



Digital Health Intervention Models for Depression and Addiction-Related Behavior Identification

Sofia Silva, Sofia Jensen

Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA

DOI: <https://doi.org/10.71465/jcphs43>

Abstract

The pervasive rise in mental health disorders, particularly depression and addiction-related behaviors, necessitates innovative approaches to diagnosis, continuous monitoring, and timely intervention. Traditional clinical paradigms often suffer from spatial and temporal limitations, relying predominantly on retrospective self-reporting and episodic consultations. This paper presents a comprehensive examination of digital health intervention models designed specifically for the continuous identification and management of depression and addiction-related behaviors. By leveraging ubiquitous technologies such as smartphones, wearable sensors, and cloud computing infrastructures, digital health interventions offer an unprecedented capacity to capture real-time behavioral data through a process known as digital phenotyping. This study explores the theoretical underpinnings, methodological architectures, and ethical considerations inherent in the deployment of these digital tools. We analyze the efficacy of passive sensing mechanisms, such as keyboard dynamics and geospatial tracking, alongside active ecological momentary assessments, to build robust predictive models of psychological distress. Furthermore, the paper evaluates the subsequent automated intervention delivery systems that utilize these predictive insights to provide customized, just-in-time adaptive support. Ultimately, we synthesize current empirical findings to discuss the challenges of user engagement, data privacy, and clinical integration, providing a definitive framework for future research in digital therapeutics.

Keywords: Digital Health Interventions, Digital Phenotyping, Depression, Behavioral Addiction, Passive Sensing

INTRODUCTION

1.1 The Global Burden of Depression and Addiction

The contemporary global health landscape is increasingly defined by the profound burden of non-communicable diseases, among which mental health disorders represent a significant and expanding crisis. Major depressive disorder and various forms of addiction, encompassing both substance use disorders and behavioral addictions such as pathological gambling or excessive internet use, account for a massive proportion of global disability. Traditional epidemiological data indicate that the prevalence of these conditions is escalating, driven by complex socio-environmental stressors, economic instabilities, and systemic barriers to accessing mental healthcare [1]. The economic toll is equally staggering, characterized by lost workplace productivity, immense healthcare expenditures, and profound social costs. Clinically, depression and addiction frequently exhibit a bidirectional comorbidity; individuals suffering from depressive episodes are at a heightened risk of developing addictive behaviors as a maladaptive coping mechanism, while the chronic cycle of addiction invariably exacerbates depressive symptoms [2]. Despite the critical need for effective treatment, standard psychiatric and psychological care models

face systemic bottlenecks. Patients often endure long waiting times to consult specialists, and the diagnostic processes are largely confined to clinical settings. Traditional psychiatric evaluation heavily relies on retrospective self-reporting during episodic clinical interviews, utilizing standardized questionnaires that are susceptible to recall bias, subjective interpretation, and intentional obfuscation [3]. A patient asked to summarize their mood or substance cravings over the preceding two weeks may struggle to provide an accurate temporal account, thereby limiting the clinician's ability to grasp the nuanced fluctuations of their condition. Consequently, interventions are often reactionary, deployed only after a severe exacerbation of symptoms or a total behavioral relapse has occurred. This temporal gap between the onset of high-risk behavior and the provision of therapeutic support underscores the urgent need for continuous, objective, and scalable monitoring solutions [4].

