

The Future of Psychosocial Healthcare Research: Integrating Artificial Intelligence, Equity Science, Behavioral Medicine, and Human-Centered Care

Lina Hassan^{1*}, Omar Khalaf², Reem Jaber¹, Rania Mostafa³

¹Department of Psychosocial Healthcare Research and AI, Faculty of Medicine, University of Jordan, Amman, Jordan.

²Department of Equity Science and Behavioral Medicine, Faculty of Medicine, Jordan University of Science and Technology, Irbid, Jordan.

³Department of Human-Centered Care and AI Integration, Faculty of Medicine, Al-Balqa Applied University, Salt, Jordan.

*E-mail ✉ lina.hassan@ju.edu.jo

Abstract

Psychosocial healthcare research is entering a period in which artificial intelligence, continuous behavioral measurement, adaptive intervention, equity science, and digitally mediated care are developing simultaneously rather than independently. Their convergence creates opportunities to measure behavior in everyday settings, tailor support dynamically, augment communication, and make social context more visible to care systems. It also creates substantial risks if computational performance is mistaken for psychological understanding, behavioral prediction for clinical benefit, personalization for patient-centeredness, or increased data availability for equitable representation. This Horizon Review examines emerging signals across digital psychiatry, generative artificial intelligence, multimodal behavioral data, precision behavioral medicine, digital phenotyping, algorithmic fairness, social determinants, and relational care. The scan treats future horizons as evidence-bounded research trajectories rather than forecasts and distinguishes observed developments from proposed convergence and weak signals. We propose that future psychosocial healthcare research should prioritize four forms of validity simultaneously: empirical validity, relational validity, equity validity, and implementation validity. Nearer-term research is likely to center on bounded artificial intelligence augmentation and improved governance; intermediate horizons may connect multimodal sensing with adaptive behavioral support; longer-term possibilities involve increasingly continuous and hybrid human-digital systems but remain substantially uncertain. Progress will depend less on maximizing technological capability in isolation than on determining when computational outputs remain psychologically interpretable, clinically useful, relationally acceptable, structurally fair, and feasible within real health systems.

Keywords: Psychosocial healthcare, Artificial intelligence, Digital phenotyping, Precision behavioral medicine, Algorithmic fairness, Human-centered care

Introduction

Digital technologies are increasingly embedded within mental-health and psychosocial care, spanning mobile interventions, conversational systems, remote

assessment, social-media-derived information, sensing technologies, and virtual environments [1]. Yet the future of psychosocial healthcare cannot be understood as a simple progression from less technology to more technology. Even established digital mental-health interventions show uneven effectiveness, engagement, implementation, and usability across populations and settings [2]. Similarly, digital phenotyping can transform continuously generated behavioral traces into potentially informative measures, but the transition from observable behavior to clinically meaningful inference requires explicit attention to purpose, quality, safety, privacy, and

Access this article online

<https://smerpub.com/>

Received: 06 June 2026; Accepted: 09 September 2026

Copyright CC BY-NC-SA 4.0

How to cite this article: Hassan L, Khalaf O, Jaber R, Mostafa R. The Future of Psychosocial Healthcare Research: Integrating Artificial Intelligence, Equity Science, Behavioral Medicine, and Human-Centered Care. Int J Soc Psychol Asp Healthc. 2026;6(2):220-34. <https://doi.org/10.51847/F9USg3CUJK>

contextual validity [3]. Adaptive intervention science further demonstrates that personalization is not merely matching a person to a program; it involves deciding when, why, and under what changing conditions an intervention should be delivered [4].

These developments expose a larger conceptual problem. Psychological experience, observable behavior, therapeutic relationship, service participation, clinical status, organizational capability, structural disadvantage, and digital access are related but noninterchangeable domains. A model may classify language accurately without understanding its social meaning. A sensor may detect behavioral change without identifying its cause. A chatbot may be experienced as supportive without producing durable clinical benefit. An efficient workflow may nevertheless erode trust, exclude people with limited digital access, or amplify inequities embedded upstream. This Horizon Review therefore examines the future of psychosocial healthcare as a problem of cross-domain convergence under uncertainty. Its central argument is that technological capability should not be the principal unit of future readiness. Instead, emerging approaches should be evaluated jointly for empirical performance, psychological interpretability, relational consequences, equity conditions, and health-system implementability. The resulting synthesis is explicitly prospective and proposed where it extends beyond established evidence.

