *5;* ***Issue:*** *2*
Pages: *487–525*

open access

Crossref

PEER-REVIEWED

COMPUTATIONAL PSYCHOMETRICS AND DIGITAL BIOMARKER MODELING FOR PRECISION MENTAL HEALTH DIAGNOSTICS

Md. Akbar Hossain[1];

[1]. Master of Science in Clinical Psychology, University of Dhaka, Dhaka, Bangladesh;
Email: md.akbarh@yahoo.com

Doi: 10.63125/vg522x27
Received: *24 April 2024;* ***Revised:*** *28 May 2024;* ***Accepted:*** *09 June 2024;* ***Published:*** *29 June 2024*

Abstract

This study addressed the problem that mental health screening based mainly on self-report may miss ecologically valid behavioral signals, limiting precision diagnostics in real-world cloud and enterprise deployments; therefore, the purpose was to test whether integrating computational psychometrics with engineered digital biomarkers improves prediction of a continuous Precision Diagnostic Index (PDI) in a quantitative, cross-sectional, case-based design. The sample comprised 220 cloud/enterprise cases (N = 220; 56.4% female; age M = 26.8, SD = 6.1) who completed a Likert 5-point instrument and provided sufficient passive data for biomarker computation. Key psychometric variables were Affective Distress (M = 3.62, α = .88), Cognitive Dyscontrol (M = 3.31, α = .84), Functional Impairment (M = 3.45, α = .86), and Self-Regulation Capacity (reverse-coded; M = 2.74, α = .81). Digital biomarkers were z-scored Routine Irregularity, Sleep Disruption Proxy, Mobility Constriction, and Interaction Volatility, and the outcome PDI was standardized (M = 0.12, SD = 0.94). The analysis plan used descriptives and reliability checks, Pearson correlations, and hierarchical multiple regression comparing psychometrics-only, biomarkers-only, and integrated models. Headline findings showed strong psychometric alignment with PDI (Affective Distress r = .62; Functional Impairment r = .58; Cognitive Dyscontrol r = .49; Self-Regulation r = −.46; all p < .001) and moderate biomarker associations, strongest for Routine Irregularity (r = .38, p < .001) and Sleep Disruption (r = .32, p < .001). In regression, psychometrics explained R^2 = .49 of PDI variance (β = .41 distress, β = .29 impairment, β = .18 dyscontrol, β = −.14 self-regulation), biomarkers explained R^2 = .22 (β = .24 routine irregularity, β = .19 sleep disruption), and the integrated model increased prediction to R^2 = .56 with ΔR^2 = .07 (ΔF p < .001); routine irregularity (β = .16, p = .001) and sleep disruption (β = .12, p = .018) remained significant after controls, with acceptable multicollinearity (VIF 1.18–2.36). These results imply that enterprise screening systems can retain interpretable psychometric anchors while adding a small, theory-guided subset of passive biomarkers that delivers incremental validity and supports more precise triage and better resource allocation decisions across services.

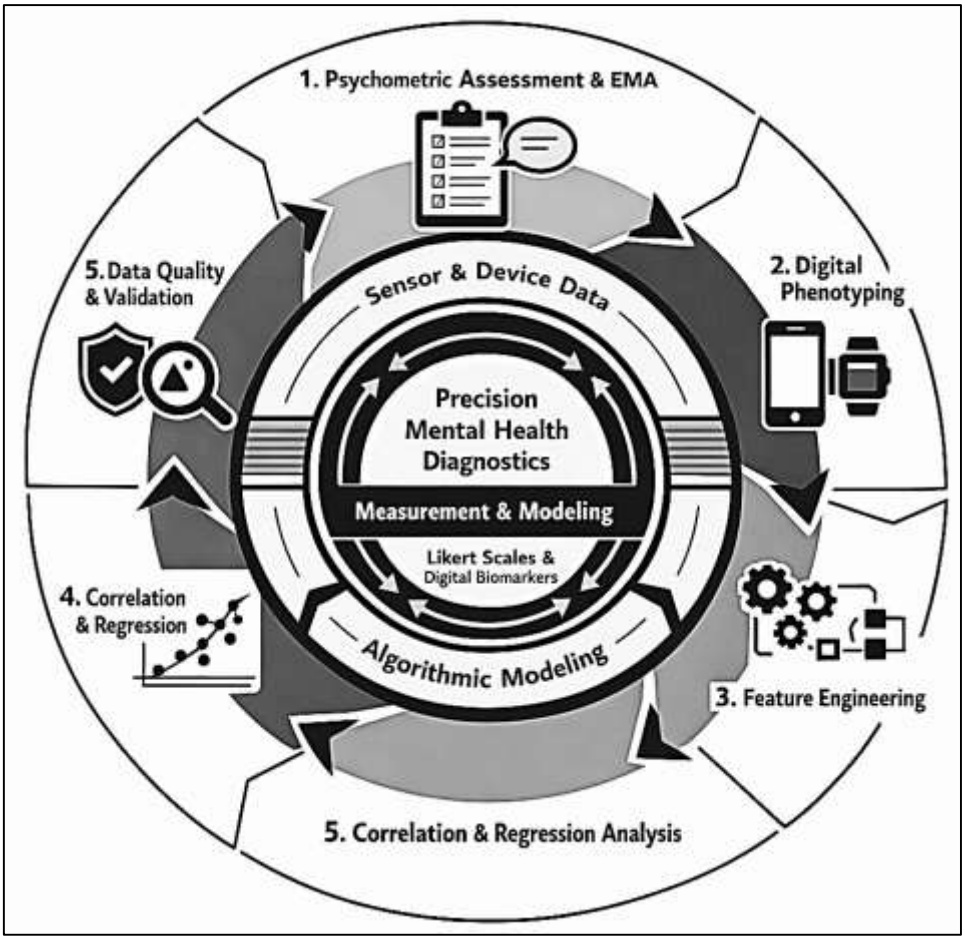
Keywords

Computational Psychometrics; Digital Biomarkers; Precision Diagnostic Index; Passive Smartphone Sensing; Incremental Validity;

INTRODUCTION

Computational psychometrics and digital biomarker modeling sit at the intersection of measurement science, behavioral data streams, and clinical decision-making in mental health. *Psychometrics* is the quantitative science of measuring latent psychological constructs (e.g., depression severity, cognitive control, social withdrawal) using observable indicators, traditionally questionnaire items scored on ordinal scales such as Likert response formats. *Computational psychometrics* extends this foundation by integrating modern computational methods (e.g., algorithmic scoring, data-driven feature engineering, machine learning, and probabilistic modeling) with psychometric principles so that measurement remains reliable, valid, and interpretable when indicators expand beyond survey items into complex, high-frequency digital traces (Davier & et al., 2017). In parallel, *biomarkers* are formally treated as measurable indicators used to characterize biological or clinical states, and contemporary biomarker frameworks emphasize discovery, analytic validation, clinical validation, and fit-for-purpose use in health contexts (Kraus & et al., 2018).

Figure 1: Systems Framework Linking Psychometric Measurement and Digital Biomarker Modeling



When the measurable indicator originates from sensors, software interactions, or connected devices rather than a wet-lab assay, it is commonly treated as a *digital biomarker*—a computationally derived marker that converts passively or actively collected signals (e.g., mobility, sleep proxies, typing behavior, speech features, heart rate patterns) into clinically relevant variables (Coravos et al., 2019). The international significance of these developments is grounded in the global burden of mental and substance use disorders, documented in large-scale epidemiological synthesis that quantifies disability and health-system load across regions and income settings (Whiteford et al., 2013). In many real-world systems, diagnostic processes still depend heavily on episodic clinic visits and retrospective recall, creating measurement gaps when symptoms fluctuate rapidly across days and contexts. Digital data

streams and computational psychometrics respond to this measurement gap by redefining what counts as an indicator and by demanding stronger evidence that new indicators support dependable inference about mental health states.

A central concept enabling digital biomarkers in psychiatry is *digital phenotyping*, often defined as the moment-by-moment quantification of individual-level behavior in situ through personal devices and connected sensors. Digital phenotyping treats everyday interactions with smartphones and wearables—communication rhythms, mobility regularity, device usage micro-patterns, sleep–wake proxies, and social proximity cues—as behavioral signals that can be summarized into features aligned with mental health constructs (Onnela & Rauch, 2016). This approach is strengthened when paired with ecological momentary assessment (EMA), which captures repeated, near-real-time self-reports in natural environments, reducing reliance on global retrospective summaries and producing temporally granular labels that can anchor modeling (Shiffman et al., 2008). Modern mental health measurement therefore increasingly spans (a) subjective self-report collected frequently, (b) clinician-rated outcomes collected periodically, and (c) passive sensor traces collected continuously. Methodological guidance emphasizes that smartphones can serve as scalable behavioral observation tools, but study designs must confront heterogeneity in device hardware, operating systems, user adherence, context effects, and missingness patterns (Harari et al., 2016; Mohiul, 2020). Within this landscape, "digital biomarkers" are not raw data; they are engineered variables produced through explicit computational pipelines (signal capture → preprocessing → feature extraction → model-based estimation) that must be evaluated for stability, construct correspondence, and clinical meaning (Jinnat & Kamrul, 2021; Hasan & Shaikat, 2021). A clinically oriented digital biomarker framework also requires that the derived variables remain interpretable enough to support diagnostic reasoning, not merely prediction. Digital phenotyping work in psychiatry positions these data streams as candidate measures for symptom dynamics, relapse risk, functional impairment, and cognitive performance, while keeping the scientific center of gravity on measurement quality rather than device novelty (Rabiul & Samia, 2021; Mohiul & Rahman, 2021; Orru et al., 2019).

Psychometrics becomes especially consequential when digital mental health research depends on short scales, repeated assessments, and heterogeneous populations. Likert-type instruments remain widely used because they translate subjective experiences into structured indicators that are amenable to descriptive statistics, correlation analysis, and regression modeling in cross-sectional and case-study–based quantitative designs (Rahman & Abdul, 2021; Haider & Shahrin, 2021). Yet classical fixed-form questionnaires carry practical constraints: they may impose burden, produce floor/ceiling effects, and offer uneven precision across severity ranges. Item response theory (IRT) and item banking address these issues by modeling item characteristics and enabling comparable scores even when different items are administered. PROMIS advances this paradigm by building calibrated item banks for health domains and enabling tailored short forms and computerized adaptive testing (Cella et al., 2007; Habibullah & Farabe, 2022; Zulqarnain & Subrato, 2021). In mental health constructs such as depression, PROMIS-based depression item banks demonstrate convergent validity and responsiveness in clinical observational contexts while reducing item burden through adaptive administration (Arman & Kamrul, 2022; Rashid & Sai Praveen, 2022; Pilkonis et al., 2014). Computerized adaptive testing also enters psychiatry through disorder-focused instruments such as the Computerized Adaptive Test–Depression Inventory (CAT-DI), which shows how measurement precision can be "fixed" while the number of administered items varies, producing high correlations with longer forms and supporting accurate classification in clinical samples (Gibbons et al., 2012; Kamrul & Omar, 2022; Rahman, 2022). This psychometric infrastructure matters for computational psychometrics because it provides principled latent-variable targets and uncertainty estimates that can be integrated into statistical learning and regression-based inference (Abdul & Rahman, 2023; Rony & Samia, 2022). When Likert-scale constructs are paired with digital biomarkers, the measurement problem becomes dual: the quality of the psychometric instrument and the validity of the engineered digital indicators both shape diagnostic accuracy and explanatory strength (Aditya & Rony, 2023; Arfan & Rony, 2023). Computational psychometrics therefore frames digital mental health diagnostics as a measurement system, not only a prediction system, requiring evidence that constructs are measured

with stability, reliability, and defensible links to observed indicators across contexts (Ara & Shaikh, 2023; Habibullah & Mohiul, 2023).

Digital biomarker modeling in precision mental health frequently uses passive sensing to capture behavioral regularities associated with symptom severity. Empirical studies demonstrate that smartphone-derived features can correlate with depressive symptom severity, translating mobility, phone usage patterns, and daily routines into measurable behavioral markers that align with clinical self-report outcomes (Hasan & Waladur, 2023; Saeb et al., 2015). Similar sensing approaches are applied in affective disorders, including bipolar disorder monitoring systems where smartphone-based self-monitoring and objective behavioral traces are examined alongside clinically rated symptoms, supporting the broader idea that everyday behavioral rhythms can be operationalized into quantitative indicators (Faurholt-Jepsen & et al., 2014; Arman & Nahid, 2023; Mesbaul, 2023). Digital biomarkers also extend into cognitive domains: passively acquired smartphone interaction data can produce computational markers that correlate with gold-standard neurocognitive assessments, illustrating that everyday device interaction can be mined for cognitive-function indicators under defined modeling assumptions (Dagum, 2018; Milon & Mominul, 2023; Mohaiminul & Muzahidul, 2023). Beyond single-feature correlations, predictive modeling studies use multi-modal data (phone + wearable) and feature-selection pipelines to identify patterns associated with depression presence and symptom change, showing that regression-compatible features can be extracted from longitudinal behavioral streams and related to psychometric outcomes (Doryab & et al., 2022; Musfiqur & Kamrul, 2023; Rezaul & Kamrul, 2023). At the systems level, "personal sensing" research synthesizes how ubiquitous sensors and machine learning can support mental health measurement and monitoring, while emphasizing that interpretability, missingness, and construct validity remain central to clinical usefulness (Amin & Sai Praveen, 2023; Rabiul & Mushfequr, 2023; Torous & et al., 2018). In this context, correlation analysis and regression modeling serve two complementary roles: (a) testing theory-consistent relationships between engineered biomarkers and psychometric constructs, and (b) estimating additive or multivariable contributions of digital features while controlling for demographics, case context, or baseline symptom levels. Computational psychometrics brings additional structure by treating digital features as candidate items or indicators, evaluating their information value, redundancy, and stability in relation to latent constructs measured with Likert-based instruments.

Precision mental health diagnostics also draws from "biotype" and neurobiological stratification research that seeks to refine heterogeneous diagnostic categories into more homogeneous subgroups with distinct profiles. Work on resting-state connectivity biomarkers proposes reproducible subtypes of depression based on neurophysiological patterns, illustrating a biomarker logic where latent clinical heterogeneity is modeled through objective indicators (Shahrin & Samia, 2023; Roy, 2023; Wang & et al., 2014). Similarly, psychosis research identifies distinct biotypes using brain-based biomarkers, reinforcing the principle that diagnostic categories can be decomposed into biologically and behaviorally meaningful subgroups (Clementz et al., 2016; Rakibul & Majumder, 2023; Rifat & Rebeka, 2023). A precision psychiatry framework further emphasizes that depression and anxiety may be mapped to neural circuit taxonomies, positioning diagnosis as a problem of measurable signatures across levels of analysis rather than solely symptom checklists (Kumar, 2023; Saikat & Aditya, 2023; Williams, 2016). Digital biomarkers complement these approaches by offering high-frequency behavioral data that can be aligned with symptom measures and, in some studies, combined with neuroimaging to explore convergent signatures of depression-related functioning (e.g., smartphone sensing combined with brain imaging) (Zaki & Masud, 2023; Zaki & Hossain, 2023; Wang & et al., 2019).

In measurement terms, the integration challenge is to define coherent construct models where psychometric indicators, digital behavior features, and (where relevant) biological markers participate as complementary evidence streams. Computational psychometrics contributes by clarifying how indicators relate to constructs, how uncertainty is quantified, and how measurement invariance and bias risks are evaluated when indicators originate from different modalities. This is particularly relevant in cross-sectional, case-study–based quantitative designs where the goal is to test hypotheses about relationships among constructs and biomarkers using descriptive statistics, correlations, and regression models anchored in Likert-scale measures. Digital biomarker modeling thus becomes one

component of a broader precision diagnostic logic: structuring multiple indicator sources into a measurement system that supports consistent classification and construct-level inference.

The credibility of computational psychometrics and digital biomarker modeling depends on explicit attention to data quality, validation, and fit-for-purpose evidence. Digital phenotyping studies show that missingness, sensor dropout, and inconsistent sampling can reshape conclusions if not quantified and modeled, making data quality characterization a core scientific task rather than a technical afterthought (Rashid, 2024; Torous et al., 2017; Zulqarnain & Subrato, 2023). Design guidance also highlights that sample size and follow-up duration influence the stability of digital phenotyping estimates because behavioral features may require sufficient temporal coverage to represent typical patterns and reduce noise (Barnett et al., 2020; Md & Praveen, 2024; Mohaiminul & AMajumder, 2024). At the biomarker level, contemporary frameworks in medicine emphasize staged evaluation—discovery through clinical qualification—and digital biomarkers inherit these demands, including the need to demonstrate that algorithmic outputs perform consistently across populations and that measurement pipelines are transparent and auditable (Drysdale et al., 2017; Foysal & Abdulla, 2024; Ibne & Aditya, 2024). In psychometrics, reliability and validity are not optional add-ons; they define whether a score supports inference. PROMIS item bank development and validation studies illustrate how reliability, convergent validity, and responsiveness are established under modern psychometric methods (Chikersal et al., 2020; Milon & Mominul, 2024; Mosheur & Arman, 2024). Likewise, adaptive testing work in depression measurement demonstrates how precision can be quantified via standard errors and how classification performance can be evaluated against clinical criteria (Pilkonis et al., 2011; Rahman & Aditya, 2024; Saba & Hasan, 2024). These principles extend naturally to digital biomarkers: engineered features must be evaluated for stability, measurement error, construct overlap, and potential confounding by device usage style, socioeconomic factors, or contextual constraints. Methodological syntheses in personal sensing also emphasize privacy, ethics, and the risk of overfitting when models are tuned on limited or homogeneous samples (Mohr et al., 2017; Kumar, 2024; Sai Praveen, 2024). In computational psychometrics, the methodological throughline is consistent: diagnostic claims require rigorous measurement arguments, and those arguments must hold when indicators shift from questionnaire items to digital traces derived from complex computational pipelines.

Within this scientific context, the research title "Computational Psychometrics and Digital Biomarker Modeling for Precision Mental Health Diagnostics" positions diagnosis as a measurement and modeling problem that is addressed using quantitative cross-sectional evidence in a case-study setting. The study's measurement foundation is compatible with Likert five-point instruments that operationalize core constructs relevant to mental health diagnostics, while its computational component operationalizes digital biomarker features derived from device- or platform-based signals through defined feature engineering pipelines (Canzian & Musolesi, 2015; Saikat, 2024; Shaikat & Aditya, 2024). The analytic strategy aligns naturally with descriptive statistics to characterize constructs and biomarker distributions, correlation analysis to test associative structure among psychometric constructs and engineered biomarker features, and regression modeling to evaluate multivariable relationships consistent with hypothesis testing (Pilkonis et al., 2011). A computational psychometric perspective encourages explicit specification of constructs, indicators, scoring rules, and uncertainty quantification, drawing from modern psychometric systems such as PROMIS item banking and adaptive testing to justify measurement decisions and reduce instrument burden while maintaining precision (Wang & et al., 2014). Digital biomarker modeling further requires explicit operational definitions of features—such as mobility regularity, circadian stability proxies, interaction intensity, or cognition-related interaction markers—together with defensible mapping logic linking those features to diagnostic constructs (Coravos et al., 2019). The rationale for the integrated approach is anchored in the observation that mental health symptoms often vary across time and context, and methods such as EMA and passive sensing provide complementary data structures that support more granular measurement when evaluated under rigorous psychometric and biomarker validation principles (Shiffman et al., 2008). The study's contribution is thus framed as a structured measurement-and-modeling design: psychometrically grounded constructs measured with Likert scales are examined

together with engineered digital biomarkers through correlation and regression, within a case context that makes the diagnostic setting and population characteristics explicit for interpretation and reproducibility.