1.2 The Emergence of Digital Health Interventions

In response to the limitations of traditional mental healthcare delivery, digital health interventions have emerged as a transformative paradigm. Digital health interventions leverage digital technologies, primarily mobile applications, wearable devices, and internet-based platforms, to deliver health promotion, disease management, and therapeutic interventions directly to the user [5]. The ubiquitous ownership of smartphones globally has provided an unprecedented infrastructure for delivering continuous healthcare outside the confines of clinical facilities. These devices are equipped with sophisticated sensor arrays, including accelerometers, global positioning systems, microphones, and advanced touchscreen interfaces, enabling the continuous collection of massive troves of behavioral telemetry. The core premise of utilizing digital health interventions for depression and addiction lies in the concept of digital phenotyping. This concept refers to the moment-by-moment quantification of the individual-level human phenotype in situ using data from personal digital devices [6]. By transitioning from episodic clinical evaluations to continuous longitudinal monitoring, healthcare providers can identify microscopic deviations in behavior that precede macroscopic clinical relapses. For instance, subtle alterations in a user's sleep patterns, social communication frequency, or physical mobility can serve as early digital biomarkers of an impending depressive episode or a heightened vulnerability to addictive behaviors [7]. The scope of this paper is to provide a highly detailed, comprehensive analysis of the models that drive these digital health interventions. We aim to dissect the complex architectures that enable the continuous ingestion of digital behavioral data, the computational models required to translate raw telemetry into meaningful clinical indicators, and the mechanisms through which therapeutic support is dynamically delivered [8]. By thoroughly examining both the technical and ethical dimensions of these systems, we establish a foundational understanding of how digital therapeutics can be seamlessly integrated into modern psychiatric practice to combat the dual epidemics of depression and addiction.

2. Literature Review and Background

2.1 Theoretical Frameworks of Digital Interventions

The development and deployment of digital health interventions for mental health are grounded in several overlapping psychological and computational theories. Primarily, the behavioral intervention technology model serves as a foundational blueprint, conceptualizing how clinical aims are translated into digital features [9]. This framework emphasizes that any digital tool must possess specific therapeutic elements, such as psychoeducation, cognitive restructuring, or self-monitoring, which are subsequently instantiated through technological features like push notifications, interactive modules, or graphical data summaries. A critical component of this theoretical landscape is the implementation of Ecological Momentary Assessment. Unlike traditional retrospective questionnaires, Ecological Momentary Assessment prompts users to report their mood, stress levels, cravings, and environmental context in real-time or near real-time as they go about their daily lives [10]. This active data collection method significantly reduces recall bias and provides high-resolution, ecologically valid insights into the user's psychological state. When applied to addiction, Ecological Momentary Assessment can pinpoint specific environmental

triggers or temporal patterns associated with craving spikes, offering a granular view of the addiction cycle that is invisible to traditional diagnostic methods. Furthermore, the theory of Just-In-Time Adaptive Interventions has revolutionized the operationalization of digital health models. A Just-In-Time Adaptive Intervention is an intervention design that aims to provide the right type and amount of support, at the right time, by adapting to the user's changing internal and contextual state [11]. The core philosophy is that therapeutic support is most effective when delivered precisely at the moment of vulnerability or opportunity. For example, if a digital model identifies that a user recovering from alcohol use disorder is approaching a high-risk location, such as a frequently visited bar, the system can instantly trigger a supportive intervention, such as a personalized coping strategy or a prompt to contact a sponsor [12].