Scope and horizon definition

Horizon scanning is used here to identify developments that may become consequential for psychosocial research and care, not to predict their probability of occurrence. Health-related horizon-scanning methods typically combine identification, assessment, and prioritization of emerging technologies or practices, but methodological approaches vary substantially and do not provide a universal forecasting rule [5]. Accordingly, the horizons in this review represent differences in evidence proximity and translational dependency rather than deterministic timelines.

The scope includes technologies and research architectures that can alter how psychosocial states, behaviors, relationships, social conditions, or care processes are measured, interpreted, supported, or governed. The horizon therefore extends beyond mental-health applications narrowly defined. It includes computational support for patient communication, multimodal behavioral sensing, adaptive behavioral interventions, algorithmic fairness, social-context data,

relational design, and workforce transformation where these developments affect psychological or psychosocial care.

Three distinctions bound the analysis. First, a signal is an observable development in research, technology, implementation, governance, or patient experience. Second, an emerging horizon is an evidence-informed trajectory constructed from related signals. Third, a potential discontinuity is a plausible but poorly established development capable of changing assumptions about care. None of these categories implies a validated probability of future occurrence.

Horizon-scanning approach

The scan used targeted searches across PubMed-indexed scholarship, journal and publisher records, and related scholarly indexing sources covering horizon-scanning methodology and the principal domains defined in advance for this review. Search families addressed generative artificial intelligence and mental health, digital sensing and phenotyping, adaptive behavioral intervention, algorithmic bias and health equity, social determinants and digital readiness, patient trust, participatory design, and technology-related workforce change. Contemporary evidence was emphasized while retaining earlier methodological or conceptual work where necessary.

Published horizon-scanning and foresight research shows considerable variation in signal sources, stakeholder involvement, search techniques, and reporting practices [6]. For this reason, the scan did not treat methodological heterogeneity as a numerical ranking problem. Candidate signals were instead examined for relevance to psychosocial healthcare, conceptual distinctiveness, evidence maturity, plausibility of cross-domain influence, equity implications, and identifiable validation requirements.

Because the retrieval environment used ranked scholarly searches rather than a fixed database export suitable for reproducible prevalence estimation, the scan is not interpreted as an exhaustive census of publications. Search-result frequency is therefore never used as evidence that one future domain is intrinsically more important than another.

Signal sources and evidence domains

Signals were drawn from analytically distinct evidence domains. Computational signals included generative models, prediction systems, sensing platforms, and

multimodal models. Behavioral signals concerned repeated measurement, contextual adaptation, engagement, and dynamic intervention. Relational signals included patient trust, communication, perceived clinician presence, co-design, and human oversight. Structural signals concerned social determinants, differential digital access, representation, and power. Organizational signals included workflow integration, workforce burden, implementation infrastructure, and governance.

Keeping these domains separate is important because digital interventions can generate inequity at multiple stages, including access, uptake, adherence, and effectiveness [7]. Information abundance itself can also be misleading; technical examination of mental-health-related applications has demonstrated that actual data-sharing behavior may diverge from privacy representations [8]. Signals were consequently assessed not only for what a technology could do, but also for the conditions under which its outputs are generated, interpreted, and acted upon.

Signal identification and screening

Signal identification emphasized developments capable of altering at least one of five psychosocial research functions: measurement, interpretation, intervention, relationship, or system organization. Signals were retained when they had a plausible connection to future care and enough empirical or methodological substance to define what would need to be tested next. Technological novelty alone was insufficient.

Screening also distinguished evidence type. Randomized intervention evidence was not considered equivalent to observational sensing research, qualitative participatory evidence, algorithm audits, consensus recommendations, or patient-attitude surveys. This separation prevents a frequent horizon-scanning error in which diverse signals become visually or rhetorically comparable merely because they concern the same technology.

Signals with clear evidence of current function but uncertain transfer were retained as context-dependent developments. Signals requiring several untested assumptions between current evidence and future clinical application were retained only as weaker horizons. Claims that depended primarily on speculative technological capability without a defensible psychosocial or health-system pathway were not used to establish priority.

Thematic clustering

Signals were clustered by the scientific problem they create rather than by technology brand or data modality. Five clusters emerged as particularly useful for organizing the scan: computational psychosocial care, continuous and multimodal behavioral measurement, adaptive behavioral intervention, equity-aware data and algorithms, and human-centered relational systems.