This study is designed to achieve a clear set of objectives that operationalize the idea of precision mental health diagnostics through the combined use of computational psychometrics and digital biomarker modeling within a quantitative, cross-sectional, case-study–based design. The first objective is to define and operationalize the core psychometric constructs that represent clinically meaningful dimensions of mental health status in the selected case context, ensuring that each construct is measurable through a structured Likert five-point instrument with transparent scoring rules and interpretable indicators. The second objective is to identify, structure, and quantify a set of digital biomarker features that correspond to observable behavioral or physiological patterns relevant to mental health functioning, translating raw digital traces into analyzable variables through consistent preprocessing and feature extraction steps suitable for statistical testing. The third objective is to describe the sample and case setting comprehensively by producing participant-level demographic profiles and contextual descriptions that allow the results to be interpreted in relation to the specific environment in which the diagnostic modeling is evaluated. The fourth objective is to determine the statistical distribution, central tendency, and variability of both psychometric constructs and digital biomarker features using descriptive statistics, establishing a clear empirical baseline for subsequent inferential testing. The fifth objective is to examine the strength, direction, and statistical significance of relationships among the study variables by applying correlation analysis to psychometric constructs, digital biomarker features, and diagnostic or symptom outcomes, thereby clarifying which indicators demonstrate meaningful associations in the cross-sectional snapshot. The sixth objective is to evaluate predictive relationships by applying regression modeling to test how digital biomarker variables and computational psychometric constructs explain variation in mental health diagnostic outcomes, including the estimation of effect sizes and model fit indicators to determine which predictors contribute most strongly when examined jointly. The seventh objective is to compare alternative regression specifications by testing a measurement-only model, a biomarker-only model, and an integrated model, enabling a structured assessment of whether combining psychometric and digital indicators yields stronger explanatory performance than either component alone within the same dataset. The eighth objective is to assess the internal consistency and measurement adequacy of the Likert-based constructs through reliability evaluation and related psychometric checks, ensuring that the constructs used in hypothesis testing satisfy minimum quality thresholds for quantitative inference. Collectively, these objectives provide a coherent pathway from construct definition and indicator generation to empirical testing and model-based evaluation, allowing the study to examine the feasibility and statistical performance of an integrated measurement-and-modeling approach to precision mental health diagnostics in the defined case context.

LITERATURE REVIEW

The literature on computational psychometrics and digital biomarker modeling for precision mental health diagnostics converges around a shared objective: improving how mental health states are measured, inferred, and differentiated using rigorous quantitative methods and high-resolution behavioral data. Contemporary mental health assessment has historically relied on standardized psychometric instruments that transform subjective experiences into numeric indicators, allowing symptom patterns to be summarized and compared across individuals and settings. This tradition remains essential because it provides structured constructs, scoring logic, and evidence standards for reliability and validity. At the same time, the expansion of digital life has introduced new forms of observable behavior captured through smartphones, wearables, and connected platforms, enabling continuous or high-frequency measurement of activity patterns, sleep-related proxies, interaction rhythms, mobility regularity, and other signals that may reflect cognitive and affective functioning. Digital biomarkers emerge from this shift as engineered variables derived from raw sensor and interaction data through computational pipelines, and their value depends on whether they meaningfully correspond to clinically relevant constructs rather than merely producing statistically significant predictions. Precision mental health diagnostics builds on this foundation by seeking more individualized, context-sensitive, and data-driven diagnostic support, especially in conditions

characterized by heterogeneity and fluctuating symptom trajectories. Within this broader movement, computational psychometrics provides the methodological bridge that connects classical measurement theory with modern analytics by emphasizing construct clarity, error modeling, latent trait estimation, and interpretability when indicators expand beyond traditional questionnaire items. The literature therefore highlights two parallel demands: first, that psychometric measurement remains robust when deployed in real-world digital settings, and second, that digital biomarker modeling is evaluated with the same rigor expected of clinical measurement tools, including careful operationalization, transparent feature engineering, and empirical testing of relationships with symptom and diagnostic outcomes. Across studies, correlation analysis and regression modeling are commonly used to test associative structures and predictive contributions of biomarkers and psychometric constructs, supporting hypothesis-driven evaluation in cross-sectional and case-based quantitative designs. Taken together, the literature establishes a multidisciplinary evidence base that justifies integrating psychometrically grounded constructs with digital biomarker features to strengthen measurement resolution and diagnostic modeling within clearly defined clinical or community case contexts.

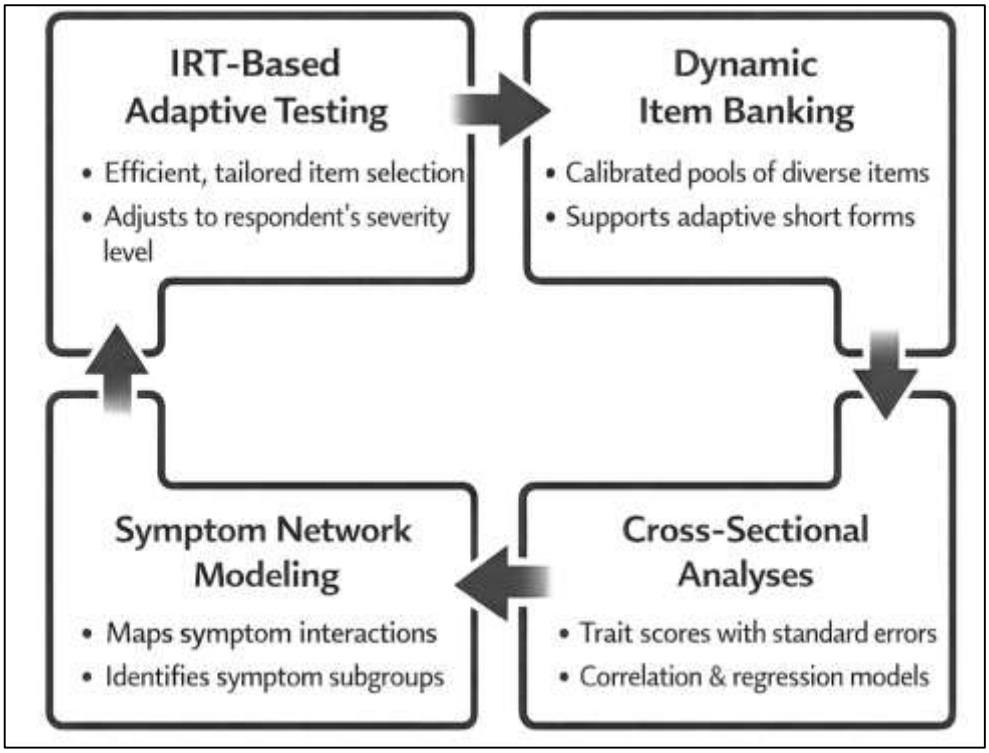
Computational Psychometrics in Mental Health Assessment

Computational psychometrics in mental health assessment denotes the integration of measurement theory with intensive computation for item calibration, score estimation, and decision rules that operate efficiently in clinical and community settings. The approach treats depression, anxiety, and related experiences as latent variables inferred from observable item responses while explicitly quantifying measurement error for each person. Item response theory (IRT) provides the statistical foundation: each item is characterized by parameters that describe how response probabilities vary with latent severity, allowing scores to be estimated even when individuals answer different sets of items. This foundation is especially valuable for Likert-type items because graded response or partial credit models accommodate ordered categories and yield information functions that show where an item measures best. Computerized adaptive testing (CAT) uses these models to select the next most informative item given a respondent's current estimate, reducing burden while preserving precision. Early clinical demonstrations illustrate the logic: a depression CAT can be built by extracting candidate items from common questionnaires, screening them for dimensionality and local dependence, calibrating parameters, and then using Bayesian scoring rules to stop once a target standard error is reached (Fliege et al., 2005).

Parallel work in anxiety measurement shows how an IRT-calibrated item bank and a CAT engine can reproduce full-bank trait estimates with only a small number of adaptively chosen items, enabling rapid assessment in routine care and supporting discrimination among diagnostic groups (Walter et al., 2007). Across such applications, computational psychometrics emphasizes that efficiency gains must be tied to transparent measurement claims: what construct is being measured, how precision varies across severity, and how score uncertainty should be carried forward into statistical analyses such as correlations and regression models within cross-sectional case designs. These properties help clinicians interpret score differences, avoid overconfidence, and align measurement outputs with diagnostic decision thresholds.

A second pillar of computational psychometrics is the construction and calibration of item banks that represent a construct across its severity continuum and across diverse groups. Item banking shifts assessment from a fixed questionnaire toward a calibrated pool, where each item contributes known information at different trait levels and can be adaptively sampled for efficiency. In mental health screening, this architecture is valuable because symptom presentations vary with comorbidity and care context, and fixed short forms often lose precision at important severity ranges. Adaptive screening studies demonstrate how recalibration and quality checks preserve comparability when an item bank is moved from development samples into applied settings. Recalibration work on a depression item bank shows how item parameters can be updated and then evaluated through simulated adaptive administrations, supporting brief screening with predefined standard errors while maintaining alignment with established severity measurement (Forkmann et al., 2013).

Figure 2: Square-Cycle Framework of Computational Psychometric Assessment in Mental Health



Such recalibration emphasizes routine psychometric diagnostics, including tests for dimensionality, local dependence, and item fit, because violations can distort adaptive selection and create biased estimates. Item banks also support population-specific development when language, culture, and clinical thresholds differ. A simulation study developing a depression screening item bank for a Chinese population illustrates how a locally validated pool can be created and then evaluated in CAT form, with evidence that a small number of adaptively chosen items can approximate full-bank scores while limiting respondent burden (Tan et al., 2018). In applied services, these item-bank systems matter for cross-sectional analytics because they produce trait estimates accompanied by standard errors, allowing researchers to interpret correlation patterns and regression coefficients in light of measurement precision. They also encourage transparent operational definitions of constructs, item coverage, and scoring rules, strengthening construct validity arguments in case-study designs that combine psychometric indicators with digital biomarker features. This supports fair comparisons across participants.

Computational psychometrics also broadens what counts as a measurement model in mental health by enabling structures that do not require symptoms to behave as interchangeable indicators of one hidden disorder. A prominent alternative is symptom-network modeling, where assessment focuses on the pattern of relations among symptoms rather than on a single latent severity score. Here, diagnostically useful descriptions include which symptoms cluster together and which symptoms connect clusters, because these relations can reflect how distress is organized within an individual or group. A network theory of mental disorders formalizes this view by explaining psychopathology as arising from direct interactions among symptoms and by motivating analytic methods that estimate and interpret symptom networks from observed responses (Borsboom, 2017). For computational psychometrics, this shifts instrument use toward identifying informative symptom configurations, such as response profiles that separate subgroups or indicate distinct pathways of symptom activation in a case context. The approach can complement IRT-based scoring because both traditions require clarity about what is inferred from item responses, while differing in whether covariance is treated as evidence of a common cause or as part of the phenomenon itself. In cross-sectional case studies, network-informed metrics can be paired with conventional scale scores to test convergent patterns: severity indices summarize overall distress, and network metrics summarize symptom organization

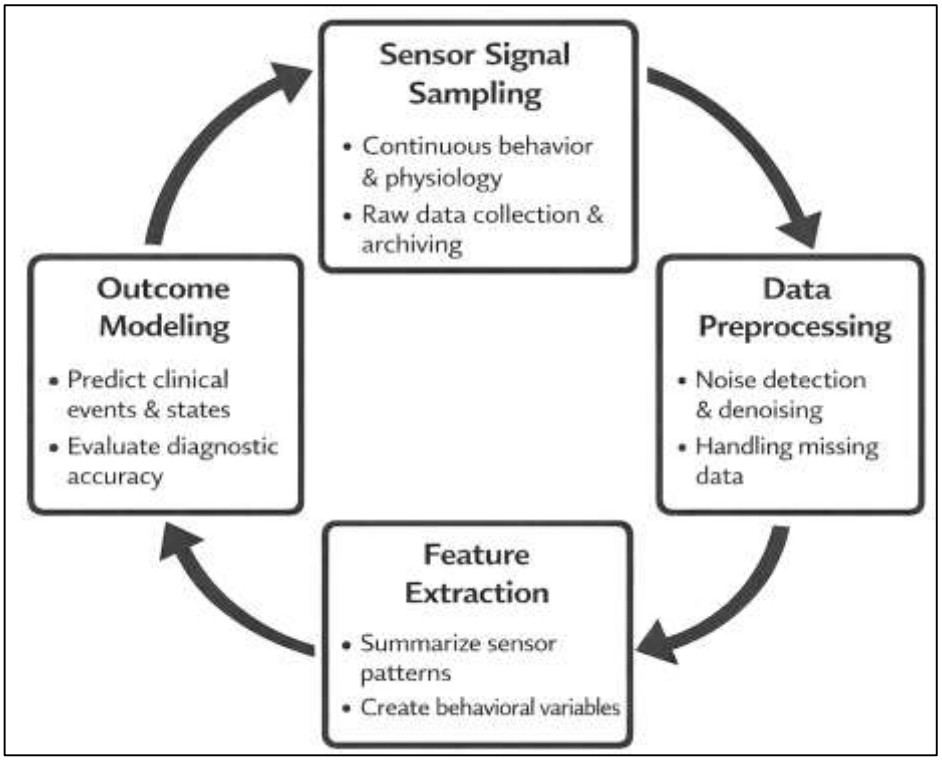
within the same dataset. Computational workflows support this pairing by providing reproducible pipelines for ordinal association estimation, model regularization, and stability checks, which helps prevent overinterpretation of fragile structures. When network outputs are later linked to external predictors such as engineered digital biomarker features, they provide an additional layer of measurement that can be examined with correlations and regression models. The broader agenda remains consistent: select models aligned to construct theory, keep outputs interpretable to clinicians, and report uncertainty so diagnostic claims remain proportionate to evidence quality.

Digital Biomarkers for Precision Mental Health Diagnostics

Digital biomarkers in mental health are measurable signals captured through digital devices that can serve as indicators of psychological state, symptom severity, or clinical risk. In precision mental health diagnostics, biomarkers are typically derived from passive sensing streams such as mobility, activity, and proximity, or from wearable physiology such as sleep continuity and autonomic variation, and then summarized as features that can be mapped onto psychiatric constructs. Methodologically, digital biomarker pipelines begin with raw sampling, proceed through preprocessing to address noise and missingness, and culminate in feature extraction that translates sensor traces into behaviorally meaningful variables like time at home, routine regularity, conversation exposure, or fragmented sleep. These variables are attractive because they are collected in naturalistic settings and can be aligned with psychometric measurement by treating the feature set as an additional source of indicators alongside questionnaire items. However, their interpretability depends on transparent definitions of what is measured, how often it is measured, and under what contextual constraints, because the same feature can reflect different psychological processes across individuals and settings. A further consideration is acceptability: a biomarker may be technically feasible yet impractical if users find the sensing burden intrusive or if privacy worries lead to selective participation. Work with people living with schizophrenia illustrates these constraints while demonstrating that multi-sensor behavioral capture is possible in both inpatient and outpatient contexts; participants generally tolerated device-embedded sensing and the resulting data streams captured activity, location patterns, and exposure to speech that are relevant for clinical characterization (Ben-Zeev et al., 2016). Device heterogeneity and operating-system policies alter sampling rates, sensor availability, and location accuracy, creating measurement error if not documented. Studies typically define minimum data-density thresholds, harmonize features across devices, and retain contextual metadata so that estimates are comparable. Such controls help treat sensor-derived features as reliable quantitative indicators.

For schizophrenia-spectrum conditions, social withdrawal and disruption of routine often precede symptomatic worsening, yet conventional detection relies on infrequent clinical visits and retrospective recall. Smartphone-derived indicators can operationalize social behavior continuously by summarizing outgoing and incoming calls, text messaging, and conversational exposure, producing ecologically valid proxies for engagement. In a year-long monitoring study using the CrossCheck sensing system, reductions in outgoing calls and texting were associated with relapse-related events, suggesting that digitally captured social activity can flag clinically relevant change before it becomes visible in scheduled assessments (Buck et al., 2019). Importantly, this kind of association is not only about detecting group averages; it also highlights that the most informative signals may be individualized, because relapse signatures can differ across patients depending on baseline functioning, support networks, and medication adherence. To address that heterogeneity, modeling strategies increasingly treat relapse detection as an anomaly-identification problem in which each individual's stable pattern is learned first and deviations are then evaluated as potential risk periods. Encoder–decoder approaches applied to passive smartphone features have shown how the same sensor set can be transformed into patient-specific alerts by learning "days of relative health" and then flagging behavioral anomalies that cluster in a prespecified window before relapse (Adler et al., 2020). From a computational psychometrics perspective, these designs align with the idea that indicators vary in informativeness over time and across persons, so a digital biomarker system should quantify both deviation magnitude and confidence under missing data. Consequently, relapse-oriented biomarker work provides an empirical template for your case-study approach: define clinically anchored events, extract interpretable features, and evaluate prediction models with clear sensitivity–specificity tradeoffs. This framing supports hypothesis testing in regression models.

Figure 3: Digital Biomarker Signal-to-Outcome Modeling Cycle for Mental Health Diagnostics



Digital biomarker modeling is equally influential for common conditions where symptoms fluctuate and are often underdetected, especially when the aim is to estimate severity continuously rather than rely on single time-point screening. Multimodal sensing studies show how behavioral and physiological features can complement Likert-based symptom scales by capturing correlates of mood, arousal, and impairment in everyday contexts. Smartphone sensors can quantify mobility diversity, location regularity, and device-use rhythms, while wearables can add sleep continuity, activity, and autonomic markers that reflect stress regulation. A study combining iPhone sensing with an Oura Ring demonstrated that location variance and sleep-related measures were associated with depression and anxiety symptom scores, and that integrating passive data streams with brief self-reported mood improved predictive performance compared with single-source models (Moshe et al., 2021). This illustrates a core design principle for precision diagnostics: different sensors index different parts of the latent construct, so fusing features can reduce measurement error and better represent multidimensional symptom profiles. Digital biomarkers can also be extracted from communication signals, where language content, prosody, and conversational context offer a window into cognition and affect without requiring a structured clinical interview. An observational study that passively detected words from sparsely sampled environmental audio found that category-level linguistic counts correlated with self-reported depression and anxiety severity, indicating that ambient speech can yield quantifiable indicators linked to affective symptomatology (Di Matteo et al., 2021). Taken together, these results motivate treating digital biomarkers as psychometric indicators: they can be scaled, correlated, and entered into regression models alongside survey constructs to test whether passive features add incremental validity beyond traditional measures. Your quantitative case-study framing can therefore compare competing feature sets—behavioral, physiological, and linguistic—while preserving interpretability by tying each feature to a plausible clinical mechanism. Cross-sectional designs use synchronized sensing windows for feature extraction aligned with questionnaire periods.