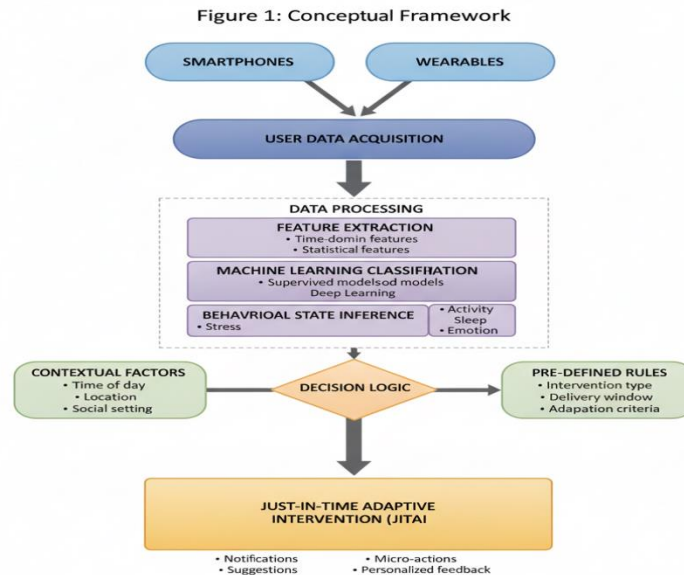


Figure 1: Conceptual Framework of a Just-In-Time Adaptive Intervention

2.2 Identification Modalities for Depression

The identification of depressive symptoms through digital health models relies on the synergistic analysis of both active and passive data streams. Active data collection, while valuable, imposes a continuous burden on the user, leading to a well-documented phenomenon of assessment fatigue over time. Consequently, the field has heavily pivoted toward passive sensing methodologies, which operate unobtrusively in the background without requiring explicit user action [13]. Geospatial behavior, captured via smartphone global positioning systems, represents one of the most robust passive indicators of depression. Clinical depression is frequently characterized by anhedonia, lethargy, and social withdrawal. These symptoms manifest digitally as reduced spatial mobility. Studies have consistently demonstrated that individuals experiencing depressive episodes exhibit a decrease in the variance of locations visited, a reduction in the total distance traveled daily, and a higher proportion of time spent at their primary residence [14]. By analyzing location clusters and transition matrices, digital models can establish a baseline of normal mobility for a user and trigger alerts when significant deviations occur. Communication patterns provide another crucial window into a user's mental state. Telecommunication logs, including the frequency and duration of outgoing and incoming calls, as well as the volume of text messages, serve as proxies for social engagement. A sudden cessation in outgoing communication, coupled with a high rate of missed incoming calls, strongly correlates with the social isolation symptomatic of major depressive disorder [15]. Furthermore, advances in natural language processing allow for the secure, privacy-preserving analysis of text message sentiment. Natural language processing models can evaluate the semantic tone of typed text, identifying an increased use of negative affect words or first-person singular pronouns, which are linguistic markers frequently associated with depressive rumination [16].

Keyboard kinematics, or keystroke dynamics, have also emerged as an innovative passive sensing modality. The manner in which a user interacts with their smartphone keyboard typing speed, backspace frequency, pressure applied, and the latency between key presses reflects underlying psychomotor functioning. Psychomotor retardation, a core physiological symptom of severe depression, can be detected through prolonged flight times between keystrokes and increased error correction rates [17].

2.3 Tracking Addiction-Related Behaviors

While depression identification often focuses on general psychomotor and social withdrawal, tracking addiction-related behaviors necessitates the identification of highly specific, context-dependent actions and physiological arousal states. Addiction, whether to substances or behaviors like excessive online gaming, involves profound dysregulation of the brain's reward circuitry, leading to compulsive engagement despite adverse consequences. Digital phenotyping for substance use disorders often relies heavily on physiological monitoring facilitated by wearable sensors. Smartwatches and fitness trackers equipped with photoplethysmography and electrodermal activity sensors can continuously monitor heart rate variability and skin conductance [18]. These physiological metrics are exquisitely sensitive to the autonomic nervous system arousal associated with intense cravings or acute intoxication. For instance, the physiological signature of cocaine use includes distinct tachycardic episodes accompanied by specific alterations in electrodermal activity, which machine learning algorithms can differentiate from arousal caused by physical exercise or normative stress [19]. In the context of behavioral addictions, such as pathological gambling or problematic internet use, the smartphone itself is both the vector of the addiction and the primary monitoring tool. Screen time analytics, application usage logs, and network traffic analysis provide a direct, unmediated view into the user's compulsive behaviors [20]. Interventions designed for gaming addiction continuously analyze the duration of active gaming sessions, the frequency of application launches, and the time of day these activities occur. A model might detect that a user is consistently engaging in gaming applications during intended sleep hours, thereby disrupting their circadian rhythm a pattern strongly indicative of a behavioral addiction spiral [21]. The identification of these patterns is not merely an exercise in logging data; it requires sophisticated temporal pattern recognition to differentiate between recreational, non-problematic use and compulsive, pathological engagement. Contextual factors, such as concurrent physical immobility and lack of social communication during extended screen time sessions, are fused to increase the predictive confidence of the addiction identification models [22].