This clustering deliberately allows overlap. Digital phenotyping, for example, belongs to behavioral measurement but can affect adaptive intervention; generative AI belongs to computational care but can also influence clinician workload and therapeutic interaction. Cross-cluster links are therefore treated as hypotheses about convergence rather than as evidence that the clusters have already become one integrated system.

Novelty, plausibility, impact, and uncertainty assessment

Novelty was interpreted as a meaningful change in capability, research design, or implementation context rather than the appearance of a new technical label. Plausibility required a credible pathway from the observed signal to a psychosocial research or care consequence. Potential impact concerned the breadth or depth of the change that could follow if the pathway proved valid. Uncertainty captured missing validation, population transfer, measurement ambiguity, implementation dependency, and possible countervailing effects.

These dimensions were not combined into a numerical score. A highly novel system may have low clinical plausibility. A technically mature tool may have high implementation uncertainty. A modest technological change may have substantial relational consequences. Treating these dimensions independently better reflects the asymmetric evidence typical of emerging psychosocial technologies.

Prioritization framework

We propose that a psychosocial healthcare signal should receive greater research priority when four conditions converge. First, its underlying measurement or intervention has enough empirical validity to justify further translation. Second, its meaning remains interpretable in relation to psychological, behavioral, or clinical processes rather than merely correlating with them. Third, its operation has identifiable equity conditions, including who is represented, who can participate, and whose data are missing. Fourth, it has

plausible relational and implementation validity, meaning that patient trust, clinician roles, workflow, and governance can be examined rather than assumed.

Adaptive intervention science illustrates why these conditions matter. Time-sensitive tailoring requires explicit decision points, changing tailoring variables, and defined intervention options rather than generic personalization [4]. At the same time, digital-health inequality can enter before, during, or after intervention delivery [7], and institutional trust cannot be inferred from technical performance. Privacy practices can further constrain whether repeated personal data collection is acceptable [8].

Priority should therefore attach not simply to the most advanced technology but to the most testable

convergence of capability and context. Signals are lower priority for immediate translation when the computational component is mature but the psychosocial interpretation remains ambiguous, when evidence is promising but population transport is unknown, or when deployment requires institutional conditions that have not been demonstrated.

The distinctions required for research prioritization are easier to compare as propositions than as a ranked list; **Table 1** therefore organizes the patient and care-context states that determine what must be tested before an emerging capability can be treated as a viable psychosocial-care horizon.

Table 1. Testability conditions that separate technically promising signals from research-ready psychosocial healthcare priorities

Patient or care-context state	What is currently observable	Mechanism that requires testing	Context that can change interpretation	Competing explanation or unresolved uncertainty	Boundary before translation	Proposed research consequence
Repeated behavioral change is detectable	Digital activity or interaction patterns vary over time	Whether the pattern tracks a clinically meaningful within-person state	Device use, routines, disability, work, culture, access	Behavioral change may reflect environment rather than psychological change	Association cannot be treated as direct measurement of internal state	Test within-person validity and context sensitivity before triggering care
Adaptive support can be delivered at selected moments	A system can modify content or timing	Whether timing corresponds to a state in which intervention can help	Burden, receptivity, competing demands, treatment stage	Increased contact may not improve engagement or outcome	Personalization is not equivalent to effectiveness	Predefine decision rules and proximal outcomes
Computational communication appears fluent or understandable	AI can produce patient-facing language or draft responses	Whether communication remains accurate, safe, relationally appropriate	Literacy, language, clinical complexity, cultural expectations	Fluency may conceal omission or error	Readability cannot substitute for clinical validation	Evaluate comprehension, accuracy, escalation, trust, and downstream action separately
Model performance is acceptable in development data	Predictions or classifications meet technical benchmarks	Whether performance and meaning transfer across populations	Representation, structural disadvantage, prevalence, measurement practice	Performance differences may reflect biased proxies or unequal data generation	Aggregate accuracy cannot establish fairness	Require subgroup and context-specific validation with transparent dataset documentation
Patients can access a digital intervention	Enrollment or technical connectivity is feasible	Whether participation remains acceptable and sustainable	Digital readiness, privacy concerns, disability, language, resources	Access may not produce use, adherence, or benefit	Availability cannot be interpreted as equitable reach	Measure access, uptake, experience, and outcome as distinct stages
AI is incorporated	Clinicians can view or use	Whether augmentation	Staffing, EHR design, liability,	Apparent automation may	Workflow adoption cannot	Evaluate workload, cognitive burden,

into a clinical workflow	generated material	improves care without shifting hidden work or responsibility	oversight, organizational resources	redistribute rather than remove workload	establish patient or clinician benefit	patient experience, error recovery, and accountability together
--------------------------	--------------------	--	-------------------------------------	--	--	---

These states are proposed research-priority categories, not validated scoring dimensions.