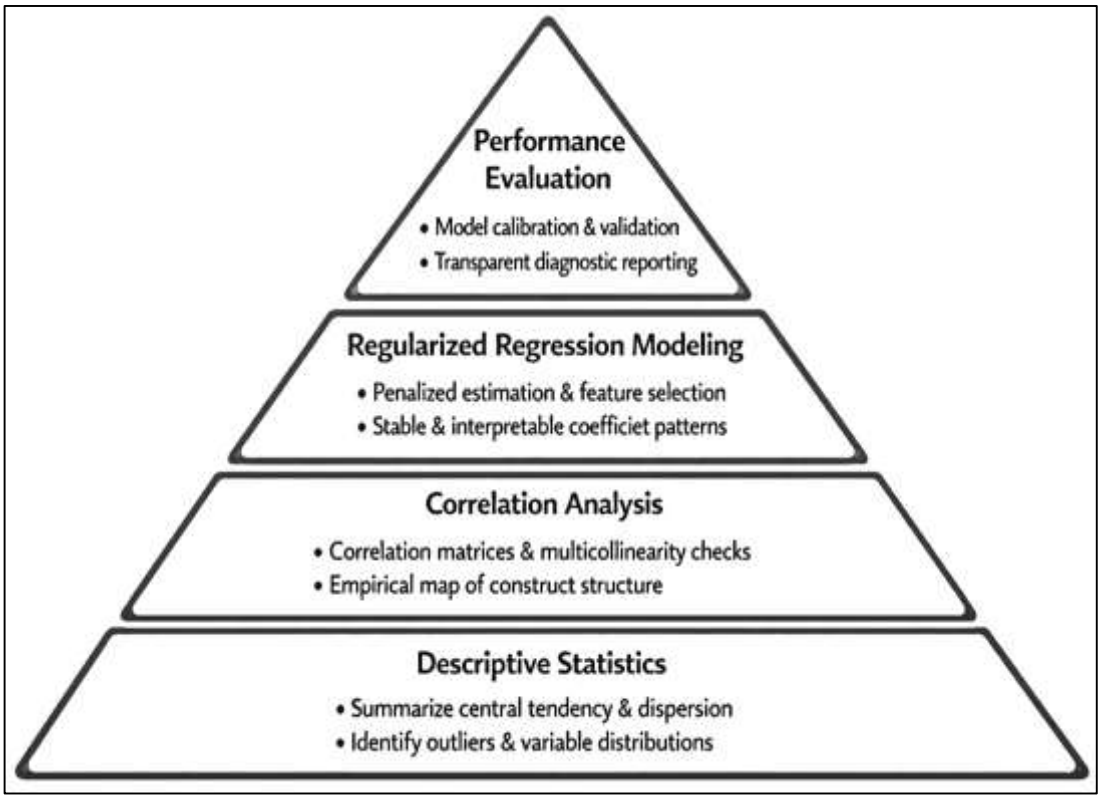
Data-Driven Diagnostic Modeling Approaches

Data-driven diagnostic modeling in precision mental health translates multi-source indicators into quantitative estimates of symptom severity, risk status, or diagnostic class while retaining interpretability for clinical reasoning. In quantitative cross-sectional designs, this process often begins with descriptive statistics to summarize central tendency and dispersion, followed by correlation

analysis to examine linear associations among Likert-scale constructs and engineered digital biomarker features. Correlation matrices help identify redundant predictors, clarify expected directions among variables, and provide an initial empirical map of the construct structure before multivariable modeling is introduced. Regression modeling then formalizes these relationships by estimating the unique contribution of each predictor while controlling for covariates such as age, gender, or context-specific factors embedded in the case-study setting. When outcomes are continuous (e.g., symptom severity indices), multiple linear regression supports effect-size interpretation through standardized coefficients; when outcomes are categorical (e.g., risk group, diagnostic class), logistic regression supports probability-based inference. The core challenge in this setting is that digital biomarkers often produce many correlated predictors relative to sample size, increasing the risk of unstable coefficients and inflated model optimism. Regularization provides a principled strategy to stabilize estimation and improve generalization in high-dimensional predictor spaces by shrinking coefficients and enabling feature selection, especially when predictors are correlated and may represent overlapping behavioral processes. Elastic net regularization is widely used because it blends ridge and lasso penalties, encouraging grouped selection among correlated predictors while still permitting sparsity in the final model, which is well aligned with digital biomarker feature sets that naturally co-vary (Zou & Hastie, 2005). This approach supports regression-based hypothesis testing by providing stable coefficient patterns and reducing sensitivity to small perturbations in the dataset.

Model building in precision mental health must also be operationally feasible, reproducible, and consistent with the measurement properties of inputs. Likert-scale constructs are ordinal indicators that are often treated as approximately continuous at scale level, yet modeling should remain attentive to distributional properties and to multicollinearity among constructs.

Figure 4: Structured Diagnostic Modeling Pipeline with Performance Evaluation



Digital biomarker features require preprocessing decisions—aggregation windows, missingness thresholds, and normalization—that directly affect correlation strength and regression estimates, so the modeling pipeline must be explicitly documented. Regularized generalized linear models make it possible to integrate psychometric constructs and biomarker features within one coherent regression framework while controlling complexity. Practical adoption accelerated with efficient coordinate-

descent algorithms that compute full regularization paths rapidly, enabling researchers to explore families of models and choose tuning parameters using cross-validation without excessive computational burden. The glmnet approach is frequently used because it supports linear and logistic regression with lasso, ridge, and elastic net penalties under a unified optimization framework, facilitating transparent comparisons between unpenalized regressions and shrinkage-based alternatives in the same dataset (Friedman et al., 2010). In cross-sectional case-study research, this makes it straightforward to test nested specifications such as (a) psychometric-only models, (b) biomarker-only models, and (c) integrated models, while using a common validation strategy to avoid selecting a model purely on in-sample fit. Even when the analysis emphasizes correlation and regression rather than complex machine learning, these tools support better control of model flexibility, strengthen interpretability, and reduce the likelihood that apparent relationships are artifacts of collinearity or overfitting. The result is a modeling approach that remains compatible with classical inference reporting (coefficients, confidence intervals, p-values) while improving stability and out-of-sample credibility.

Performance evaluation and transparent reporting are essential because diagnostic models can appear strong in a single dataset yet fail when applied outside the original setting due to sampling differences, predictor measurement shifts, or outcome definition changes. Discrimination is commonly assessed using receiver operating characteristic analysis for binary outcomes, which summarizes how sensitivity and specificity change across thresholds and provides a threshold-free view of separability between outcome groups (Fawcett, 2006). However, discrimination alone is insufficient for diagnostic decision support, because well-separated predictions may still be poorly calibrated, meaning predicted probabilities do not align with observed outcome frequencies. For this reason, model evaluation typically pairs discrimination summaries with calibration checks and clear reporting of how thresholds are selected for classification in applied contexts. Reporting standards provide structure for this transparency: the TRIPOD statement specifies key items that should be documented in prediction model studies, including participant selection, predictor definitions, outcome ascertainment, handling of missing data, model specification, and the reporting of performance measures, which improves reproducibility and interpretability across readers and reviewers (Collins et al., 2015). Beyond reporting, quality appraisal frameworks help evaluate whether a model is likely to be biased or inappropriately generalized; PROBAST offers domain-based signaling questions that focus attention on participant flow, predictor handling, outcome definition, and analytic choices that can introduce bias or inflate apparent accuracy (Moons et al., 2019). Together, this evaluation and reporting principles anchor correlation-and-regression–based diagnostic modeling in rigorous evidence standard suitable for precision mental health studies that integrate psychometric constructs with digital biomarker features.

RDoC for Precision Diagnostics

A suitable theoretical foundation for computational psychometrics and digital biomarker modeling in precision mental health diagnostics is built by combining (i) dimensional, mechanism-oriented classification logic and (ii) person-specific measurement assumptions. The Research Domain Criteria (RDoC) framework formalizes the first pillar by positioning mental health research around *functional domains* (e.g., negative valence, cognitive systems) rather than symptom checklists alone, and by encouraging integration of multiple "units of analysis" (self-report, behavior, physiology, neural systems) within a single measurement logic (Insel et al., 2010). This theoretical stance is directly aligned with digital biomarkers because passive sensing features and wearable physiology can be treated as behavioral/physiological units of analysis that complement Likert-scale self-report constructs in a single model of inference. The RDoC program also emphasizes that diagnostic understanding improves when the same construct is measured through convergent indicators, enabling triangulation between what a person reports and what their daily patterns suggest. In this study's context, that stance supports defining constructs (e.g., affective dysregulation, stress reactivity, behavioral activation) and then operationalizing them using both psychometric indicators and engineered digital features. A second pillar is articulated by the person-specific paradigm, which argues that many psychological processes are *nonergodic*—meaning group-level relationships may not generalize to individual-level dynamics—and therefore measurement and modeling should respect intraindividual variability and

individualized patterns when possible (Molenaar & Campbell, 2009). Precision diagnostics inherits this logic because digital biomarkers often reflect stable personal baselines (routine, mobility radius, sleep regularity) and clinically relevant deviations from that baseline. When RDoC dimensional constructs are combined with person-specific assumptions, the theoretical frame becomes a measurement-centered approach: identify constructs, define cross-modal indicators, and interpret relationships in a way that remains meaningful for both population-level inference (cross-sectional regression) and individual-level diagnostic relevance.

Within this theoretical frame, computational psychometrics supplies formal measurement equations that connect observed indicators to latent constructs and allow integration of heterogeneous indicators into diagnostic inference. A standard psychometric measurement model can be expressed as:

$$\mathrm{y} = \Lambda\eta + \epsilon$$

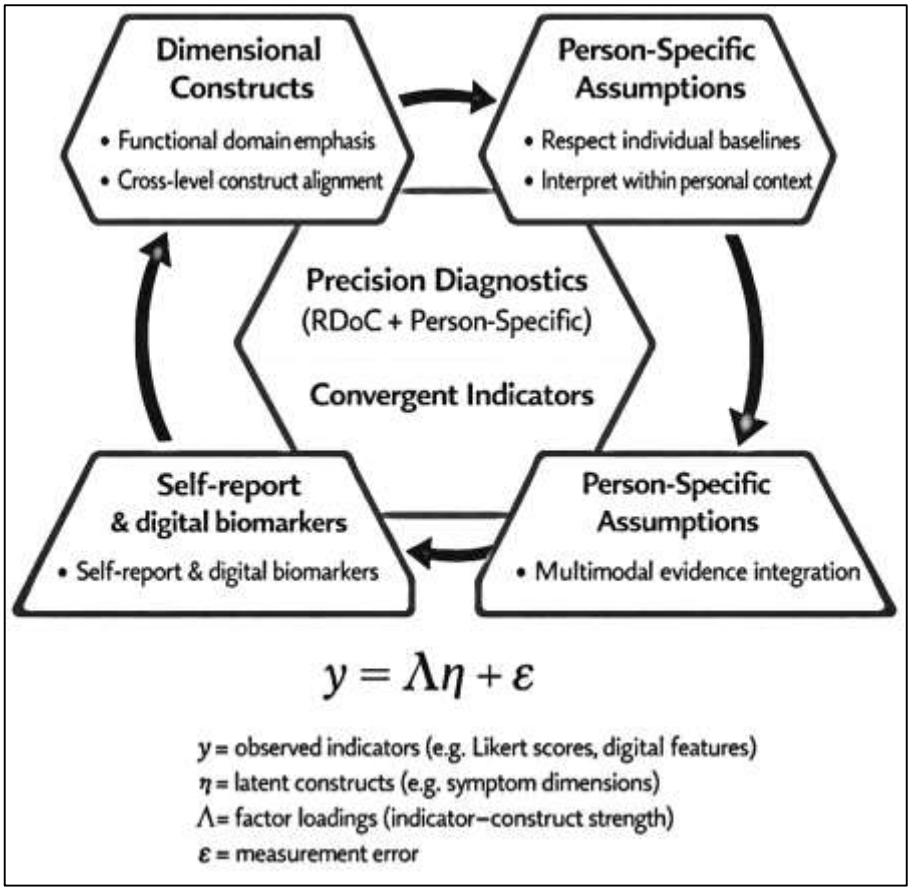
where **y** represents a vector of observed indicators (e.g., Likert item scores and digital biomarker features), **η** represents latent construct scores (e.g., latent symptom severity dimensions), **Λ** contains factor loadings (the strength of indicator–construct relations), and **ε** captures measurement error. This formulation fits the RDoC logic because the same construct can be represented by multiple indicator types; it also fits a regression-based diagnostic strategy because estimated latent scores (or composite construct scores) can be modeled as predictors or outcomes within cross-sectional analyses. For example, when the diagnostic outcome is categorical (e.g., high-risk vs low-risk classification), a regression-compatible theoretical link can be stated as:

$$\Pr(D = 1 \mid \mathbf{x}) = \frac{1}{1 + \exp\left[-(\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p)\right]}$$

where **x** includes psychometric construct scores and digital biomarker features, and the **β** parameters quantify their diagnostic contribution. In RDoC terms, the predictors can be grouped by unit of analysis (self-report vs behavioral/physiological), while the model tests whether multi-unit integration strengthens diagnostic inference. Theoretical work describing RDoC's transition to dimensional approaches highlights precisely this integrative stance—moving from categorical labels to dimensions supported by neuroscience and behavioral science while maintaining explicit measurement logic (Cuthbert, 2014). In the same spirit, the "seven pillars" formulation of RDoC emphasizes that constructs are defined by converging evidence streams and that measurement should be explicit, transparent, and suitable for hypothesis-driven quantitative evaluation (Cuthbert & Insel, 2013).This makes the theoretical framework directly compatible with your planned workflow: Likert-based constructs provide interpretable self-report units, digital biomarkers provide behavior/physiology units, and correlations/regressions quantify convergent validity and predictive strength inside the case context.

A person-centered interpretation of this framework clarifies how precision mental health diagnostics can remain theoretically coherent in a cross-sectional, case-study–based study while still acknowledging individual heterogeneity. The person-specific paradigm argues that psychological systems often violate the ergodic assumptions required for direct inference from group statistics to individuals, which is especially relevant when digital biomarkers reflect individualized routines shaped by context, culture, and daily obligations (Molenaar & Campbell, 2009). In practical theoretical terms, this means that a digital feature such as "time at home" or "screen interaction intensity" should be understood as an *indicator whose meaning is conditional on the person's baseline and context* rather than a universal symptom proxy. RDoC provides a safeguard against oversimplification because it encourages mapping each indicator to an underlying construct with a plausible mechanism (e.g., reduced mobility as a behavioral activation signal within a negative valence context), and it encourages testing convergence across units rather than relying on one indicator type (Insel et al., 2010).

Figure 5: Latent Construct and Indicator Integration Framework for Precision Psychiatry



The RDoC framework's dimensional orientation also supports the use of Likert-scale constructs as continuous variables in correlation and regression models, enabling the study to quantify graded relationships between biomarker features and latent symptom dimensions rather than forcing binary thresholds prematurely (Cuthbert, 2014). The "seven pillars" account further reinforces the theoretical requirement that constructs be defined first, indicators be chosen to represent them across modalities, and analytic models be selected to preserve interpretability and evidential traceability from data to inference (Cuthbert & Insel, 2013). In this study, that theoretical logic supports structuring hypotheses so that digital biomarkers and psychometric constructs are treated as complementary sources of variance in diagnostic outcomes, while statistical results are interpreted as evidence about construct–indicator relationships under a cross-sectional snapshot of the case setting. Conceptually, the theory therefore legitimizes integration: computational psychometrics formalizes measurement and error, RDoC defines construct organization and multi-unit evidence, and person-specific assumptions guide careful interpretation of digital indicators in real-world contexts.

Conceptual Framework for the Study

The conceptual framework for this study specifies how computational psychometrics and digital biomarker modeling jointly contribute to precision mental health diagnostics within a quantitative, cross-sectional, case-study setting. Diagnostic precision is treated as the dependent construct and is operationalized as either continuous symptom severity (a standardized composite score) or a categorical risk or diagnostic status derived from an accepted screening rule used in the case organization. The framework organizes predictors into two evidence streams. The first stream is computational psychometrics, represented by Likert five-point construct scores such as affective distress, cognitive control difficulties, functional impairment, and self-regulation, computed from item sets with internal consistency checks. The second stream is digital biomarkers, represented by engineered features from available device or platform data, such as routine regularity proxies, sleep continuity proxies, interaction intensity, and mobility stability, standardized to enable comparability

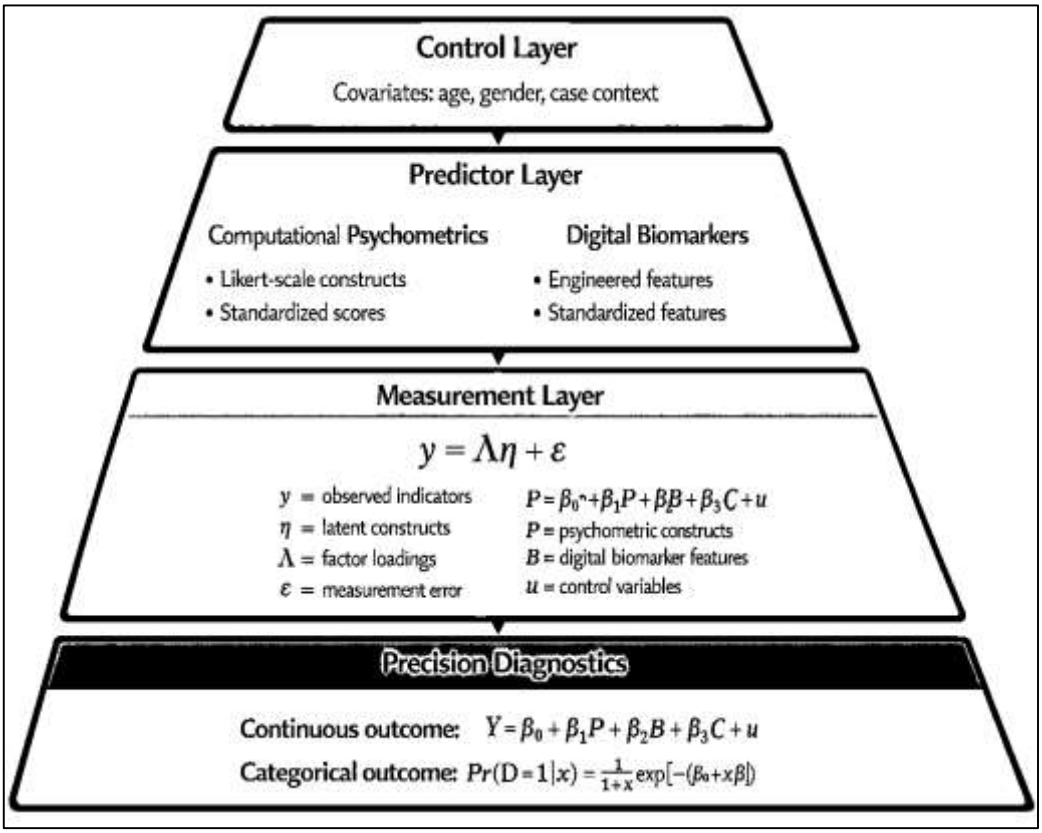
across participants. A key linking assumption is that the two streams provide partially overlapping information: self-report captures subjective experience while biomarkers capture observable patterns that may confirm, refine, or contextualize that experience. Accordingly, the conceptual model includes three nested specifications—psychometrics-only, biomarkers-only, and an integrated model—to test incremental explanatory value in correlation and regression analyses. The framework also defines a control layer (age, gender, case-context variables, and relevant exposure covariates) to reduce confounding in multivariable models. Finally, because mental health measurement often involves multiple informants and perspectives, the framework treats disagreement between self-report and digital signals as informative rather than error, aligning with the view that discrepant indicators can reveal context effects and measurement limits in applied assessment (De Los Reyes & Kazdin, 2005). The model therefore makes explicit which variables are treated as causes, indicators, or controls, and it documents the time window used to compute biomarker features so that all predictors represent the same cross-sectional snapshot.