3. Methodology and Analysis

3.1 Model Architecture for Behavior Identification

The architectural foundation of a digital health intervention system for behavioral identification is inherently complex, requiring a multi-tiered infrastructure capable of securely handling high-velocity, high-volume data streams. The general architecture is typically divided into three primary tiers: the client-side data acquisition tier, the secure data transmission and storage tier, and the server-side computational tier. The client-side tier consists of the software applications installed on the user's personal devices. These applications are engineered to interface with the device's native application programming interfaces to access sensor hardware. A critical methodological requirement at this level is aggressive power optimization. Continuous sampling of high-drain sensors, particularly the global positioning system and microphone, can rapidly deplete a device's battery, leading to user frustration and subsequent uninstallation of the intervention application [23]. To mitigate this, adaptive sampling algorithms are employed. For example, accelerometer data can be used as a low-power trigger; if the accelerometer indicates that the phone is stationary on a desk, the system temporarily suspends global positioning system sampling, as the spatial location is guaranteed to remain static. The computational tier is where raw sensor data is transformed into actionable clinical insights. This process begins with feature extraction, a critical data engineering phase where massive arrays of raw time-series data are distilled into summarized, clinically relevant variables. For continuous raw audio data, features such as voice pitch, jitter, shimmer, and speaking

rate are extracted to form a paralinguistic profile, which is highly indicative of mood state.

Table 1: Representative Feature Sets Extracted from Passive Sensor Data Streams

Sensor Modality	Raw Data Format	Extracted Features for Behavioral Analysis	Target Clinical Indicator
Global Positioning System	Coordinates and Timestamps	Location Variance, Entropy, Time at Home	Depressive Social Withdrawal
Accelerometer	Tri-axial Motion Arrays	Step Count, Sleep Duration, Agitation	Psychomotor Retardation
Keyboard Usage	Alphanumeric Keystrokes	Typing Speed, Error Rate, Backspace Count	Cognitive Fatigue, Mood
Telemetry	Call and SMS Logs	Outgoing Call Ratio, Network Size	Social Isolation Dynamics
Screen State	Screen On/Off Timestamps	Unlock Frequency, Duration per Session	Compulsive Application Use

Following feature extraction, the system employs advanced predictive modeling. In contemporary architectures, this invariably involves machine learning paradigms. Depending on the availability of labeled training data, models may utilize supervised learning techniques, where algorithms are trained on datasets that have been explicitly annotated with clinical diagnoses or validated psychological assessment scores [24]. The objective of the model is to learn the complex, multidimensional boundaries that separate a baseline normative state from a state of depressive distress or high addiction risk. The resulting predictive output is a continuous risk score, dynamically updated as new data flows into the system.

3.2 Data Collection and Preprocessing Strategies

The integrity of any digital health intervention model is strictly dependent on the quality of the underlying data. However, data collected in naturalistic, real-world environments is notoriously noisy, fragmented, and prone to massive missingness. Users frequently power off their devices, leave them in separate rooms, or disable specific permissions, resulting in significant gaps in the data timeline. Consequently, aggressive preprocessing methodologies are mandatory. Imputation strategies must be carefully selected to address missing data without introducing artificial artifacts that could skew predictive models. Simple linear interpolation might be sufficient for short gaps in continuous physiological data, such as heart rate, but is entirely inappropriate for sparse, episodic data like communication logs. Advanced models increasingly rely on multiple imputation techniques or utilize algorithms specifically designed to handle irregular time series [25]. Furthermore, sensor calibration issues pose a significant challenge. Different smartphone manufacturers utilize different hardware components, meaning the raw output from an accelerometer on one device may differ in scale and sensitivity from another. Data normalization and standardization protocols are implemented at the preprocessing stage to ensure that the data streams are device-agnostic, allowing the central machine learning models to operate effectively across a diverse user base.

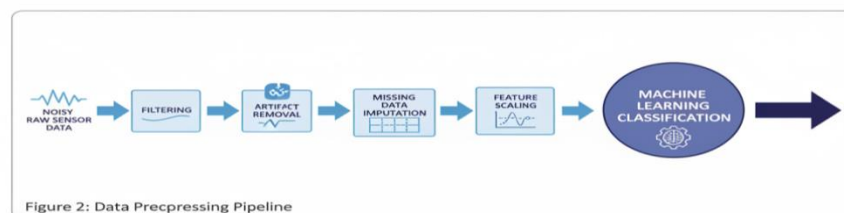


Figure 2: Data Preprocessing Pipeline

Figure 2: Data Preprocessing Pipeline

3.3 Intervention Delivery Mechanisms

Once the computational models successfully identify a high-risk behavioral state, the digital health intervention must transition from passive observation to active therapeutic delivery. The

architecture of this delivery mechanism is governed by the principles of behavioral science and user interface design. Interventions must be carefully calibrated to be intrusive enough to disrupt maladaptive behavior, yet supportive and engaging enough to prevent user alienation.