The same reasoning is represented visually in **Figure 1** because prioritization depends on crossing several

evidence and interpretation boundaries rather than progressing through a single pipeline.

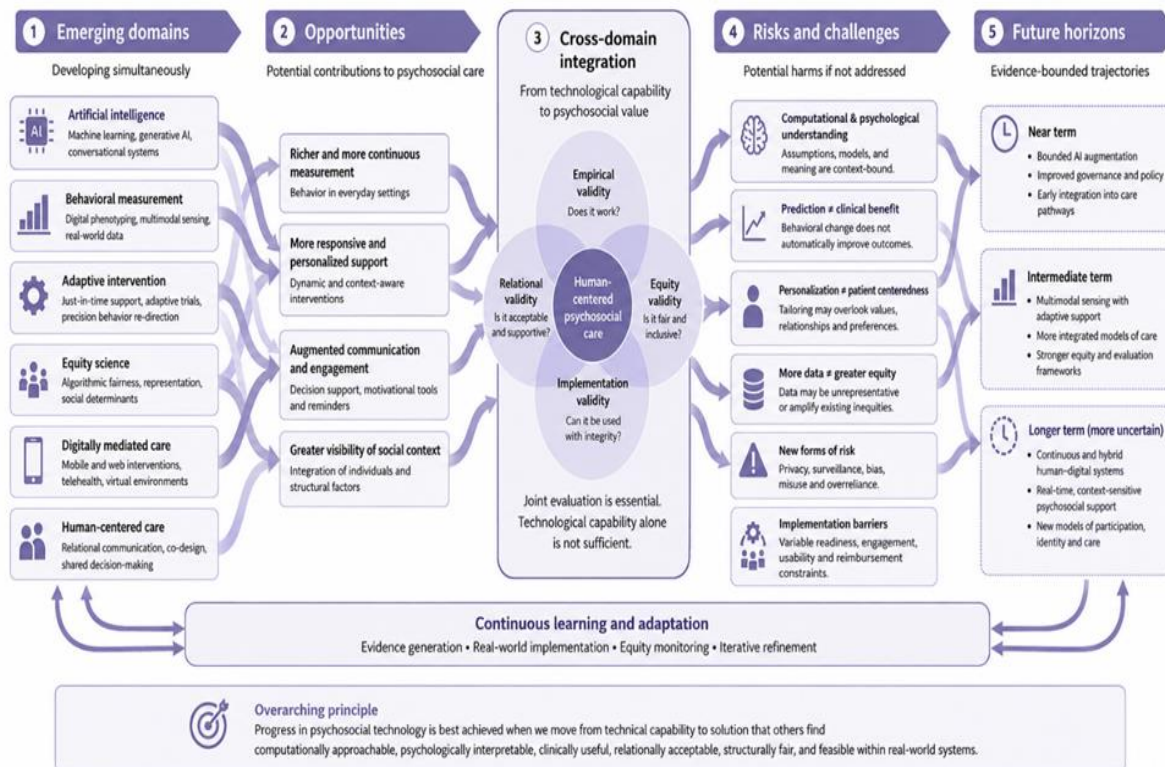


Figure 1. Where psychosocial technology signals become research-ready horizons: evidence maturity, interpretability, equity, and implementation boundaries

Emerging horizons

Three emerging patterns are particularly visible. First, generative AI is moving from demonstration tasks toward direct mental-health support, communication, and clinical workflow applications. Recent reviews describe rapid expansion but also heterogeneous evaluation, limited standardization, and unresolved problems involving clinical validity, cultural competence, safety, trust, and relational adequacy [9-11]. A meta-analysis of generative mental-health chatbots suggests potential symptom benefit on average, but variation across interventions and populations makes broad effectiveness claims premature [12].