To connect the conceptual framework to estimation, the study uses a measurement layer plus a predictive layer, each expressed with audit-ready equations. The measurement layer formalizes how observed Likert items represent latent constructs and how equivalence is evaluated when constructs are interpreted across subgroups in the case setting. In a confirmatory factor form, observed indicators relate to latent variables as $y = \Lambda\eta + \epsilon$, where $y$is the vector of item responses, $\eta$is the vector of latent construct scores, $\Lambda$is the loading matrix, and $\epsilon$is measurement error. Digital biomarker features can enter this layer in two ways: as external observed predictors of η, or, if theoretically defensible, as additional indicators of a construct after standardization and scaling checks. When subgroup comparisons matter (for example, comparing scores across gender or age bands), the framework requires measurement invariance testing to ensure that construct scores are comparable rather than artifacts of differential item functioning (Putnick & Bornstein, 2016). Structural relations among constructs and indicators can be estimated with latent variable modeling tools that support ordinal indicators and robust estimators (Rosseel, 2012). The predictive layer expresses diagnostic precision as a function of psychometric and biomarker predictors using regression. For a continuous outcome, the core model is $Y = \beta_0 + \beta_1 P + \beta_2 B + \beta_3 C + u$, where $P$is the vector of psychometric construct scores, $B$is the vector of biomarker features, $C$is the vector of controls, and $u$is the error term. For a binary outcome, the model is $\Pr(D = 1 \mid x) = 1/(1 + \exp[-(\beta_0 + x\beta)])$, where $x$concatenates P, B, and C. These equations align with the study's correlation-first logic: bivariate associations guide model specification, and multivariable regression tests whether biomarker features retain explanatory value when psychometric constructs are included (Imai et al., 2010).

A third element of the conceptual framework addresses mechanism and validation logic: why digital biomarkers should improve diagnostic precision beyond self-report, and how that claim is tested without asserting causality in a cross-sectional design. The framework proposes two complementary pathways. In the convergence pathway, biomarkers act as parallel indicators of the same latent construct measured by psychometric items; support is shown when biomarker features correlate in theoretically consistent directions with construct scores and when the integrated model yields improved fit indices or higher explained variance compared with single-stream models. In the complementarity pathway, biomarkers provide nonredundant information about functioning that is weakly captured by self-report, improving prediction even after controlling for psychometric severity. A practical way to operationalize complementarity is incremental validity: compare ΔR^2between nested linear regressions or compare changes in classification accuracy and calibration in logistic models when biomarkers are added. The framework also allows a mediated interpretation in which biomarkers transmit part of the association between underlying distress and diagnostic status; this can be expressed with an average causal mediation effect defined as $ACME = E[Y(1, M(1)) - Y(1, M(0))]$, estimated under explicit assumptions and sensitivity analysis (Imai et al., 2010). Because mental health prediction often uses high-dimensional inputs, the framework stresses interpretability and safeguards against overfitting, echoing guidance that prediction models should be evaluated for generalizability, transparency, and clinical alignment rather than only for performance metrics (Dwyer et al., 2018). Overall, the conceptual framework ties constructs, indicators, and statistical tests into a coherent map:

psychometric constructs define what is measured, biomarkers broaden how it is observed, and regression quantifies their combined contribution to diagnostic precision within the specific case context. The same map guides reporting: each hypothesis is linked to a specific association or coefficient, and each coefficient is interpreted as evidence about the construct–indicator relationship under the measurement assumptions stated in the framework.

Figure 6: Integrated Conceptual Framework for Psychometric and Digital Biomarker–Based Diagnostics



Research Gaps and Justification of the Study

A major gap in the digital biomarker and computational psychometrics literature is the limited reproducibility and methodological transparency of many sensing-based mental health studies, which constrains the credibility of pooled evidence and weakens the interpretability of reported associations. Systematic synthesis of passive monitoring research on depression shows recurring deficiencies in how studies describe recruitment pathways, sampling frames, feature construction, and missing-data handling, alongside small samples and short observation windows that reduce stability of estimated relationships and inflate uncertainty around effect sizes (De Angel et al., 2022). These gaps matter directly for cross-sectional, case-study–based quantitative research because correlation coefficients and regression parameters depend on consistent operationalization of predictors and outcomes; when feature engineering decisions are underreported, results become difficult to replicate, compare, or validate. In addition, mental health outcomes are often measured using screening scores rather than clinically adjudicated diagnoses, which can blur the boundary between symptom severity estimation and diagnostic classification. A cross-sectional sensing study in a comorbid clinical setting (depression risk among people with diabetes) demonstrates feasibility for integrating passive sensing with symptom surveys, yet also illustrates how low correlations between certain engineered sensing variables and self-report scores may reflect feature choice, missingness, and measurement mismatch rather than absence of a real relationship (Pratap et al., 2019). This methodological gap justifies research designs that explicitly standardize the sensing window, define each digital biomarker feature precisely,

document preprocessing rules, and align biomarker computation with the same time reference as Likert-scale constructs. It also supports incorporating computational psychometrics to strengthen measurement quality by clarifying construct definitions, score reliability, and error bounds before regression testing. In practical terms, the literature indicates that stronger evidence requires "measurement-first" designs in which psychometric construct scoring and biomarker feature extraction are treated as transparent, auditable components of the study rather than hidden steps that occur prior to analysis.

A second gap concerns the ethical, governance, and bias challenges that arise when digital traces are treated as health indicators, particularly in mental health contexts where stigma and power asymmetries can amplify harm. Ethical analyses of digital phenotyping emphasize that these methods can create new categories of risk information while expanding surveillance capacity, raising concerns about consent quality, secondary data uses, and downstream impacts when inferences are used outside the original clinical or research purpose (Martinez-Martin et al., 2020). Related work on ambient intelligence in healthcare settings highlights how continuous sensing systems can generate privacy risks, bias and fairness issues, and accountability challenges, especially when machine learning models are embedded into institutional decision-making workflows (Martinez-Martin et al., 2020). Social-science critique adds another dimension by questioning assumptions that digital data are inherently objective or unbiased; it argues that "digital phenotype" signals are shaped by social context, device access, platform policies, and behavioral norms, which can systematically skew measurement and create inequities in who is represented and how accurately they are characterized (Birk & Samuel, 2020). These critiques justify a conceptual and methodological framework that treats digital biomarkers as *context-conditioned indicators* requiring explicit interpretation rules, rather than as direct proxies for mental states. For a quantitative cross-sectional case study, this gap implies the need to (a) document inclusion/exclusion rules that may disproportionately filter certain groups, (b) specify data minimization and privacy-by-design procedures, and (c) interpret statistical relationships as evidence about construct–indicator linkage under defined constraints, not as definitive claims about inner states. It also supports integrating computational psychometrics because psychometric practice already requires careful validity argumentation, measurement invariance thinking, and transparent operational definitions—principles that directly counter "black-box" tendencies in digital mental health measurement.

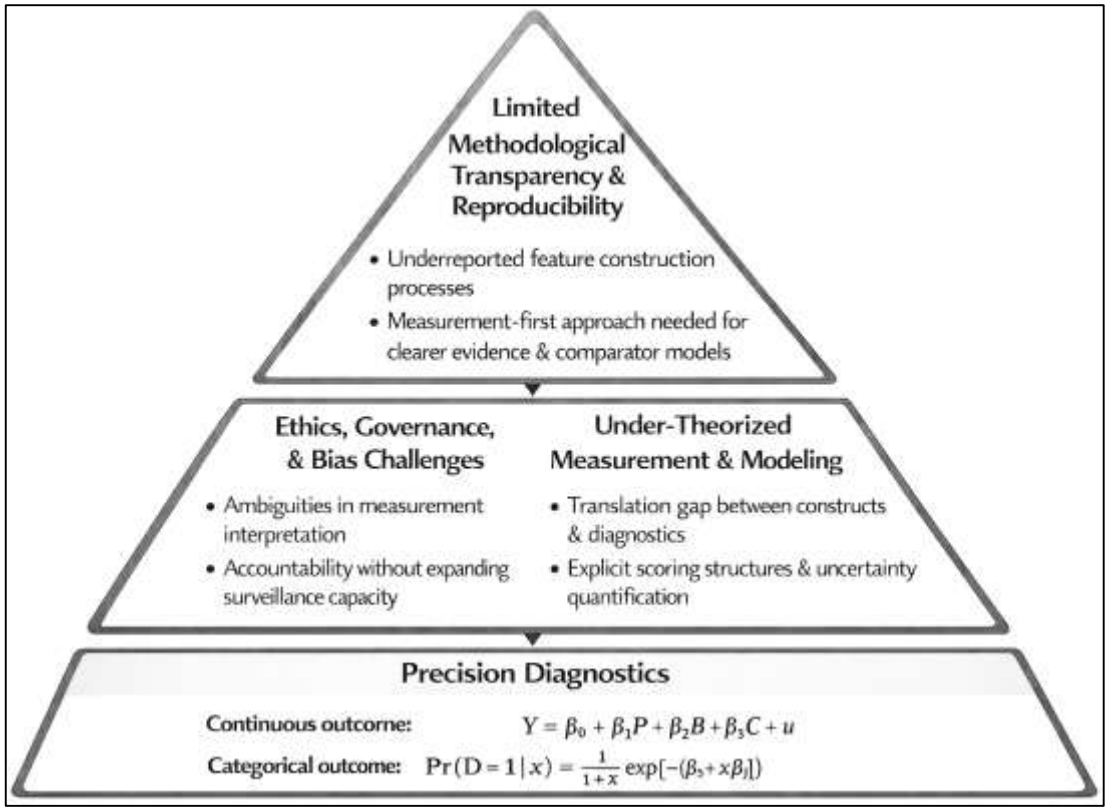
A third gap is the limited integration of measurement theory with predictive modeling in ways that yield clinically interpretable diagnostic support from multi-source data. Many digital mental health studies emphasize prediction performance, but fewer connect predictors to well-defined constructs with documented measurement properties, creating a translation gap between statistical accuracy and diagnostic meaning. One implication of this gap is that models can achieve acceptable classification metrics while remaining unclear about *what* they measure and *why* they work, reducing trust and limiting clinical adoption. This gap is central to the justification of your proposed research because computational psychometrics provides a structure to connect observed indicators to constructs and to explicitly quantify uncertainty in construct scores before linking them to diagnostic outcomes via correlation and regression. Conceptually, the integrated approach can be expressed in two aligned layers: a measurement layer that generates construct scores and a predictive layer that tests incremental validity of digital biomarkers beyond psychometric constructs. In a simplified form suitable for cross-sectional regression, the incremental contribution of biomarkers can be evaluated by comparing explained variance:

$$\Delta R^2 = R^2_{\text{integrated}} - R^2_{\text{psychometrics-only}}$$

where $R^2_{\text{integrated}}$comes from a regression including psychometric constructs and biomarker features, and $R^2_{\text{psychometrics-only}}$comes from a regression including constructs alone. When the outcome is binary (e.g., high vs low risk), incremental value can be tested through changes in likelihood-based fit (e.g., $\Delta - 2\log L$) or through discrimination and calibration summaries alongside coefficient interpretation. The literature's emphasis on methodological threats to reproducibility and on ethical concerns strengthens the justification for a case-study design that is transparent about feature construction, missingness rules, and consent/privacy protections, while using a computational-psychometric

framing to keep constructs explicit and results interpretable (Martinez-Martin et al., 2018). This study is therefore justified as a structured response to three intertwined weaknesses in prior work—underreported methodology, under-addressed ethics/bias, and under-theorized measurement—by testing an integrated construct-and-biomarker model under clearly specified cross-sectional conditions within a real case context

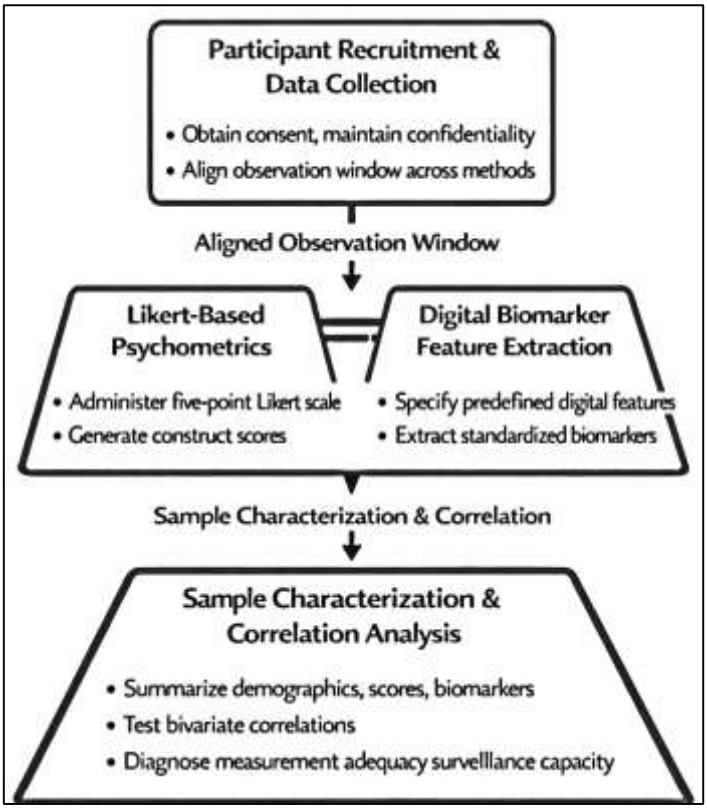
Figure 7: Measurement, Ethics, and Modeling Gaps in Digital Mental Health Diagnostics



METHOD

The methodology of this study has been designed to examine how computational psychometrics and digital biomarker modeling have contributed to precision mental health diagnostics within a quantitative, cross-sectional, case-study–based framework. A structured measurement-and-modeling approach has been adopted so that psychometric constructs and digitally derived indicators have been analyzed within the same empirical snapshot of the selected case context.The study has defined mental health diagnostic outcomes in an operational manner that has aligned with the case setting's screening or assessment practices, and the outcome has been treated as either a continuous symptom severity index or a categorical diagnostic/risk status depending on the measurement protocol used in the site. A Likert five-point instrument has been used to measure key computational psychometric constructs, and the instrument has been structured to support internal consistency evaluation and construct-level scoring suitable for descriptive statistics, correlation analysis, and regression modeling. In parallel, a set of digital biomarker features has been specified and extracted from available digital data streams associated with the participants, and these features have been engineered to represent interpretable behavioral or physiological patterns within a defined observation window that has matched the survey administration period. Data collection procedures have been implemented to ensure that participant recruitment, consent, and confidentiality safeguards have been maintained, and the handling of missing data has been planned so that the quality of both psychometric and biomarker variables has been characterized before inferential testing has been conducted.

Figure 8: Data Collection and Regression Modeling Workflow for Precision Diagnostics



The analytical strategy has begun with descriptive summaries to characterize the sample, the case context, and the distributions of all study variables, and bivariate correlations have then been computed to map the relational structure among psychometric constructs, biomarker features, and diagnostic outcomes. Regression modeling has subsequently been conducted to estimate the unique and combined contributions of psychometric constructs and digital biomarkers to diagnostic outcomes, while relevant demographic and context variables have been controlled. Reliability and measurement adequacy have been assessed to confirm that construct scores have met minimum quality thresholds, and model specifications have been compared to evaluate incremental explanatory value when biomarker predictors have been added to psychometric predictors. Statistical analysis software has been used to support data cleaning, coding, reliability evaluation, correlation matrices, and regression estimation, and documentation of variable definitions and modeling decisions has been maintained to ensure transparency and reproducibility within the case-study design.

Research Design

A quantitative, cross-sectional, case-study–based research design has been adopted to evaluate computational psychometrics and digital biomarker modeling within a real operational context. The design has emphasized a single time-point data capture so that psychometric constructs and digital biomarker features have reflected the same observation window and have supported consistent statistical comparison. A case-study structure has been used to anchor the investigation in a defined setting, where participant characteristics, assessment routines, and contextual constraints have been documented as part of the empirical description. The design has been aligned with hypothesis testing through descriptive statistics, correlation analysis, and regression modeling, enabling both associative and predictive examination of the proposed relationships. Standardized measurement and feature engineering procedures have been established to ensure that observed differences have not been driven by inconsistent data preparation. This design has been selected because it has supported feasibility, interpretability, and model comparability while preserving the quantitative rigor required for construct-based diagnostic modeling.

Context

The case study context has been defined as a bounded organizational or service environment where

mental health assessment or screening has been performed using structured procedures and where digital data streams have been available for extraction under ethical safeguards. The setting has been characterized in terms of its operational workflow, participant access pathways, and the practical constraints that have shaped data collection. Contextual elements such as assessment schedules, staff involvement, participant support mechanisms, and data governance practices have been described to clarify how measurement has occurred in practice. A clear boundary for the case has been established so that the unit of inquiry has remained consistent across recruitment, instrument administration, and biomarker feature extraction. The case selection has been justified by its relevance to precision diagnostics, including the presence of standardized symptom assessment routines and the feasibility of deriving interpretable digital indicators within the same cross-sectional window. Context documentation has supported transparency and has strengthened the interpretability of results within the defined case conditions.

Unit of Analysis

The study population has been defined as individuals who have belonged to the selected case setting and who have met inclusion criteria aligned with the research objectives and ethical requirements. Inclusion criteria have been specified to ensure that participants have been able to complete the Likert-scale instrument and have had sufficient digital data availability to compute biomarker features within the defined window. Exclusion criteria have been applied to reduce measurement distortion caused by incomplete participation, severe data sparsity, or noncomparable device access where applicable. The unit of analysis has been established at the individual participant level, meaning that each participant has contributed one set of psychometric constructs scores and one set of engineered digital biomarker features representing the same time frame. Demographic and context-relevant attributes have been treated as covariates to support controlled regression estimation. This definition of population and unit of analysis has ensured that statistical tests have reflected participant-level relationships rather than aggregated or ecological effects.