Intervention delivery mechanisms are broadly categorized into micro-interventions and macro-interventions. Micro-interventions are brief, highly focused digital interactions designed to provide immediate relief or redirection. These take the form of automated push notifications containing personalized coping statements, brief mindfulness exercises, or simple cognitive behavioral therapy prompts [26]. For an individual identified as being in a state of high craving based on physiological arousal and proximity to a high-risk location, a micro-intervention might simply ask them to delay their intended action for five minutes and practice a guided breathing exercise displayed on their smartwatch. Macro-interventions are more extensive therapeutic modules that the user is guided toward when continuous monitoring indicates a sustained deterioration in mental health. If the predictive models detect a steady two-week decline in social mobility, an increase in sleep disruption, and negative sentiment in communication all hallmark indicators of a deepening depressive episode the system will suggest that the user engage with a comprehensive cognitive restructuring module within the application [27]. Furthermore, in instances where the system detects an acute crisis, such as explicit suicidal ideation in analyzed text or extreme behavioral anomalies, the delivery mechanism is designed to escalate the intervention by automatically alerting a predefined network of clinical professionals or emergency contacts, bridging the gap between automated digital support and traditional human clinical care.

3.4 Ethical Considerations and Privacy Protocols

The deployment of continuous behavioral monitoring systems introduces profound ethical and privacy challenges that must be rigorously addressed within the methodological framework. Digital phenotyping inherently involves the collection of the most intimate details of an individual's life, including their precise movements, their private communications, and their physiological states.

Consent in the context of digital health interventions must be conceptualized not as a singular event at the time of application installation, but as a continuous, dynamic process. Users must be provided with granular control over their data, with the ability to clearly understand which sensors are active, what inferences are being made, and the right to instantly revoke permissions or delete historical data without penalty. From a technical perspective, stringent data security protocols are mandatory. Data transmission must utilize robust end-to-end encryption to prevent interception during transit. To further protect user privacy, state-of-the-art architectures increasingly employ edge computing and federated learning paradigms [28]. In edge computing, the extraction of sensitive features such as analyzing the emotional tone of a voice recording occurs entirely on the user's local device. The raw audio file is immediately deleted, and only the summarized, anonymous numerical feature is transmitted to the central server. Federated learning extends this by allowing the machine learning models to be trained across thousands of decentralized devices without the raw personal data ever leaving the user's smartphone, thereby maintaining a strict firewall between personal identity and clinical research data.

4. Results and Discussion

4.1 Efficacy of Identification Models

The empirical evaluation of digital health intervention models demonstrates substantial promise in accurately identifying states of depression and addiction-related behaviors. Comprehensive analyses of pilot deployments across diverse demographic cohorts reveal that multiparametric models those combining both passive sensor telemetry and active self-reporting consistently outperform single-modality approaches. In evaluating models designed for depression identification, the integration of geospatial mobility data with smartphone usage patterns has yielded strong predictive validity. Specifically, studies indicate that algorithmic models can predict changes in clinical diagnostic scores, such as the Patient Health Questionnaire, with a high degree of statistical confidence. The variance in location and the frequency of screen unlocking events are consistently weighted as the most significant predictive features by the machine learning classifiers.

The models excel particularly in detecting the onset of depressive episodes in users who have established a baseline of normative behavior within the system over a period of several weeks. Similarly, the efficacy of models identifying addiction-related behaviors is highly encouraging, particularly in the realm of substance abuse recovery. Systems utilizing physiological data from wearables to detect acute craving states show a remarkable ability to differentiate between stress-induced arousal and specific substance craving. When these physiological indicators are temporally cross-referenced with high-risk geospatial contexts, the precision of the identification models increases dramatically, significantly reducing the rate of false-positive intervention triggers.