Second, passive and multimodal data are shifting research from episodic self-report toward denser representations of everyday behavior. Third, adaptive intervention methods are increasingly capable of linking changing states to changing support. These developments are potentially complementary, but complementarity should not be confused with validated integration. A behavioral state inferred from a smartphone does not automatically identify the correct intervention; an adaptive intervention does not establish that the sensed state was psychologically meaningful.

The emerging horizons are therefore best understood as mechanism-context cases rather than technology

categories. **Table 2** makes explicit how apparently similar developments can diverge once equity, measurement, relational, and organizational conditions are considered.

Table 2. Equity-sensitive pathways through which emerging psychosocial technologies may diverge in real-world care

Emerging mechanism-context case	Immediate capability	Psychosocial pathway under consideration	Equity-sensitive modifier	Evidence uncertainty	Condition that could reverse the expected value	Proposed horizon interpretation
Generative conversational support	Produces rapid natural-language interaction	May extend psychoeducation, self-reflection, or low-intensity support	Language, cultural norms, disability access, digital literacy, privacy	Relational quality and durable clinical benefit are incompletely established	Unsafe advice, cultural mismatch, false reassurance, or delayed escalation	Nearer-term use is more defensible when bounded and supervised than when autonomous
AI-assisted patient communication	Rewrites or drafts patient-facing content	May improve accessibility and reduce communication friction	Health literacy, language, trust, clinical complexity	Fluency does not guarantee accuracy or comprehension	Omission, hallucination, altered tone, unclear responsibility	Evaluate as communication augmentation with explicit accuracy and escalation checks
Passive behavioral sensing	Captures activity and interaction patterns longitudinally	May detect within-person deviations relevant to changing need	Device ownership, work patterns, disability, housing stability, surveillance concerns	Behavioral features are context-dependent proxies	Environmental or social change mimics psychological deterioration	Useful horizon if inference remains personalized and context-aware
Multimodal assessment	Combines language, voice, video, or sensor features	May improve characterization of heterogeneous behavioral signals	Representation across culture, language, age, phenotype, and setting	External transport and construct validity remain limited	Model learns demographic or contextual correlates rather than target construct	Research priority should emphasize external and cross-context validation
Just-in-time adaptive support	Changes timing or content as states change	May align intervention with moments of need or receptivity	Notification burden, autonomy, resource constraints, treatment access	Optimal decision rules and engagement mechanisms remain unsettled	Increased intervention frequency becomes burdensome or mistimed	Medium-horizon integration depends on within-person causal testing
Social-context data integration	Makes selected social needs visible to care systems	May connect structural context to care planning	Disclosure risk, trust, stigma, organizational capacity, available services	Missingness and willingness to disclose are informative rather than random	Data collection occurs without meaningful capacity to respond	Value depends on governance and service linkage, not data capture alone
Human-AI collaborative care	Combines computational support with clinician judgment or facilitation	May preserve relational accountability while extending capacity	Workforce resources, patient preference, institutional trust	Appropriate division of responsibility remains unsettled	Automation shifts hidden workload or weakens continuity	Promising only when clinician role, patient choice, and accountability remain explicit

Entries represent a proposed horizon interpretation synthesized from the reviewed evidence rather than validated future-state categories.

Artificial intelligence and computational psychosocial care

The strongest near-term AI signal is not autonomous psychosocial treatment but diversification of bounded computational roles. In an exploratory randomized trial, a guardrailed generative-AI digital mental-health intervention was feasible over a short period and produced encouraging user-experience findings, but the design does not establish general clinical effectiveness or long-term safety [13]. A separate randomized trial of topic-based mental-health chatbots found improvements in selected self-care and well-being outcomes that were not uniformly sustained over follow-up [14]. These findings support continued evaluation while arguing against treating conversational capability itself as therapeutic validation.

Communication augmentation may mature faster because its output can be inspected before action. Generative rewriting of inpatient discharge summaries substantially improved readability and understandability, while clinician review also identified omissions and inaccuracies with potential safety relevance [15]. AI-drafted replies to patient portal messages can exhibit useful communication characteristics [16], yet integration of draft replies into electronic workflows did not automatically reduce physician reply time [17]. Computational quality, workflow efficiency, and patient benefit are therefore separate outcomes.