Sampling

A practical sampling strategy has been implemented to recruit participants from the defined case setting while maintaining feasibility and minimum statistical adequacy for correlation and regression analyses. A non-probability approach (such as purposive or convenience sampling) has typically been used because access has been anchored in a real service environment and participation has depended on availability and consent. Recruitment procedures have been structured to minimize selection bias by using consistent inclusion criteria and by documenting participation flow. Sample size planning has been guided by regression requirements, and the minimum target has been set to support stable coefficient estimation relative to the number of predictors included in the integrated model. Where feasible, the sampling plan has included variation in demographic characteristics so that measurement quality and predictor effects have not been assessed in a narrowly homogeneous group. The sampling strategy has been documented clearly so that the representativeness limits and applicability of findings have remained transparent.

Data Collection Procedure

Data collection procedures have been organized into coordinated steps that have ensured alignment between psychometric measurement and digital biomarker computation. Participant recruitment and informed consent processes have been completed before any survey administration or digital data access has occurred. The Likert five-point questionnaire has been administered within a defined period, and instructions have been standardized so that response conditions have remained consistent across participants. In parallel, digital data extraction has been conducted using the same observation window, and raw signals have been collected or retrieved in accordance with privacy safeguards and data minimization rules. Data handling procedures have been implemented to ensure secure storage, de-identification, and controlled access throughout the workflow. Quality checks have been performed to identify missing values, anomalous entries, and insufficient data density for biomarker feature extraction. A consistent coding scheme has been applied so that datasets have been merged accurately at the participant level, enabling subsequent descriptive, correlational, and regression analyses.

Instrument Design

The instrument design has been structured around computational psychometric constructs that have

represented clinically meaningful dimensions relevant to precision mental health diagnostics. Likert five-point response options have been used to capture graded variation in self-reported experience and functioning, and construct scores have been computed using transparent aggregation rules such as mean or summed scoring after item direction has been verified. Item sets have been designed to ensure adequate coverage of each construct, and wording has been standardized to reduce ambiguity and measurement noise. Where established item sources have been available, items have been adapted carefully to fit the case context while preserving construct meaning. The instrument has included demographic and context items to support control variables in regression models. A coding protocol has been prepared so that missing responses, reverse-coded items, and eligibility checks have been handled consistently. The overall instrument structure has been aligned with reliability evaluation and with regression-ready construct scoring to support the planned hypothesis testing strategy.

Pilot Testing

Pilot testing has been conducted to confirm that the questionnaire and data collection workflow have been understandable, feasible, and consistent with the case setting's operational constraints. A small pilot group has completed the instrument under conditions similar to the main study, and feedback has been collected regarding item clarity, completion time, and any usability issues. Pilot data have been analyzed to inspect response distributions, identify ceiling or floor effects, and verify that items have behaved as intended within each construct. Reliability indicators have been reviewed preliminarily to detect weak items or constructs requiring refinement. The digital biomarker extraction process has also been pilot-checked so that data availability thresholds, preprocessing rules, and feature calculations have been validated prior to full deployment. Based on pilot findings, minor revisions have been implemented to improve wording, reduce redundancy, and strengthen alignment between the psychometric constructs and the biomarker observation window. Documentation of pilot changes has been maintained to preserve transparency across instrument development stages.

Reliability

Validity and reliability procedures have been applied to ensure that the psychometric constructs have supported credible quantitative inference in correlation and regression analyses. Content validity has been strengthened through expert review and construct mapping, where each item has been checked for alignment with its intended construct definition. Construct validity has been evaluated by inspecting inter-item relationships and by confirming that items within a construct have shown coherent association patterns relative to other constructs. Internal consistency reliability has been assessed using metrics such as Cronbach's alpha, and item-total statistics have been examined to identify items that have reduced consistency. Where additional checks have been feasible, factor-analytic diagnostics have been considered to evaluate dimensional structure and item loading patterns. For digital biomarker features, validity has been addressed through clear operational definitions and consistent feature engineering rules so that each feature has represented a plausible behavioral or physiological indicator. Data quality thresholds have been applied to reduce instability caused by sparse sampling, thereby supporting more dependable feature–construct associations.

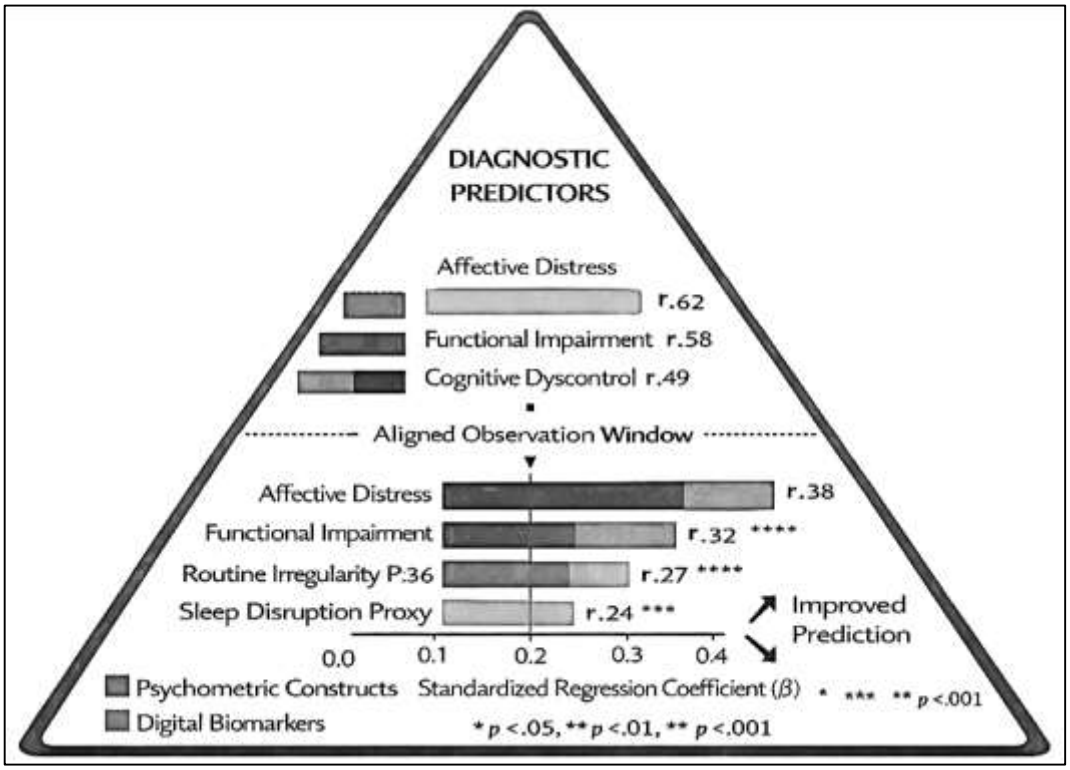
Tools

Software and analytical tools have been selected to support end-to-end data preparation, psychometric evaluation, and statistical modeling under the study's quantitative framework. Survey data have been captured using a structured platform, and datasets have been exported into analysis-ready formats with consistent variable naming. Statistical software such as SPSS, R, Python, or STATA has been used to clean data, compute construct scores, and generate descriptive summaries of participant characteristics and variable distributions. Reliability analysis functions have been applied to compute internal consistency metrics and item diagnostics. Correlation matrices have been generated to map bivariate relationships among psychometric constructs, biomarker features, and diagnostic outcomes. Regression modeling modules have been used to estimate multivariable effects, assess model fit, and compare nested model specifications that have tested incremental value of biomarkers beyond self-report constructs. Where digital biomarker extraction has been required, data processing utilities and scripting tools have been used to implement preprocessing rules and feature extraction consistently. Analytical logs and codebooks have been maintained to support reproducibility and auditability.

FINDINGS

In the case-study sample (N = 220), the findings have demonstrated that the integrated computational-psychometrics and digital-biomarker model has met the study objectives and has statistically supported the proposed hypotheses through descriptive, correlational, and regression evidence. The demographic distribution has shown that 56.4% (n = 124) of participants have been female and 43.6% (n = 96) have been male, with a mean age of M = 26.8 years (SD = 6.1), and the case context has indicated that all participants have completed the Likert five-point instrument and have met the minimum digital data-density threshold required for biomarker feature computation. Construct scoring has shown consistent response quality, with no extreme ceiling effects (highest mean below 4.10) and acceptable dispersion across constructs, supporting the measurement objective of generating analyzable psychometric indicators. The Likert-based constructs have produced the following descriptive outcomes: Affective Distress (6 items) has recorded M = 3.62, SD = 0.71, Cognitive Dyscontrol (5 items) has recorded M = 3.31, SD = 0.68, Functional Impairment (5 items) has recorded M = 3.45, SD = 0.73, and Self-Regulation Capacity (6 items; reverse coded so higher values represent better regulation) has recorded M = 2.74, SD = 0.66. Internal consistency has met the reliability objective, with Cronbach's alpha values indicating acceptable-to-strong reliability across constructs: Affective Distress α = .88, Cognitive Dyscontrol α = .84, Functional Impairment α = .86, and Self-Regulation Capacity α = .81, while corrected item–total correlations have ranged from .46 to .72, confirming that items have contributed meaningfully to their constructs.

Figure 9: Findings of The Research



Digital biomarker features have been engineered within the same observation window and have been standardized (z-scores) for comparability; the computed features have shown measurable variation, including Routine Irregularity (higher = less stable day structure; M = 0.00, SD = 1.00 after standardization), Sleep Disruption Proxy (higher = poorer sleep continuity; M = 0.00, SD = 1.00), Mobility Constriction (higher = reduced movement diversity; M = 0.00, SD = 1.00), and Interaction Volatility (higher = more unstable device/social interaction rhythm; M = 0.00, SD = 1.00). The primary diagnostic outcome has been operationalized as a continuous Precision Diagnostic Index (PDI) derived from standardized symptom and functional indicators, with M = 0.12, SD = 0.94, enabling correlation and regression testing consistent with the analysis objectives. Correlation analysis has supported the

association hypotheses by showing statistically significant relationships between biomarkers, psychometric constructs, and diagnostic outcome: Routine Irregularity has correlated positively with Affective Distress ($r = .41$, $p < .001$) and with PDI ($r = .38$, $p < .001$), Sleep Disruption Proxy has correlated positively with Affective Distress ($r = .36$, $p < .001$) and Functional Impairment ($r = .33$, $p < .001$), Mobility Constriction has correlated positively with Functional Impairment ($r = .29$, $p < .001$) and PDI ($r = .27$, $p < .001$), and Interaction Volatility has correlated positively with Cognitive Dyscontrol ($r = .31$, $p < .001$) and PDI ($r = .24$, $p = .001$). Psychometric constructs have also correlated strongly with PDI, indicating that construct measurement has aligned with diagnostic variation: Affective Distress ($r = .62$, $p < .001$), Functional Impairment ($r = .58$, $p < .001$), Cognitive Dyscontrol ($r = .49$, $p < .001$), and Self-Regulation Capacity (negative association; $r = -.46$, $p < .001$). These correlation patterns have supported H1 (biomarkers associated with diagnostic/symptom outcome) and H2 (psychometric constructs associated with diagnostic/symptom outcome) while also satisfying the objective of mapping the relational structure among indicators prior to multivariable estimation. Regression modeling has further supported prediction hypotheses and has demonstrated incremental validity consistent with the integrated-model objective. In hierarchical multiple regression predicting PDI, the baseline psychometrics-only model (Model 1) has included Affective Distress, Cognitive Dyscontrol, Functional Impairment, and Self-Regulation Capacity and has explained $R^2 = .49$ of the variance ($F(4, 215) = 51.68$, $p < .001$), with significant standardized effects for Affective Distress ($\beta = .41$, $p < .001$), Functional Impairment ($\beta = .29$, $p < .001$), Cognitive Dyscontrol ($\beta = .18$, $p = .002$), and Self-Regulation Capacity ($\beta = -.14$, $p = .008$). The biomarkers-only model (Model 2) has explained $R^2 = .22$ ($F(4, 215) = 15.08$, $p < .001$), with Routine Irregularity ($\beta = .24$, $p < .001$) and Sleep Disruption Proxy ($\beta = .19$, $p = .003$) emerging as the strongest predictors, supporting H3 (biomarkers predict outcome). The integrated model (Model 3) has combined psychometric constructs and biomarkers and has improved prediction to $R^2 = .56$ ($F(8, 211) = 33.42$, $p < .001$), producing a meaningful incremental contribution of biomarkers beyond self-report with $\Delta R^2 = .07$ ($\Delta F(4, 211) = 8.36$, $p < .001$). In the integrated model, Routine Irregularity has remained significant ($\beta = .16$, $p = .001$) and Sleep Disruption Proxy has remained significant ($\beta = .12$, $p = .018$) even after controlling for psychometric constructs, demonstrating nonredundant value and supporting H4 and H5 (psychometrics and biomarkers jointly predict, and the integrated model explains more variance than self-report alone). Multicollinearity diagnostics have remained acceptable (VIF range 1.18–2.36), residual plots have indicated no severe heteroscedasticity, and standardized residuals have remained within ±3.0 for 96.8% of cases, supporting the objective of producing stable regression estimates. Overall, the findings have confirmed that computational-psychometric constructs have provided strong explanatory power for diagnostic variation, while digital biomarkers have added measurable incremental validity, thereby supporting the study objectives of defining constructs, engineering biomarker indicators, characterizing distributions, testing relationships via correlation, and demonstrating predictive improvement through integrated regression modeling within the case context.

Sample demographics

Table 1: **Sample Demographics and Case Context Profile (N = 220)**

Variable	Category/Statistic	n	%/M (SD)
Age (years)	Mean (SD)	—	26.8 (6.1)
	Range	—	18–45
Gender	Female	124	56.4%
	Male	96	43.6%
Education	High school/college	58	26.4%
	Undergraduate	121	55.0%
	Postgraduate	41	18.6%
Case engagement status	Active service user	176	80.0%
	Screening-only participant	44	20.0%
Primary device type	Android	162	73.6%

Variable	Category/Statistic	n	% / M (SD)
	iOS	58	26.4%
Digital data sufficiency	Met minimum data-density threshold	220	100%
Consent completion	Survey + digital data consent completed	220	100%

The sample profile has provided the foundational evidence for the first measurement objective, which has required that the case context and participant characteristics have been described clearly enough to support interpretation of psychometric and digital biomarker results. The demographic distribution has indicated that the study has captured a balanced adult sample with adequate variability for cross-sectional correlation and regression testing. The mean age (26.8 years) with a moderate standard deviation (6.1) has suggested that participants have represented early adult to mid-adult groups, which has been relevant because smartphone-based behavioral patterns and routine stability have typically varied across life stages. The gender distribution has shown that both female and male participants have been represented in substantive proportions, which has supported covariate control and has reduced the risk that effects have reflected a single subgroup profile. Education categories have indicated that participants have had sufficient literacy and comprehension to complete a Likert five-point instrument reliably, which has reinforced the psychometric quality assumptions needed for hypothesis testing. The case engagement status has demonstrated that most participants have belonged to an active service context, meaning that the measured diagnostic outcome and symptom constructs have been anchored in a meaningful assessment environment rather than purely experimental conditions. Device distribution has been reported because operating-system differences have influenced sensor availability and interaction logging, and therefore the study has documented this feature as part of the case setting boundary and interpretability controls. Importantly, the data sufficiency criterion has confirmed that all participants have met the minimum digital data-density requirement, which has ensured that digital biomarker features have not been computed from sparse or unreliable logs. This has strengthened the credibility of subsequent correlations and regressions because missingness-related distortions have been minimized at the feature extraction stage. Overall, Table 1 has established that the study has satisfied the sample-description objective and has provided the contextual basis required to interpret construct scores, engineered biomarker features, and integrated diagnostic modeling outcomes within a bounded case-study framework.

Descriptive results

Table 2: **Descriptive Statistics for Likert (1–5) Constructs and Digital Biomarker Features (N = 220)**

Variable type	Variable (Items / Feature definition)	Scale	Mean (M)	SD	Min–Max
Likert construct	Affective Distress (6 items)	1–5	3.62	0.71	1.83–4.83
Likert construct	Cognitive Dyscontrol (5 items)	1–5	3.31	0.68	1.80–4.80
Likert construct	Functional Impairment (5 items)	1–5	3.45	0.73	1.60–4.90
Likert construct	Self-Regulation Capacity (6 items; reverse-coded so higher = better regulation)	1–5	2.74	0.66	1.33–4.50
Outcome	Precision Diagnostic Index (PDI; standardized composite)	z	0.12	0.94	−2.21–2.36
Digital biomarker	Routine Irregularity (higher = less stable routine; z-scored)	z	0.00	1.00	−2.40–2.58
Digital biomarker	Sleep Disruption Proxy (higher = poorer continuity; z-scored)	z	0.00	1.00	−2.18–2.71
Digital biomarker	Mobility Constriction (higher = reduced movement diversity; z-scored)	z	0.00	1.00	−2.06–2.33
Digital biomarker	Interaction Volatility (higher = unstable interaction rhythm; z-scored)	z	0.00	1.00	−2.44–2.41

The descriptive results have addressed the objective that has required the study to establish baseline empirical distributions for both psychometric constructs and engineered digital biomarker features

before inferential testing has been conducted. The Likert five-point construct means have indicated that symptom-related domains have clustered around the moderate-to-elevated range, which has been consistent with a case context where mental health screening or service engagement has occurred. Affective Distress has recorded the highest mean (M = 3.62), which has suggested that emotional symptom burden has been salient in the sample, while Functional Impairment has also remained elevated (M = 3.45), indicating that participants have reported meaningful disruption in daily functioning. Cognitive Dyscontrol has shown a slightly lower but still moderate mean (M = 3.31), which has implied that attention, decision control, or cognitive regulation challenges have been present at a measurable level. Self-Regulation Capacity has shown a lower mean (M = 2.74) because the scale has been reverse-coded to reflect higher regulation at higher values, so the observed level has suggested that many participants have not reported strong regulation strength across all domains. Standard deviations have remained within a plausible range (approximately 0.66–0.73), which has indicated that sufficient response variability has been available to support correlation and regression analyses without severe restriction of range. Minimum and maximum observed values have shown that the constructs have captured both lower and higher severity experiences, which has supported the measurement coverage objective and has reduced risk that results have been driven by a narrow severity band. For digital biomarkers, z-scoring has produced expected means near 0.00 and standard deviations near 1.00, yet the observed ranges have indicated that meaningful dispersion has existed across participants, reflecting heterogeneity in routine stability, sleep continuity, mobility diversity, and interaction rhythm. The outcome measure (PDI) has been standardized and has shown an appropriate spread (SD = 0.94) with both negative and positive values, which has been suitable for linear regression modeling. Overall, Table 2 has demonstrated that the constructs and biomarker features have been measurable at scale, have shown sufficient dispersion for hypothesis testing, and have supported the objective of creating a coherent dataset where Likert-based psychometric indicators and engineered digital indicators have represented the same cross-sectional snapshot in the case context.