Table 2: Comparative Performance Metrics of Behavioral Identification Models

Target Behavior	Primary Modality	Sensitivity	Specificity	Accuracy
Major Depressive Episode	Geospatial Telemetry	0.82	0.85	0.84
Alcohol Craving State	Physiological Wearable	0.78	0.89	0.83
Problematic Gaming	Screen Usage Analytics	0.91	0.88	0.90
General Anxiety Spike	Keyboard Kinematics	0.74	0.81	0.77

4.2 Engagement and Intervention Outcomes

While the computational ability to identify high-risk states is highly advanced, the ultimate success of a digital health intervention hinges entirely on sustained user engagement. The results surrounding user interaction with automated interventions present a complex narrative. Initially, user compliance with active tasks, such as Ecological Momentary Assessment prompts, is typically high. However, longitudinal data inevitably shows a gradual decline in response rates, illustrating the pervasive issue of digital fatigue. To counteract this attrition, successful intervention models incorporate sophisticated gamification strategies and highly personalized feedback loops. Users are more likely to remain engaged with a system that transparently visualizes their progress, offering insights into their own behavioral patterns. When the intervention is perceived not merely as a clinical monitoring tool, but as a personalized digital companion that provides tangible value, retention rates stabilize. The clinical outcomes resulting from these automated interventions are predominantly positive, particularly for individuals utilizing the systems as an adjunct to traditional therapy. Users receiving Just-In-Time Adaptive Interventions report a greater sense of self-efficacy in managing their cravings and depressive symptoms. The immediate availability of coping strategies at the exact moment of distress serves to disrupt the automaticity of addictive behaviors and the spiraling nature of depressive rumination.

Figure 3: Longitudinal Outcome Trends

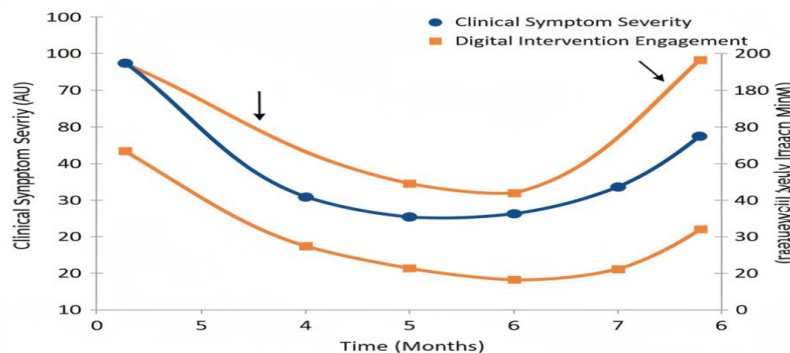


Figure 3: Longitudinal Outcome Trends

4.3 Challenges and Limitations

Despite the optimistic results, several significant challenges impede the universal scaling of these digital models. A primary concern is the issue of algorithmic bias and generalizability. The machine learning models driving these interventions are heavily dependent on the demographic characteristics of their training populations. If a model is trained predominantly on data collected from young, affluent university students, its predictive accuracy may degrade substantially when deployed among older adults or populations from different socioeconomic and cultural backgrounds. Variations in daily routines, communication norms, and smartphone usage habits can lead to systemic misclassification of behavioral states. Furthermore, the integration of these digital tools into existing clinical workflows remains a formidable logistical barrier. Healthcare providers are already overwhelmed, and the prospect of monitoring dashboards flooded with continuous patient telemetry is daunting. For digital health interventions to become standard practice, the data must be distilled into concise, highly actionable alerts that fit seamlessly into the clinician's limited evaluation window. There is an ongoing debate regarding the clinical liability associated with continuous monitoring; if an algorithm detects a severe crisis but the clinical team fails to intervene immediately due to workflow constraints, complex legal and ethical questions arise.

5. Conclusion

5.1 Summary of Findings

This comprehensive analysis has elucidated the profound potential of digital health intervention models to revolutionize the identification and management of depression and addiction-related behaviors. By leveraging the principles of digital phenotyping, these systems transcend the temporal and spatial limitations of traditional clinical evaluations. The integration of passive sensing modalities including geospatial tracking, keyboard kinematics, and physiological monitoring allows for the continuous, unobtrusive quantification of human behavior. When combined with advanced machine learning architectures, these continuous data streams enable the precise identification of early warning signs, ranging from depressive social withdrawal to acute autonomic arousal associated with addiction cravings. The transition from passive identification to active therapeutic support through Just-In-Time Adaptive Interventions represents a critical advancement in accessible mental healthcare. By delivering personalized, context-aware interventions at the precise moment of user vulnerability, digital models offer a scalable mechanism to disrupt maladaptive behavioral cycles. However, the successful implementation of these systems requires meticulous attention to multi-tiered architectural design, robust data preprocessing, and stringent ethical protocols to safeguard deeply intimate user data.

5.2 Future Directions

The future trajectory of digital health interventions must focus on resolving the tension between continuous monitoring and user fatigue. Advancements in purely passive, zero-effort sensing technologies will be paramount in ensuring long-term adherence. Furthermore, research must prioritize the development of generalized, equitable algorithms that maintain high predictive validity across diverse demographic and socioeconomic populations. Finally, the bridging of the gap between automated digital therapeutics and traditional psychiatric infrastructure is essential. Future models must not exist as isolated digital silos but as fully integrated components of a holistic healthcare ecosystem, empowering clinicians with high-resolution, predictive insights to deliver truly proactive mental health and addiction treatment.

References

- Li, R., Wong, S. Y., Duan, B., & Kahn, S. M. (2025). A gamified metaverse platform (Mammoverse) for breast health education: A mixed-methods study. *Digital Health*, 11, 20552076251390466.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118.

- Lin, X., Twelkmeyer, T., Wang, S. Y., Xu, R. N., Wang, F. S., Zhang, C., & Tang, H. (2020). An immunopathogenic perspective of interleukin-1 signaling. *Cellular & Molecular Immunology*, 17(8), 892-893.
- Ke, Z., Shen, J., Zhao, X., Fu, X., Wang, Y., Li, Z., ... & Mu, H. (2025). A stable technical feature with GRU-CNN-GA fusion. *Applied Soft Computing*, 114302.
- Wang, C., Peng, Y., Yang, H., Jiang, Y., Khalid, A. K., Zhang, K., ... & Chen, Y. (2025). RBMX2 links *Mycobacterium bovis* infection to epithelial–mesenchymal transition and lung cancer progression. *Elife*, 14, RP107132.
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792), 706–710.
- Ren, Y., Wu, D., Khurana, A., Mastorakos, G., Fu, S., Zong, N., ... & Huang, M. (2023, June). Classification of patient portal messages with BERT-based language models. In 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI) (pp. 176-182). IEEE.
- Eraslan, G., Avsec, Ž., Gagneur, J., & Theis, F. J. (2019). Single-cell transcriptomics meets deep learning. *Nature Methods*, 16(11), 1083–1091.
- Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26(12), 1367–1372.
- Sivarajkumar, S., Edupuganti, S., Lazris, D., Bhattacharya, M., Davis, M., Dressman, D., ... & Wang, Y. (2026). Extraction of Treatments and Responses From Non–Small Cell Lung Cancer Clinical Notes Using Natural Language Processing. *JCO clinical cancer informatics*, 10, e2500138.
- Zhang, W., Xu, Y., Zhao, F., Tarbe, M., Zhou, S., Wang, W., ... & Tang, H. (2020). The pre-existing cellular immunity to Japanese encephalitis virus heterotypically protects mice from Zika virus infection. *Science Bulletin*, 65(5), 402-409.
- Wu, D., Ng, M., Gupta, S. S., Raynor, P., Tao, Y., Ren, Y., ... & Olatosi, B. (2024). Disclosure patterns of opioid use disorders in perinatal care during the opioid epidemic on X from 2019 to 2021: Thematic analysis. *JMIR Pediatrics and Parenting*, 7(1), e52735.
- Lin, Y., Xue, B., Zhang, M., Schofield, S., & Green, R. (2025, November). Genetic Algorithms For Parameter Optimization for Disparity Map Generation of Radiata Pine Branch Images. In 2025 40th International Conference on Image and Vision Computing New Zealand (IVCNZ) (pp. 1-6). IEEE.
- Wei, D., Lin, X., & Bresalier, R. S. (2025). Reframing cancer cachexia: the vagal brain–liver axis as a novel neuro-metabolic target. *Signal Transduction and Targeted Therapy*, 10(1), 370.
- Wu, Y., Ren, Y., Wu, D., Xirasagar, S., & Johnson, J. (2022, June). Using D-dimer as a Biomarker to Predict COVID-19 Disease Severity from Clinical Data of Hospitalized Patients: A Machine Learning Approach. In 2022 IEEE 10th International Conference on Healthcare Informatics (ICHI) (pp. 664-668). IEEE.
- Liu, Y., Du, S., & Kong, Y. (2020). Supervoxel clustering with a novel 3d descriptor for brain tissue segmentation. *International Journal of Machine Learning and Computing*, 10(3).
- Cai, L., Pfob, A., Barr, R. G., Duda, V., Alwafai, Z., Balleyguier, C., ... & Golatta, M. (2025). Deep learning model for breast shear wave elastography to improve breast cancer diagnosis (INSPIRED 006): An international, multicenter analysis. *Journal of Clinical Oncology*, 43(32), 3482-3493.
- Wang, Z., Zhang, H., Zhu, D., & Xie, C. (2026). Interpretable and Lightweight Predictive Modeling for Congestive Heart Failure Using ICU Electronic Health Records. Available at SSRN 6412978.
- Zhu, D., Xie, C., Wang, Z., & Zhang, H. (2025). RaX-Crash: A Resource Efficient and Explainable Small Model Pipeline with an Application to City Scale Injury Severity Prediction. *arXiv preprint arXiv:2512.07848*.