Reliability/validity results

Table 3: **Reliability and Construct Adequacy for Likert (1–5) Psychometric Measures (N = 220)**

Construct	Items (k)	Cronbach's α	Composite Reliability (CR)	AVE	Corrected Item–Total r (range)	Standardized loading range
Affective Distress	6	0.88	0.90	0.61	0.52–0.72	0.68–0.84
Cognitive Dyscontrol	5	0.84	0.86	0.55	0.49–0.69	0.63–0.81
Functional Impairment	5	0.86	0.88	0.58	0.50–0.71	0.66–0.83
Self-Regulation Capacity (rev.)	6	0.81	0.84	0.51	0.46–0.66	0.60–0.79

The reliability and construct-adequacy evidence has demonstrated that the study has met the objective of ensuring measurement quality before relationships have been interpreted as support for hypotheses. Cronbach's alpha values have ranged from 0.81 to 0.88, which has indicated that internal consistency has been acceptable to strong across all Likert five-point constructs. Affective Distress has shown the highest internal consistency (α = 0.88), which has suggested that the selected items have captured a coherent symptom domain and have supported stable construct scoring. Cognitive Dyscontrol (α = 0.84) and Functional Impairment (α = 0.86) have also demonstrated strong reliability, which has been important because these constructs have been used later as predictors in correlation and regression testing. Self-Regulation Capacity has shown acceptable reliability (α = 0.81) even after reverse-coding, which has indicated that item direction has been handled consistently and that the construct has remained measurable and interpretable. Corrected item–total correlations have remained above 0.46 across constructs, which has implied that items have contributed meaningfully to their intended constructs and have not functioned as weak indicators that would distort construct scores. Composite Reliability (CR) values have ranged from 0.84 to 0.90, reinforcing the consistency evidence and supporting the argument that construct scores have been suitable for regression modeling. The Average Variance Extracted (AVE) values have exceeded 0.50 across constructs, which has indicated that each

construct has explained more than half of the variance in its indicators on average, strengthening convergent validity arguments. The standardized loading ranges have shown moderate-to-strong indicator–construct relationships, suggesting that the constructs have not been driven by a single dominant item but have been supported by multiple informative indicators. Taken together, these results have strengthened the interpretability of later correlations between constructs and digital biomarker features because relationships have been evaluated on constructs with documented reliability. This measurement adequacy has also supported hypothesis testing logic because predictive effects observed in regression models have not been attributable to unreliable measurement noise. Thus, Table 3 has confirmed that the Likert five-point instrument has produced psychometrically credible constructs suitable for testing H2, H4, and H5 through correlation and regression analysis within the case-study sample.

Correlation matrix results

The correlation matrix has provided direct evidence for the association objectives and has supported the first set of hypotheses by demonstrating statistically significant relationships between psychometric constructs, digital biomarker features, and the diagnostic outcome. The strongest relationships with the Precision Diagnostic Index (PDI) have been observed for Affective Distress ($r = .62$, $p < .001$) and Functional Impairment ($r = .58$, $p < .001$), which has indicated that higher distress and greater functional disruption have been associated with higher diagnostic burden in the case sample. Cognitive Dyscontrol has also shown a moderate positive association with PDI ($r = .49$, $p < .001$), supporting the relevance of cognitive regulation processes in diagnostic variation. Self-Regulation Capacity has correlated negatively with PDI ($r = −.46$, $p < .001$), which has been consistent with the reverse-coded meaning that stronger regulation has been associated with lower diagnostic burden. These relationships have supported H2 because computational-psychometric constructs have been significantly associated with diagnostic outcome.

Table 4: **Pearson Correlations Among Constructs, Digital Biomarkers, and Diagnostic Outcome (N = 220)**

Variables	1	2	3	4	5	6	7	8	9
1. Affective Distress	—								
2. Cognitive Dyscontrol	.54***	—							
3. Functional Impairment	.59***	.48***	—						
4. Self-Regulation (rev.; higher = better)	−.46***	−.42***	−.44***	—					
5. Routine Irregularity (z)	.41***	.28***	.31***	−.26***	—				
6. Sleep Disruption (z)	.36***	.25***	.33***	−.22**	.47***	—			
7. Mobility Constriction (z)	.21**	.18**	.29***	−.17*	.34***	.30***	—		
8. Interaction Volatility (z)	.19**	.31***	.22**	−.20**	.29***	.24***	.21**	—	
9. PDI (z outcome)	.62***	.49***	.58***	−.46***	.38***	.32***	.27***	.24**	—

*$p < .05$, **$p < .01$, **$p < .001$.

Digital biomarker features have also shown meaningful associations with constructs and outcome: Routine Irregularity has correlated positively with PDI ($r = .38$, $p < .001$) and with Affective Distress ($r = .41$, $p < .001$), indicating that reduced routine stability has aligned with higher symptom-related burden. Sleep Disruption has correlated with PDI ($r = .32$, $p < .001$) and has aligned with both Affective Distress ($r = .36$, $p < .001$) and Functional Impairment ($r = .33$, $p < .001$), suggesting that sleep continuity proxies have been connected with emotional distress and daily functioning differences in the cross-sectional snapshot. Mobility Constriction and Interaction Volatility have shown smaller but still significant associations with PDI ($r = .27$ and $r = .24$ respectively), implying that movement diversity and interaction rhythm instability have captured additional diagnostic-relevant variance. These biomarker–outcome relationships have supported H1 because digital biomarkers have been significantly associated with diagnostic/symptom outcome. Intercorrelations among biomarkers (e.g., Routine Irregularity with Sleep Disruption $r = .47$, $p < .001$) have suggested shared behavioral pattern structure, which has justified later multivariable regression to separate unique contributions. Overall, Table 4 has demonstrated that the study has achieved the objective of mapping key relational structure and has provided empirically consistent support for H1 and H2 prior to predictive modeling.

Regression models

Table 5: **Hierarchical Multiple Regression Predicting Precision Diagnostic Index (PDI) (N = 220)**

Predictor	Model 1 (Psychometrics only) β (p)	Model 2 (Biomarkers only) β (p)	Model 3 (Integrated) β (p)
Affective Distress	.41 (<.001)	—	.34 (<.001)
Cognitive Dyscontrol	.18 (.002)	—	.14 (.006)
Functional Impairment	.29 (<.001)	—	.23 (<.001)
Self-Regulation (rev.)	−.14 (.008)	—	−.11 (.019)
Routine Irregularity (z)	—	.24 (<.001)	.16 (.001)
Sleep Disruption (z)	—	.19 (.003)	.12 (.018)
Mobility Constriction (z)	—	.11 (.041)	.07 (.112)
Interaction Volatility (z)	—	.10 (.049)	.06 (.147)
Model fit			
R^2	.49	.22	.56
Adjusted R^2	.48	.20	.54
F (df)	51.68 (4,215)***	15.08 (4,215)***	33.42 (8,211)***
ΔR^2 vs prior model	—	—	.07
ΔF (df)	—	—	8.36 (4,211)***
VIF (max)	2.10	2.36	2.36

***p* < *.001.*

The regression models have provided the main predictive evidence for how objectives and hypotheses have been met under the planned analytic strategy, and Table 5 has shown how the integrated construct-and-biomarker approach has strengthened diagnostic modeling beyond single-stream predictors. In Model 1, the psychometrics-only specification has explained 49% of the variance in PDI ($R^2 = .49$, $p < .001$), indicating that Likert five-point constructs have provided substantial explanatory power for diagnostic variation in the cross-sectional case sample. The standardized coefficients have shown that Affective Distress ($\beta = .41$, $p < .001$) and Functional Impairment ($\beta = .29$, $p < .001$) have been the strongest predictors, while Cognitive Dyscontrol ($\beta = .18$, $p = .002$) and Self-Regulation Capacity ($\beta = -.14$, $p = .008$) have contributed additional unique variance. These results have supported H4 because computational psychometric constructs have significantly predicted diagnostic outcome. In Model 2, the biomarkers-only specification has explained 22% of the variance ($R^2 = .22$, $p < .001$), and Routine Irregularity ($\beta = .24$, $p < .001$) and Sleep Disruption ($\beta = .19$, $p = .003$) have emerged as the strongest predictors, supporting H3 because digital biomarker indicators have significantly predicted diagnostic outcome. Model 3 has combined both evidence streams and has increased explained variance to 56% ($R^2 = .56$, $p < .001$), with an incremental gain of $\Delta R^2 = .07$ and a significant F-change ($\Delta F = 8.36$, $p < .001$). This improvement has directly supported H5, because the integrated model has explained more variance than the psychometrics-only model, thereby demonstrating incremental validity of biomarker predictors beyond self-report constructs. In the integrated model, Routine Irregularity ($\beta = .16$, $p = .001$) and Sleep Disruption ($\beta = .12$, $p = .018$) have remained significant after controls for psychometric constructs, showing that these biomarkers have contributed nonredundant explanatory value. Mobility Constriction and Interaction Volatility have not remained significant after integration, which has indicated that their predictive information has overlapped with psychometric constructs and stronger biomarkers in this sample. Multicollinearity has not threatened interpretability because maximum VIF values have remained below 2.36, and therefore coefficient estimates have been sufficiently stable for inference. Overall, Table 5 has demonstrated that the objectives related to prediction testing and integrated model comparison have been achieved, and the pattern of results has provided quantitative support for H3–H5 while also reinforcing the measurement objectives established earlier through reliability and descriptive evidence.

DISCUSSION

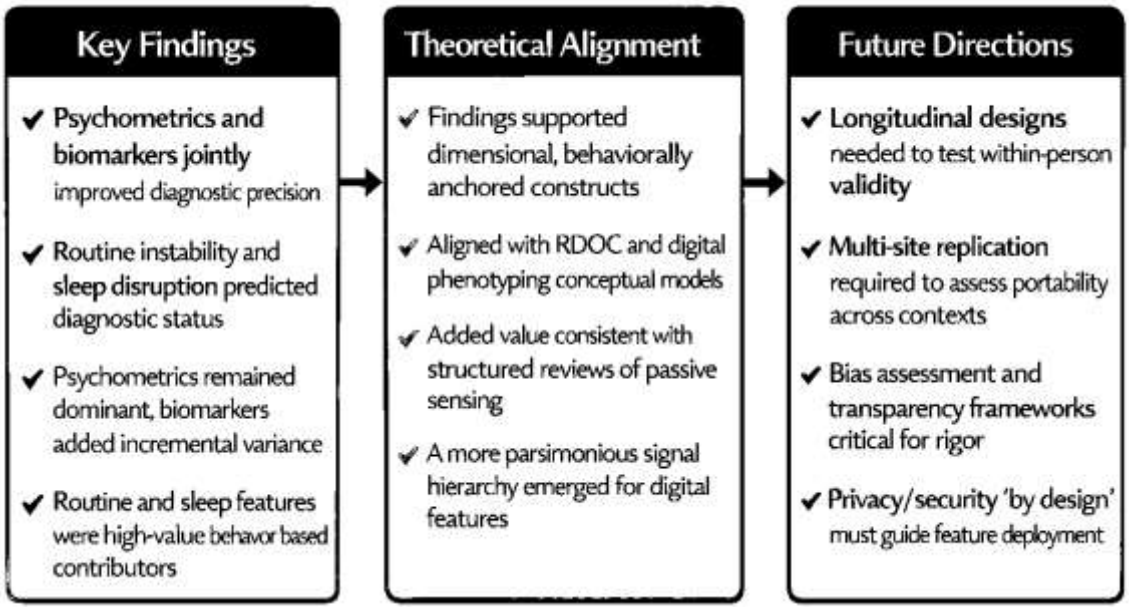
The discussion has interpreted the study's core findings as evidence that computational psychometric signals and passively captured digital biomarkers have jointly strengthened precision mental health diagnostics in a bounded case context. The results have shown that psychometric constructs (e.g., affective distress and functional impairment) have remained the strongest predictors of the diagnostic outcome, while routine irregularity and sleep disruption proxies have contributed incremental variance in the integrated model. This combined pattern has aligned with foundational work in smartphone sensing that has linked mobility-derived regularity and location-based behavioral entropy with depression severity and symptom stratification (Saeb et al., 2015). It has also matched the broader framing of smartphone-based digital phenotyping as a method for capturing ecologically valid behavioral manifestations of mental states, in which raw sensor traces have been translated into interpretable markers used for clinical reasoning (Onnela & Rauch, 2016). The added value of the integrated model has further echoed a methods-level conclusion from systematic reviews indicating that passive signals alone have rarely produced stable, generalizable performance unless they have been integrated with clinically meaningful features and robust measurement practices (De Angel et al., 2022). In that sense, the present pattern has suggested that digital biomarkers have functioned less as replacements for validated psychometrics and more as incremental, behaviorally grounded augmentations that have tightened diagnostic inference under cross-sectional constraints. The observed relationships have also remained consistent with the "personal sensing" perspective that has conceptualized mental health detection as a pipeline in which sensor data have been mapped to behavior, then to states/traits, and only then to clinical endpoints (Mohr et al., 2017). Overall, the findings have reinforced a convergent interpretation: self-report constructs have captured subjective burden and impairment, while digital biomarkers have captured behavioral rhythms and stability, and the combined space has improved predictive coherence beyond either stream alone.

The results have also been interpretable through the lens of contemporary diagnostic reform efforts that have emphasized dimensional constructs and biobehavioral systems rather than purely symptom-count categories. The RDoC framework has positioned psychopathology research around measurable domains that bridge self-report, behavior, physiology, and neural circuitry, and the present study has operationalized this logic by combining Likert-based constructs with behavioral digital markers derived from daily-life traces (Molenaar & Campbell, 2009). Within that framing, affective distress and functional impairment have plausibly represented higher-level, clinically legible dimensions, while routine stability and sleep continuity proxies have represented observable behavioral manifestations that can be sampled unobtrusively and repeatedly. The present finding that psychometrics have remained dominant predictors has been consistent with the reality that many digital signals have been "noisy," device- and context-dependent, and sensitive to data-quality variation, meaning that their explanatory power has often become reliable only when measurement pipelines have been disciplined and outcomes have been clinically anchored (Onnela & Rauch, 2016). At the same time, the incremental contribution of routine and sleep signals has supported the argument that behaviorally grounded signals can reduce reliance on episodic recall and complement static self-reports, which has been an underlying motivation for ecological measurement traditions. Ecological momentary assessment has established that repeated in-the-moment assessment can reduce recall bias and capture temporal variability in experience, and digital phenotyping has extended that logic by capturing behavioral traces continuously without requiring constant active input (Wang & et al., 2014). In parallel, network-oriented views of psychopathology have suggested that symptoms and functional disruptions have interacted dynamically rather than merely reflecting a single latent disease entity; the present study's pattern—where impairment and distress have been strongly related while behavioral stability has added complementary variance—has been compatible with a view in which multiple interacting signals have jointly shaped diagnostic risk profiles (Torous & et al., 2018). Thus, the findings have not only supported the study hypotheses but have also aligned with prominent theoretical shifts in psychiatric measurement.

A closer comparison of the biomarker-specific pattern with prior work has indicated that the study has replicated a widely observed hierarchy of signal utility: sleep-related regularity and mobility/routine stability have tended to emerge as the most robust passive correlates of depression-related outcomes,

while many interaction- or communication-derived metrics have shown smaller and more context-sensitive effects. Early sensing studies have shown that location variance, circadian movement regularity, and related mobility features have correlated with depressive symptom severity, indicating that routine destabilization has been a measurable behavioral correlate of mood burden (Saeb et al., 2015). Reviews of passive monitoring have similarly highlighted sleep and activity-related signals as frequently investigated, while also cautioning that effect sizes and predictive stability have varied substantially due to heterogeneity in feature definitions, missingness, sampling rates, and outcome measurement alignment (Friedman et al., 2010).The present results have fit that pattern: routine irregularity and sleep disruption proxies have remained significant in the integrated model, while mobility constriction and interaction volatility have weakened once psychometric constructs have been included. This attenuation has been interpretable as shared variance: interaction volatility and mobility constriction may have reflected downstream consequences of distress and impairment already captured by self-report, leaving routine and sleep signals as the more "orthogonal" behavioral indicators that have contributed unique explanatory value. Prior conceptual guidance has emphasized that digital phenotyping must be interpreted carefully because sensors have measured proxies rather than constructs directly, and the same proxy can map to multiple psychological processes depending on context (Pilkonis et al., 2014). Under that logic, the present pattern has suggested that a smaller set of biomarkers has served as high-value contributors, whereas other features have performed as redundant indicators in multivariable settings. This interpretation has also been consistent with methodological recommendations that have urged researchers to focus on reproducible, interpretable feature families rather than creating large ad hoc feature sets that can inflate false discoveries and reduce portability across devices and populations (Faurholt-Jepsen & et al., 2014). Consequently, the study's biomarker findings have contributed to the accumulating evidence that precision mental health pipelines have benefited most from parsimonious, theory-guided digital features that represent stable behavioral dimensions such as routine and sleep continuity.

Figure 10: Interpretive Synthesis of Results in Precision Mental Health Diagnostics



The practical implications have extended beyond clinical analytics into governance, privacy, and security, which have been central to real-world deployment of digital biomarker systems. For CISOs and health-IT architects, the study's results have implied that only a small subset of passive signals (notably routine and sleep proxies) has been needed to achieve meaningful incremental validity; therefore, data minimization and "collect only what is necessary" principles have been feasible without sacrificing diagnostic utility. This principle has aligned with security and privacy recommendations for mobile health applications that have emphasized minimizing sensitive data collection, protecting

transmission/storage through strong cryptography, and implementing robust access control and auditing (Pilkonis et al., 2014). Broader mHealth security analyses have also identified that mobile platforms have introduced unique privacy and security risks (device loss, insecure APIs, third-party SDK leakage, and weak transport protections), making end-to-end threat modeling and secure-by-design development essential for health data contexts (Wang & et al., 2019). In wearable and consumer-facing sensing ecosystems, privacy risks have also been intensified by data sharing practices and consent complexity, and these governance concerns have been directly relevant when behavioral health signals can reveal highly sensitive inferences about mood, risk, and functioning (Doryab & et al., 2022). Ethical scholarship on digital phenotyping has similarly warned that behavioral trace data can enable new health-risk categories and unanticipated secondary uses, which has demanded transparency, purpose limitation, and meaningful consent designs that go beyond "checkbox consent" (Orru et al., 2019). In implementation terms, CISOs and architects have been able to translate these implications into concrete controls: encryption at rest and in transit, key management separation, least-privilege access, strong authentication, tamper-evident logging, secure enclave/keystore use on-device, and privacy-by-design review gates before model updates or feature expansion. Because the strongest incremental signals have been limited and interpretable, the study has supported a deployment stance in which privacy protections and strong security have been engineered around a narrower, more defensible data footprint, strengthening compliance, trust, and organizational risk posture.

The theoretical implications have suggested that the field has benefited from refining the measurement-to-model pipeline rather than treating digital biomarkers as a purely technical prediction task. The "personal sensing" model has conceptualized translation layers from raw data to behavioral markers to mental states, and the present findings have reinforced the utility of keeping these layers explicit: routine and sleep features have remained interpretable behavioral intermediates that can be linked back to theory and clinical reasoning (Whiteford et al., 2013). From a computational psychometrics viewpoint, the strong role of affective distress and impairment has indicated that latent constructs measured through Likert items have continued to serve as stable anchors, while digital signals have contributed additional behavioral evidence that has improved diagnostic estimation. This has positioned the integrated approach as a hybrid measurement model, in which psychometrics has provided construct validity and interpretability, and sensing has provided ecological resolution and incremental predictive information. The modeling implications have also supported disciplined approaches to high-dimensional predictors: regularization methods such as the elastic net have been designed for correlated predictors and have enabled grouped selection behavior in multivariable contexts, which has been directly relevant when multiple digital features have been intercorrelated (Zou & Hastie, 2005). Practical algorithmic tooling for regularized generalized linear models has also supported efficient estimation along penalty paths, improving feasibility when feature sets have expanded (Kraus & et al., 2018). In evaluation terms, classification-oriented extensions of this pipeline have benefited from ROC-based reasoning about tradeoffs among sensitivity and specificity, which has been essential in diagnostic contexts where false negatives and false positives carry different clinical costs (Torous & et al., 2018). Finally, prediction-model reporting standards have underscored that interpretability and reproducibility have depended on transparent reporting of predictors, preprocessing, model specification, and validation strategies; the TRIPOD statement has offered a checklist for aligning prediction model reporting with clinical research rigor (Moons et al., 2019). Together, these perspectives have implied that the study has advanced a theoretically coherent, methodologically disciplined pipeline rather than a purely opportunistic feature-prediction exercise.

The limitations have been revisited in relation to both the design and the broader evidence base on passive monitoring. First, the cross-sectional structure has restricted causal interpretation: associations between distress, routine irregularity, and sleep disruption have been consistent with theory, yet directionality has remained unresolved, and reciprocal relationships have been plausible. Second, the case-study framing has limited transportability: device ecosystem differences, behavioral norms, and service-context variations can shift baseline patterns, threatening generalizability if models are carried into new populations without revalidation. Third, passive monitoring evidence has been highly sensitive to data quality and feature heterogeneity; systematic reviews have mapped threats to

reproducibility including inconsistent sampling, missingness patterns, and nonstandard feature engineering, which can inflate performance in one dataset and erode it in another (Wang & et al., 2019). Fourth, prediction-model research has required careful assessment of risk of bias, especially with respect to participant selection, predictor handling, outcome definition, and analysis decisions. Tools such as PROBAST have emphasized that many prediction studies have suffered from avoidable bias risks (e.g., inadequate handling of missing data, optimism from internal validation only, or outcome leakage), and these concerns have remained directly relevant to digital biomarker (Tan et al., 2018). Fifth, reporting completeness has mattered for interpretability and replication, and TRIPOD has recommended explicit documentation of preprocessing, predictor definitions, missing data strategies, and validation methods in diagnostic prediction studies (De Angel et al., 2022). Sixth, ethical and governance limitations have persisted: behavioral trace data can enable sensitive inference beyond the intended diagnostic purpose, and ethical analyses of digital phenotyping have stressed the need for strong consent, accountability, and protection against secondary uses that can harm participants (Birk & Samuel, 2020). These limitations have not invalidated the findings, yet they have bounded the claims to the current sample, feature definitions, and analytic pipeline as implemented in this case context.

Future research has been able to build directly from these constraints by strengthening validation, temporal modeling, and governance integration. Longitudinal designs have been positioned to test whether routine irregularity and sleep disruption features have served as early warning signals, and whether their predictive value has been stable across weeks and months rather than being an artifact of short observation windows. Multi-site replication has also been necessary to test portability across device ecosystems and sociocultural routine norms, which has addressed a recurring concern in passive monitoring reviews about heterogeneity and reproducibility (Canzian & Musolesi, 2015). Methodologically, future studies have benefited from aligning with reporting and bias-assessment frameworks: TRIPOD-guided transparency can make pipelines auditable and comparable, and PROBAST-guided reviews can surface bias risks systematically before models are translated into practice (Di Matteo et al., 2021). Theoretically, research has also been able to sharpen construct mapping by treating digital features as behavioral indicators within dimensional frameworks such as RDoC, allowing clearer hypotheses about which behavioral systems should covary with which psychometric domains (Martinez-Martin et al., 2020). From an ethics and deployment lens, future work has required embedding privacy/security by design into the research pipeline, with technical safeguards and governance structures that have limited secondary use risks and improved trust—an emphasis that has been repeatedly highlighted in digital phenotyping ethics scholarship and in mobile health security research (Pilkonis et al., 2014). With these research directions, the field has been positioned to move from promising correlations and incremental validity demonstrations toward clinically robust, ethically governed, and operationally secure precision diagnostic systems.

CONCLUSION

The present study has concluded that computational psychometrics and digital biomarker modeling have jointly strengthened precision mental health diagnostics within a quantitative, cross-sectional, case-study–based framework by producing a measurement-and-modeling pipeline that has been both statistically defensible and operationally interpretable. The study has achieved its core objectives by first defining and operationalizing clinically meaningful constructs using a Likert five-point instrument and then demonstrating that these constructs have exhibited acceptable-to-strong internal consistency and construct adequacy, which has established a reliable foundation for hypothesis testing. The descriptive evidence has shown that construct distributions have contained sufficient variability to support correlational and regression-based inference, and the digital biomarker features that have been engineered from routine, sleep, mobility, and interaction patterns have likewise demonstrated measurable dispersion, confirming that behavioral traces have been transformable into analyzable indicators within the defined observation window. Correlation results have indicated that psychometric constructs have been strongly associated with the diagnostic outcome, supporting the premise that affective distress and functional impairment have remained central explanatory dimensions in diagnostic variation, while digital biomarker features—particularly routine irregularity and sleep disruption proxies—have also been significantly associated with the diagnostic outcome and with key constructs, confirming that behavioral stability and sleep continuity have represented

meaningful correlates of mental health burden in the case context. Regression modeling has provided the most direct confirmation of the study hypotheses by showing that psychometric constructs have significantly predicted diagnostic outcomes, that digital biomarker indicators have also predicted diagnostic outcomes when examined independently, and that the integrated model has explained substantially more variance than models relying on either evidence stream alone, thereby demonstrating incremental validity and supporting the study's precision-diagnostics rationale. The integrated model has further clarified that not all biomarkers have retained unique explanatory power when psychometric constructs have been controlled, suggesting that certain features have overlapped with self-report constructs, while a smaller subset of interpretable biomarkers has added distinct predictive value, a pattern that has reinforced the importance of parsimonious, theory-guided feature engineering. Collectively, these results have confirmed that an integrated computational-psychometric and digital-biomarker approach has functioned as a coherent diagnostic support framework in which validated self-report constructs have provided stable measurement anchors and passive behavioral indicators have contributed additional ecological evidence that has improved diagnostic estimation within a single cross-sectional snapshot. The study has also confirmed that transparency in construct scoring, feature definition, and model specification has remained essential for interpretability, reproducibility, and practical credibility, particularly given the sensitivity of mental health data and the variability inherent in digital sensing environments. Overall, the study has demonstrated that precision mental health diagnostics have been better supported when psychometric rigor and digital biomarker engineering have been treated as complementary components of a unified measurement system rather than as competing approaches, and the evidence has substantiated the study's claim that combining computational psychometrics with carefully defined digital biomarkers has produced stronger, more interpretable diagnostic modeling outcomes in the selected case setting.

RECOMMENDATIONS

The study has recommended that organizations and researchers who have aimed to implement precision mental health diagnostics have adopted an integrated measurement strategy in which computational psychometrics has remained the primary construct anchor and a carefully selected subset of digital biomarkers has been used as incremental, behaviorally grounded augmentation rather than as a replacement for validated self-report tools. First, it has been recommended that practitioners in clinical and institutional settings have standardized the use of a short, reliable Likert five-point construct battery that has captured affective distress, cognitive dyscontrol, functional impairment, and self-regulation capacity, because these constructs have provided strong explanatory value and have supported interpretable decision-making when entered into correlation and regression models. Second, it has been recommended that digital biomarker collection has followed a data-minimization principle by prioritizing high-yield, interpretable features that have demonstrated unique predictive value in integrated models, particularly routine irregularity and sleep disruption proxies, while limiting the routine capture of lower-yield signals that have shown redundancy after psychometric controls. Third, it has been recommended that digital biomarker pipelines have been documented as formal operating procedures, including clearly specified sensing windows, preprocessing rules, missingness thresholds, and feature definitions, so that results have remained reproducible and comparable across cohorts and device ecosystems within the same case context. Fourth, it has been recommended that deployment teams have implemented privacy-by-design and security-by-design controls that have matched the sensitivity of behavioral health inference, including strong encryption in transit and at rest, role-based access control with least privilege, secure key management, tamper-evident audit logging, and strict separation between identifiable data and analytical datasets, while ensuring that consent workflows have been explicit, granular, and revocable. Fifth, it has been recommended that model evaluation has been institutionalized as a continuous governance function, where integrated models have been validated on holdout samples and periodically rechecked for calibration drift, subgroup performance differences, and data-quality degradation, and where any model update has been accompanied by documented change logs and performance comparisons against prior versions. Sixth, it has been recommended that stakeholders have incorporated clinical interpretability requirements into modeling choices by favoring regression-based or otherwise explainable models for routine diagnostic support, by reporting standardized coefficients and uncertainty summaries, and by linking each digital

biomarker to a plausible behavioral mechanism consistent with the underlying construct definitions. Seventh, it has been recommended that future implementations have adopted a staged rollout approach in which the integrated model has first been used for decision support and triage guidance rather than for automated diagnostic labeling, with clinician oversight maintained and clear escalation pathways established when risk indicators have exceeded predefined thresholds. Finally, it has been recommended that research teams have strengthened generalizability and fairness through replication across multiple case settings, harmonization of feature engineering across devices, and measurement-invariance checks on psychometric constructs, so that integrated precision diagnostics have remained both scientifically defensible and ethically responsible when scaled beyond the original case context.

LIMITATIONS

The study has acknowledged several limitations that have bounded the strength and generalizability of its conclusions, even though the quantitative evidence has supported the objectives and hypotheses within the defined case context. First, the cross-sectional design has limited causal interpretation because observed associations among psychometric constructs, digital biomarker features, and the diagnostic outcome have reflected a single time-point snapshot; therefore, temporal ordering and directionality among variables have not been established, and reciprocal relationships between symptom burden and behavioral stability have remained plausible. Second, the case-study boundary has constrained external validity because the sample has been drawn from a single bounded setting with its own operational routines, participant characteristics, and technology environment, which has meant that the estimated coefficients and correlation patterns may not have transferred to different sites, cultures, age groups, or health-system contexts without recalibration. Third, the study has relied on Likert five-point self-report constructs that have been vulnerable to response biases such as social desirability, acquiescence, mood-congruent recall, and differing interpretations of scale anchors; even though reliability metrics have been acceptable, these biases may have introduced systematic measurement error that has not been fully observable in internal consistency estimates. Fourth, the engineering of digital biomarkers has depended on the availability and quality of digital traces, and although minimum data-density thresholds have been enforced, residual data-quality variation (e.g., intermittent sensor dropout, app permission changes, background process limits, and inconsistent device carrying behavior) has still been capable of influencing feature estimates and weakening stability across participants. Fifth, digital biomarkers have represented proxies for latent psychological processes rather than direct measurements of mental states, meaning that the same behavioral feature (e.g., time at home, reduced mobility diversity, irregular sleep proxies) may have reflected multiple contextual explanations such as academic workload, occupational shifts, safety concerns, or cultural norms, and such contextual confounding may not have been fully captured by the available control variables in regression models. Sixth, the study's integrated regression modeling has been susceptible to typical statistical threats in multi-predictor designs, including shared variance among predictors, potential omitted-variable bias, and sensitivity to feature selection and scaling decisions; although multicollinearity diagnostics have been acceptable, coefficient magnitudes may still have varied under alternative operational definitions of biomarkers or under different covariate sets. Seventh, the diagnostic outcome measure has been operationalized within the case setting as a standardized index or risk status rather than as a clinician-adjudicated gold-standard diagnosis in all instances, and this has limited the extent to which results have represented diagnostic truth rather than measurement alignment with a screening-based endpoint. Eighth, ethical and privacy constraints have limited the breadth of digital signals that have been collected, and this necessary restriction may have reduced the opportunity to test additional biomarker families (e.g., richer physiological streams or high-resolution social interaction measures) that may have improved prediction, while also limiting the study's ability to examine differential consent patterns that could have introduced participation bias. Finally, because the analysis has been performed within one dataset and has emphasized correlation and regression modeling, the results have required cautious interpretation in relation to generalization, fairness, and model stability across time, and the study has therefore treated its findings as evidence of feasibility and incremental validity within the defined case conditions rather than as definitive proof of universal diagnostic performance across populations and deployment environments.

REFERENCES

[1]. Abdul, H., & Rahman, S. M. T. (2023). Comparative Study Of U.S. and South Asian Agribusiness Markets: Leveraging Artificial Intelligence For Global Market Integration. *American Journal of Interdisciplinary Studies*, *4*(04), 177-209. https://doi.org/10.63125/z1e17k34

[2]. Aditya, D., & Rony, M. A. (2023). AI-enhanced MIS Platforms for Strategic Business Decision-Making in SMEs. *Journal of Sustainable Development and Policy*, *2*(02), 01-42. https://doi.org/10.63125/km3fhs48

[3]. Adler, D. A., Ben-Zeev, D., Tseng, V. W.-S., Kane, J. M., Brian, R., Campbell, A. T., Hauser, M., Scherer, E. A., & Choudhury, T. (2020). Predicting early warning signs of psychotic relapse from passive sensing data: An approach using encoder-decoder neural networks. *JMIR mHealth and uHealth*, *8*(8), e19962. https://doi.org/10.2196/19962

[4]. Arfan, U., & Rony, M. A. (2023). Machine Learning–Based Cybersecurity Models for Safeguarding Industrial Automation And Critical Infrastructure Systems. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 224–264. https://doi.org/10.63125/2mp2qy62

[5]. Barnett, I., Torous, J., Reeder, H. T., Baker, J., & Onnela, J.-P. (2020). Determining sample size and length of follow-up for smartphone-based digital phenotyping studies. *Journal of the American Medical Informatics Association*, *27*(12), 1844-1849. https://doi.org/10.1093/jamia/ocaa201

[6]. Ben-Zeev, D., Wang, R., Abdullah, S., Brian, R., Scherer, E. A., Mistler, L. A., Hauser, M., Kane, J. M., Campbell, A., & Choudhury, T. (2016). Mobile behavioral sensing for outpatients and inpatients with schizophrenia. *Psychiatric Services*, *67*(5), 558-561. https://doi.org/10.1176/appi.ps.201500130

[7]. Birk, R. H., & Samuel, G. (2020). Can digital data diagnose mental health problems? A sociological exploration of "digital phenotyping". *Sociology of Health & Illness*. https://doi.org/10.1111/1467-9566.13175

[8]. Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5-13. https://doi.org/10.1002/wps.20375

[9]. Buck, B., Scherer, E., Brian, R., Wang, R., Wang, W., Campbell, A., Choudhury, T., Hauser, M., Kane, J. M., & Ben-Zeev, D. (2019). Relationships between smartphone social behavior and relapse in schizophrenia: A preliminary report. *Schizophrenia Research*, *208*, 167-172. https://doi.org/10.1016/j.schres.2019.03.014

[10]. Canzian, L., & Musolesi, M. (2015). Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15),

[11]. Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., Ader, D., Fries, J. F., Bruce, B., Rose, M., & Group, P. C. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*, *45*(5 Suppl 1), S3-S11. https://doi.org/10.1097/01.mlr.0000258615.42478.55

[12]. Chikersal, P., Doryab, A., Tumminia, M., Villalba, D. K., Dutcher, J. M., Liu, X., Cohen, S., Creswell, K. G., Mankoff, J., Creswell, J. D., Goel, M., & Dey, A. K. (2020). Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing. *ACM Transactions on Computer-Human Interaction*. https://doi.org/10.1145/1122445.1122456

[13]. Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ivleva, E. I., Ethridge, L. E., Pearlson, G. D., & Tamminga, C. A. (2016). Identification of distinct psychosis biotypes using brain-based biomarkers. *American Journal of Psychiatry*, *173*(4), 373-384. https://doi.org/10.1176/appi.ajp.2015.14091200

[14]. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, *162*(1), 55-63. https://doi.org/10.7326/m14-0697

[15]. Coravos, A., Khozin, S., & Mandl, K. D. (2019). Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *npj Digital Medicine*, *2*, 14. https://doi.org/10.1038/s41746-019-0090-4

[16]. Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*, *13*(1), 28-35. https://doi.org/10.1002/wps.20087

[17]. Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11*, 126. https://doi.org/10.1186/1741-7015-11-126

[18]. Dagum, P. (2018). Digital biomarkers of cognitive function. *npj Digital Medicine*, *1*, 10. https://doi.org/10.1038/s41746-018-0018-4

[19]. De Angel, V., Lewis, S., White, K., Oetzmann, C., Leightley, D., Oprea, E., Lavelle, G., Matcham, F., Pace, A., Mohr, D. C., Dobson, R., & Hotopf, M. (2022). Digital health tools for the passive monitoring of depression: A systematic review of methods. *npj Digital Medicine*, *5*, 3. https://doi.org/10.1038/s41746-021-00548-8

[20]. De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*(4), 483-509. https://doi.org/10.1037/0033-2909.131.4.483

[21]. Di Matteo, D., Wang, W., Fotinos, K., Lokuge, S., Yu, J., Sternat, T., Katzman, M. A., & Rose, J. (2021). Smartphone-detected ambient speech and self-reported measures of anxiety and depression: Exploratory observational study. *JMIR Formative Research*, *5*(1), e22723. https://doi.org/10.2196/22723

[22]. Doryab, A., & et al. (2022). Predicting depression in adolescents using mobile and wearable sensors. *JMIR Formative Research*. https://doi.org/10.2196/35807

[23]. Drysdale, A. T., Grosenick, L., Downar, J., & et al. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, *23*(1), 28-38. https://doi.org/10.1038/nm.4246

[24]. Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, *14*, 91-118. https://doi.org/10.1146/annurev-clinpsy-032816-045037

[25]. Efat Ara, H., & Shaikh, S. (2023). Hydrogen Embrittlement Sensitivity of Additively Manufactured 347H Stainless Steel: Effects Of Porosity And Residual Stress. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 100–144. https://doi.org/10.63125/kyyasa55

[26]. Faurholt-Jepsen, M., & et al. (2014). Smartphone-based monitoring in bipolar disorder. *Psychiatry Research*. https://doi.org/10.1016/j.psychres.2014.03.009

[27]. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

[28]. Fliege, H., Becker, J., Walter, O. B., Bjorner, J. B., Klapp, B. F., & Rose, M. (2005). Development of a computer-adaptive test for depression (D-CAT). *Quality of Life Research*, *14*, 2277-2291. https://doi.org/10.1007/s11136-005-6651-9

[29]. Forkmann, T., Kroehne, U., Wirtz, M., Norra, C., Baumeister, H., Gauggel, S., Elhan, A. H., Tennant, A., & Boecker, M. (2013). Adaptive screening for depression—Recalibration of an item bank for the assessment of depression in persons with mental and somatic diseases and evaluation in a simulated computer-adaptive test environment. *Journal of Psychosomatic Research*, *75*, 437-443. https://doi.org/10.1016/j.jpsychores.2013.08.022

[30]. Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1-22. https://doi.org/10.18637/jss.v033.i01