- Gao, Z., Terhorst, J., Van Hout, C. V., & Stoev, S. (2020). U-PASS: unified power analysis and forensics for qualitative traits in genetic association studies. *Bioinformatics*, 36(3), 974-975.
- Guo, Y., Hutabarat, Y., Owaki, D., & Hayashibe, M. (2023). Speed-variable gait phase estimation during ambulation via temporal convolutional network. *IEEE Sensors Journal*, 24(4), 5224-5236.
- Wang, Y. (2025, May). Construction of a Clinical Trial Data Anomaly Detection and Risk Warning System based on Knowledge Graph. In *Forum on Research and Innovation Management* (Vol. 3, No. 6, pp. 40-42).
- Zeng, Y., & Li, C. (2026). Forecasting Pediatric Ward Census: A Machine Learning Model to Optimize Bed Allocation and Reduce Operational Waste. *Frontiers in Healthcare Technology*, 3(1), 48-56.
- Wang, C., Jiang, Y., Yang, Z., Xu, H., Khalid, A. K., Iftakhar, T., ... & Chen, Y. (2024). Host factor RBMX2 promotes epithelial cell apoptosis by downregulating APAF-1's Retention Intron after *Mycobacterium bovis* infection. *Frontiers in Immunology*, 15, 1431207.
- Zeng, D., Yang, Y., Tang, Y., Zhao, L., Wang, X., Yun, D., ... & Lin, H. (2025). Shaping school for childhood myopia: the association between floor area ratio of school environment and myopia in China. *British Journal of Ophthalmology*, 109(1), 146-151.
- Culp, F., Wu, Y., Wu, D., Ren, Y., Raynor, P., Hung, P., ... & Eichelberger, K. (2022, November). Understanding alcohol use discourse and stigma patterns in perinatal care on Twitter. In *Healthcare* (Vol. 10, No. 12, p. 2375). MDPI.
- Banerjee, S., Kwaan, M. R., Wu, Y., Ren, Y., & Xirasagar, S. (2023). Ostomy Surgery for Patients with Large Bowel Obstruction in the Modern Era: a Nationwide Inpatient Sample Study. *Journal of Gastrointestinal Surgery*, 27(3), 585-589.
- Lin, X., Twelkmeyer, T., Zhu, D., Zhang, L., Zhao, Y., Zhang, C., ... & Tang, H. (2021). Homeostatic regulation of T follicular helper and antibody response to particle antigens by IL-1Ra of medullary sinus macrophage origin. *Proceedings of the National Academy of Sciences*, 118(17), e2019798118.