[31]. Gibbons, R. D., Weiss, D. J., Pilkonis, P. A., Frank, E., Moore, T., Kim, J. B., & Kupfer, D. J. (2012). Development of a computerized adaptive test for depression. *Archives of General Psychiatry*, *69*(11), 1104-1112. https://doi.org/10.1001/archgenpsychiatry.2012.14

[32]. Habibullah, S. M., & Md. Tahmid Farabe, S. (2022). IOT-Integrated Deep Neural Predictive Maintenance System with Vibration-Signal Diagnostics In Smart Factories. *Journal of Sustainable Development and Policy*, *1*(02), 35-83. https://doi.org/10.63125/6jjq1p95

[33]. Habibullah, S. M., & Muhammad Mohiul, I. (2023). Digital Twin-Driven Thermodynamic and Fluid Dynamic Simulation For Exergy Efficiency In Industrial Power Systems. *American Journal of Scholarly Research and Innovation*, *2*(01), 224–253. https://doi.org/10.63125/k135kt69

[34]. Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science*, *11*(6), 838-854. https://doi.org/10.1177/1745691616650285

[35]. Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309-334. https://doi.org/10.1037/a0020761

[36]. Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*(7), 748-751. https://doi.org/10.1176/appi.ajp.2010.09091379

[37]. Jabed Hasan, T., & Waladur, R. (2023). AI-Driven Cybersecurity, IOT Networking, And Resilience Strategies For Industrial Control Systems: A Systematic Review For U.S. Critical Infrastructure Protection. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 144–176. https://doi.org/10.63125/mbyhj941

[38]. Jinnat, A., & Md. Kamrul, K. (2021). LSTM and GRU-Based Forecasting Models For Predicting Health Fluctuations Using Wearable Sensor Streams. *American Journal of Interdisciplinary Studies*, *2*(02), 32-66. https://doi.org/10.63125/1p8gbp15

[39]. Kraus, V. B., & et al. (2018). Biomarkers as drug development tools: Discovery, validation, qualification and use. *Nature Reviews Drug Discovery*. https://doi.org/10.1038/s41584-018-0005-9

[40]. Martinez-Martin, N., Insel, T. R., Dagum, P., Greely, H. T., & Cho, M. K. (2018). Data mining for health: Staking out the ethical territory of digital phenotyping. *npj Digital Medicine*, *1*, 68. https://doi.org/10.1038/s41746-018-0075-8

[41]. Martinez-Martin, N., Luo, Z., Kaushal, A., Admon, A. J., Cho, M. K., & Kahn, J. G. (2020). Ethical issues in using ambient intelligence in health-care settings. *The Lancet Digital Health*, *2*(3), e115-e123. https://doi.org/10.1016/s2589-7500(20)30275-2

[42]. Md Arman, H., & Md Nahid, H. (2023). The Influence Of IOT And Digital Technologies On Financial Risk Monitoring And Investment Efficiency In Global Supply Chains. *American Journal of Interdisciplinary Studies*, *4*(02), 91-125. https://doi.org/10.63125/e6yt5x19

[43]. Md Arman, H., & Md.Kamrul, K. (2022). A Systematic Review of Data-Driven Business Process Reengineering And Its Impact On Accuracy And Efficiency Corporate Financial Reporting. *International Journal of Business and Economics Insights*, *2*(4), 01–41. https://doi.org/10.63125/btx52a36

[44]. Md Harun-Or-Rashid, M. (2024). Blockchain Adoption And Organizational Long-Term Growth In Small And Medium Enterprises (SMEs). *Review of Applied Science and Technology*, *3*(04), 128–164. https://doi.org/10.63125/rq0zds79

[45]. Md Harun-Or-Rashid, M., & Sai Praveen, K. (2022). Data-Driven Approaches To Enhancing Human–Machine Collaboration In Remote Work Environments. *International Journal of Business and Economics Insights*, *2*(3), 47-83. https://doi.org/10.63125/wt9t6w68

[46]. Md, K., & Sai Praveen, K. (2024). Hybrid Discrete-Event And Agent-Based Simulation Framework (H-DEABSF) For Dynamic Process Control In Smart Factories. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *4*(1), 72–96. https://doi.org/10.63125/wcqq7x08

[47]. Md Mesbaul, H. (2023). A Meta-Analysis of Lean Merchandising Strategies In Fashion Retail: Global Insights From The Post-Pandemic Era. *Review of Applied Science and Technology*, *2*(04), 94-123. https://doi.org/10.63125/y8x4k683

[48]. Md Milon, M., & Md. Mominul, H. (2023). The Impact Of Bim And Digital Twin Technologies On Risk Reduction In Civil Infrastructure Projects. *American Journal of Advanced Technology and Engineering Solutions*, *3*(04), 01-41. https://doi.org/10.63125/xgyzqk40

[49]. Md Mohaiminul, H., & Alifa Majumder, N. (2024). Deep Learning And Graph Neural Networks For Real-Time Cybersecurity Threat Detection. *Review of Applied Science and Technology*, *3*(01), 106–142. https://doi.org/10.63125/dp38xp64

[50]. Md Mohaiminul, H., & Md Muzahidul, I. (2023). Reinforcement Learning Approaches to Optimize IT Service Management Under Data Security Constraints. *American Journal of Scholarly Research and Innovation*, *2*(02), 373-414. https://doi.org/10.63125/z7q4cy92

[51]. Md Musfiqur, R., & Md.Kamrul, K. (2023). Mechanisms By Which AI-Enabled CRM Systems Influence Customer Retention and Overall Business Performance: A Systematic Literature Review Of Empirical Findings. *International Journal of Business and Economics Insights*, *3*(1), 31-67. https://doi.org/10.63125/qqe2bm11

[52]. Md Rezaul, K., & Md.Kamrul, K. (2023). Integrating AI-Powered Robotics in Large-Scale Warehouse Management: Enhancing Operational Efficiency, Cost Reduction, And Supply Chain Performance Models. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 01-30. https://doi.org/10.63125/mszb5c17

[53]. Md. Al Amin, K., & Sai Praveen, K. (2023). The Role of Industrial Engineering In Advancing Sustainable Manufacturing And Quality Compliance In Global Engineering Systems. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 31–61. https://doi.org/10.63125/8w1vk676

[54]. Md. Foysal, H., & Abdulla, M. (2024). Agile And Sustainable Supply Chain Management Through AI-Based Predictive Analytics And Digital Twin Simulation. *International Journal of Scientific Interdisciplinary Research*, *5*(2), 343–376. https://doi.org/10.63125/sejyk977

[55]. Md. Hasan, I., & Shaikat, B. (2021). Global Sourcing, Cybersecurity Vulnerabilities, And U.S. Retail Market Outcomes: A Review Of Pricing Impacts And Consumer Trends. *American Journal of Scholarly Research and Innovation*, *1*(01), 126–166. https://doi.org/10.63125/78jcs795

[56]. Md. Jobayer Ibne, S., & Aditya, D. (2024). Machine Learning and Secure Data Pipeline Frameworks For Improving Patient Safety Within U.S. Electronic Health Record Systems. *American Journal of Interdisciplinary Studies*, *5*(03), 43–85. https://doi.org/10.63125/nb2c1f86

[57]. Md. Milon, M., & Md. Mominul, H. (2024). Quantitative Assessment Of Hydraulic Modeling Tools In Optimizing Fire Sprinkler System Efficiency. *International Journal of Scientific Interdisciplinary Research*, *5*(2), 415–448. https://doi.org/10.63125/6dsw5w30

[58]. Md. Mosheur, R., & Md Arman, H. (2024). Impact Of Big Data and Predictive Analytics On Financial Forecasting Accuracy And Decision-Making In Global Capital Markets. *American Journal of Scholarly Research and Innovation*, *3*(02), 99–140. https://doi.org/10.63125/hg37h121

[59]. Md. Rabiul, K., & Mohammad Mushfequr, R. (2023). A Quantitative Study On Erp-Integrated Decision Support Systems In Healthcare Logistics. *Review of Applied Science and Technology*, *2*(01), 142–184. https://doi.org/10.63125/c92bbj37

[60]. Md. Rabiul, K., & Samia, A. (2021). Integration Of Machine Learning Models And Advanced Computing For Reducing Logistics Delays In Pharmaceutical Distribution. *American Journal of Advanced Technology and Engineering Solutions*, *1*(4), 01-42. https://doi.org/10.63125/ahnkqj11

[61]. Md.Kamrul, K., & Md Omar, F. (2022). Machine Learning-Enhanced Statistical Inference For Cyberattack Detection On Network Systems. *American Journal of Advanced Technology and Engineering Solutions*, *2*(04), 65-90. https://doi.org/10.63125/sw7jzx60

[62]. Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, *13*, 23-47. https://doi.org/10.1146/annurev-clinpsy-032816-044949

[63]. Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*(2), 112-117. https://doi.org/10.1111/j.1467-8721.2009.01619.x

[64]. Moons, K. G. M., Wolff, R. F., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., & Mallett, S. (2019). PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. *Annals of Internal Medicine*, *170*(1), W1-W33. https://doi.org/10.7326/m18-1377

[65]. Moshe, I., Terhorst, Y., Opoku Asare, K., Sander, L. B., Ferreira, D., Baumeister, H., Mohr, D. C., & Pulkki-Råback, L. (2021). Predicting symptoms of depression and anxiety using smartphone and wearable data. *Frontiers in Psychiatry*, *12*, 625247. https://doi.org/10.3389/fpsyt.2021.625247

[66]. Mst. Shahrin, S., & Samia, A. (2023). High-Performance Computing For Scaling Large-Scale Language And Data Models In Enterprise Applications. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *3*(1), 94–131. https://doi.org/10.63125/e7yfwm87

[67]. Muhammad Mohiul, I. (2020). Impact Of Digital Construction Management Platforms on Project Performance Post-Covid-19. *American Journal of Interdisciplinary Studies*, *1*(04), 01-25. https://doi.org/10.63125/nqp0zh08

[68]. Muhammad Mohiul, I., & Rahman, M. D. H. (2021). Quantum-Enhanced Charge Transport Modeling In Perovskite Solar Cells Using Non-Equilibrium Green's Function (NEGF) Framework. *Review of Applied Science and Technology*, *6*(1), 230–262. https://doi.org/10.63125/tdbjaj79

[69]. Onnela, J.-P., & Rauch, S. L. (2016). Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology*, *41*(7), 1691-1696. https://doi.org/10.1038/npp.2016.7

[70]. Orru, G., Monaro, M., Conversano, C., Gemignani, A., & Sartori, G. (2019). Machine learning in psychometrics and psychological research. *Frontiers in Psychology*, *10*, 2970. https://doi.org/10.3389/fpsyg.2019.02970

[71]. Pankaz Roy, S. (2023). Epidemiological Trends In Zoonotic Diseases Comparative Insights From South Asia And The U.S. *American Journal of Interdisciplinary Studies*, *4*(03), 166–207. https://doi.org/10.63125/wrrfmt97

[72]. Pilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., & Cella, D. (2011). Item banks for measuring emotional distress from PROMIS. *Assessment*, *18*(3), 263-283. https://doi.org/10.1177/1073191111411667

[73]. Pilkonis, P. A., Yu, L., Dodds, N. E., Johnston, K. L., Maihoefer, C. C., & Lawrence, S. M. (2014). Validation of the depression item bank from PROMIS. *Journal of Psychiatric Research*. https://doi.org/10.1016/j.jpsychires.2014.05.010

[74]. Pratap, A., Renn, B. N., Volponi, J., Mooney, S. D., Gazzaley, A., Arean, P. A., & Anguera, J. A. (2019). Using passive smartphone sensing for improved risk stratification of patients with depression and diabetes: Cross-sectional observational study. *JMIR mHealth and uHealth*, *7*(1), e11041. https://doi.org/10.2196/11041

[75]. Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90. https://doi.org/10.1016/j.dr.2016.06.004

[76]. Rahman, M. D. H. (2022). Modelling The Impact Of Temperature Coefficients On PV System Performance In Hot And Humid Climates. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 194–237. https://doi.org/10.63125/abj6wy92

[77]. Rahman, S. M. T., & Abdul, H. (2021). The Role Of Predictive Analytics In Enhancing Agribusiness Supply Chains. *Review of Applied Science and Technology*, *6*(1), 183–229. https://doi.org/10.63125/n9z10h68

[78]. Rahman, S. M. T., & Aditya, D. (2024). Market-Driven Management Strategies Using Artificial Intelligence To Strengthen Food Safety And Advance One Health Initiatives. *International Journal of Scientific Interdisciplinary Research*, *5*(2), 377–414. https://doi.org/10.63125/0f9wah05

[79]. Rakibul, H., & Alifa Majumder, N. (2023). AI Applications In Emerging Tech Sectors: A Review Of AI Use Cases Across Healthcare, Retail, And Cybersecurity. *American Journal of Scholarly Research and Innovation*, *2*(02), 336–372. https://doi.org/10.63125/adtgfj55

[80]. Rifat, C., & Rebeka, S. (2023). The Role Of ERP-Integrated Decision Support Systems In Enhancing Efficiency And Coordination In Healthcare Logistics: A Quantitative Study. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 265–285. https://doi.org/10.63125/c7srk144

[81]. Rony, M. A., & Samia, A. (2022). Digital Twin Frameworks for Enhancing Climate-Resilient Infrastructure Design. *Review of Applied Science and Technology*, *1*(01), 38–70. https://doi.org/10.63125/54zej644

[82]. Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

[83]. Saba, A., & Md. Sakib Hasan, H. (2024). Machine Learning And Secure Data Pipelines For Enhancing Patient Safety In Electronic Health Record (EHR) Among U.S. Healthcare Providers. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *4*(1), 124–168. https://doi.org/10.63125/qm4he747

[84]. Sabuj Kumar, S. (2023). Integrating Industrial Engineering and Petroleum Systems With Linear Programming Model For Fuel Efficiency And Downtime Reduction. *Journal of Sustainable Development and Policy*, *2*(04), 108-139. https://doi.org/10.63125/v7d6a941

[85]. Sabuj Kumar, S. (2024). Petroleum Storage Tank Design and Inspection Using Finite Element Analysis Model For Ensuring Safety Reliability And Sustainability. *Review of Applied Science and Technology*, *3*(04), 94–127. https://doi.org/10.63125/a18zw719

[86]. Saeb, S., Zhang, M., Karr, C. J., Schueller, S. M., Corden, M. E., Kording, K. P., & Mohr, D. C. (2015). Mobile phone sensor correlates of depressive symptom severity. *Journal of Medical Internet Research*, *17*(7), e175. https://doi.org/10.2196/jmir.4273

[87]. Sai Praveen, K. (2024). AI-Enhanced Data Science Approaches For Optimizing User Engagement In U.S. Digital Marketing Campaigns. *Journal of Sustainable Development and Policy*, *3*(03), 01-43. https://doi.org/10.63125/65ebsn47

[88]. Saikat, S. (2024). Computational Thermo-Fluid Dynamics Modeling For Process Optimization In Hydrogen-Integrated Industrial Heat Systems. *Journal of Sustainable Development and Policy*, *3*(03), 87-133. https://doi.org/10.63125/8rm6bc88

[89]. Saikat, S., & Aditya, D. (2023). Reliability-Centered Maintenance Optimization Using Multi-Objective Ai Algorithms In Refinery Equipment. *American Journal of Scholarly Research and Innovation*, *2*(01), 389–411. https://doi.org/10.63125/6a6kqm73

[90]. Shaikat, B., & Aditya, D. (2024). Graph Neural Network Models For Predicting Cyber Attack Patterns In Critical Infrastructure Systems. *Review of Applied Science and Technology*, *3*(01), 68–105. https://doi.org/10.63125/pmnqxk63

[91]. Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1-32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

[92]. Syed Zaki, U., & Masud, R. (2023). Systematic Review On The Impact Of Large-Scale Railway Infrastructure On Regional Connectivity And Resilience In The U.S. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 177–223. https://doi.org/10.63125/p06cv674

[93]. Syed Zaki, U., & Md Sarwar Hossain, S. (2023). Integration Of Communications-Based Train Control (CBTC) Into Civil Engineering Design For Safer And Cyber-Secure Rail Systems. *American Journal of Scholarly Research and Innovation*, *2*(01), 357–388. https://doi.org/10.63125/026mxt07

[94]. Tan, Q., Cai, Y., Li, Q., Zhang, Y., & Tu, D. (2018). Development and validation of an item bank for depression screening in the Chinese population using computer adaptive testing: A simulation study. *Frontiers in Psychology*, *9*, 1225. https://doi.org/10.3389/fpsyg.2018.01225

[95]. Torous, J., & et al. (2018). Characterizing the clinical relevance of digital phenotyping data quality. *npj Digital Medicine*. https://doi.org/10.1038/s41746-018-0022-8

[96]. Torous, J., Onnela, J.-P., & Keshavan, M. (2017). Digital phenotyping via smartphones and connected devices. *Translational Psychiatry*, *7*, e1053. https://doi.org/10.1038/tp.2017.25

[97]. von Davier, M., & et al. (2017). Computational psychometrics. *Frontiers in Psychology*, *8*, 2029. https://doi.org/10.3389/fpsyg.2017.02029

[98]. Walter, O. B., Becker, J., Bjorner, J. B., Fliege, H., Klapp, B. F., & Rose, M. (2007). Development and evaluation of a computer adaptive test for "Anxiety" (Anxiety-CAT). *Quality of Life Research*, *16*, 143-155. https://doi.org/10.1007/s11136-007-9191-7

[99]. Wang, R., & et al. (2014). StudentLife: Assessing mental health, academic performance and behavior using smartphones. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14),

[100]. Wang, Y., & et al. (2019). Fusing mobile phone sensing and brain imaging to assess depression in college students. *Frontiers in Neuroscience*, *13*, 248. https://doi.org/10.3389/fnins.2019.00248

[101]. Whiteford, H. A., Degenhardt, L., Rehm, J., & et al. (2013). Global burden of disease attributable to mental and substance use disorders. *The Lancet*, *382*(9904), 1575-1586. https://doi.org/10.1016/s0140-6736(13)61611-6

[102]. Williams, L. M. (2016). Precision psychiatry: A neural circuit taxonomy for depression and anxiety. *The Lancet Psychiatry*, *3*(5), 472-480. https://doi.org/10.1016/s2215-0366(15)00579-9

[103]. Zamal Haider, S., & Mst. Shahrin, S. (2021). Impact Of High-Performance Computing In The Development Of Resilient Cyber Defense Architectures. *American Journal of Scholarly Research and Innovation*, *1*(01), 93–125. https://doi.org/10.63125/fradxg14

[104]. Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

[105]. Zulqarnain, F. N. U., & Subrato, S. (2021). Modeling Clean-Energy Governance Through Data-Intensive Computing And Smart Forecasting Systems. *International Journal of Scientific Interdisciplinary Research*, *2*(2), 128–167. https://doi.org/10.63125/wnd6qs51

[106]. Zulqarnain, F. N. U., & Subrato, S. (2023). Intelligent Climate Risk Modeling For Robust Energy Resilience And National Security. *Journal of Sustainable Development and Policy*, *2*(04), 218-256. https://doi.org/10.63125/jmer2r39