Trust creates an additional boundary. National survey evidence indicates limited patient trust in health systems to use AI responsibly and suggests that broader institutional trust and experiences of discrimination matter to AI acceptance [18]. Future research should consequently treat trust not as a usability score appended after model evaluation but as a relational and institutional outcome influenced by transparency, clinician involvement, prior experience, perceived accountability, and power.

Multimodal behavioral and patient-generated data

Psychosocial assessment has historically relied heavily on episodic questionnaires, interviews, and clinical encounters. Digital sensing creates the possibility of observing behavioral patterns between encounters, but the research task is not merely to collect more data. Field-level work emphasizes the need for stronger standards, adequate study design, and alignment between sensing research and clinical or community priorities [19]. Longitudinal smartphone data can support individualized mood prediction within studied cohorts [20], while real-world digital and physical activity measures show that

different indicators of activity and motivation need not relate to depression and anxiety in equivalent ways [21]. Passive smartphone features have also been associated with transdiagnostic dimensions of psychopathology [22]. Such results are scientifically useful but should not be described as direct measurement of mood, motivation, identity, trust, or meaning. A reduction in mobility, for example, may reflect psychological distress, physical illness, work conditions, caregiving, environmental disruption, disability, or deliberate behavior change. Measurement becomes psychosocially informative only when competing explanations are addressed.

Multimodal approaches add further complexity. Video, speech, and text can be combined to model self-reported depression, anxiety, or posttraumatic stress symptoms [23], but additional modalities can also multiply sources of bias and context dependence. The research horizon is therefore multimodal inference with retained provenance: each modality should preserve information about what was measured, when, under what conditions, and with what plausible alternative explanations.

Precision behavioral medicine

Precision behavioral medicine should be distinguished from increasingly fine demographic segmentation. Its more consequential direction is dynamic within-person adaptation. Contemporary JITAI science emphasizes the challenge of engaging people when intervention may be useful, sustaining engagement, and incorporating social relationships into adaptive support [24]. Meta-analytic evidence suggests a small average benefit for mental-health and well-being outcomes while also documenting design heterogeneity and risk-of-bias concerns [25].

The next research step is therefore unlikely to be maximal personalization. It is better decision science about when adaptation is justified. A system may detect an inferred state yet still lack evidence that an intervention delivered at that moment changes the desired proximal process. Precision requires distinguishing prediction from treatment selection, receptivity from need, and behavioral response from clinical outcome.

This review proposes that adaptive psychosocial interventions should be evaluated as repeated decision systems. Their validity depends on whether the tailoring variable is interpretable, whether the intervention option is acceptable, whether timing has causal relevance, and whether adaptation remains equitable for people whose

daily behavior or digital access differs from the populations used to build the system.

Digital phenotyping and continuous assessment

Digital phenotyping is most useful when understood as a measurement architecture rather than a clinical verdict. The horizon scan identified converging signals from smartphone sensing, repeated ecological assessment, individualized modeling, and multimodal data, but the evidence also makes clear that continuous observation does not eliminate measurement error. Instead, it changes its form.

Three validation layers are therefore necessary. The first is signal validity: whether the device or platform reliably captures the behavior it claims to measure. The second is construct validity: whether changes in that behavior support the intended psychological or clinical interpretation. The third is decision validity: whether acting on the inferred state improves a meaningful process or outcome. Evidence can be strong at one layer and weak at another.

Continuous assessment also changes the ethics of measurement. Frequency creates potential surveillance burden, missingness may become socially patterned, and long-term observation can reveal routines extending beyond the health purpose for which data were collected. These concerns are not peripheral to measurement quality; they influence who participates, when people disengage, and whether the resulting data represent the intended population.

Equity science and algorithmic fairness

Equity science challenges the assumption that technical accuracy averaged across a population establishes clinical validity. A widely used population-health algorithm demonstrated how health care cost, when treated as a proxy for health need, could produce substantial racial bias because expenditure reflected unequal care rather than only illness burden [26]. Clinical algorithms that explicitly adjust for race raise a related problem when social categories are treated as if they were unambiguous biological variables [27]. International consensus work now emphasizes that transparent documentation of populations, datasets, intended uses, and performance is central to responsible health-data and AI development [28].

Bias can enter at multiple stages, including data generation, model development, validation, deployment, and interpretation [29]. Digital mental-health scholarship

similarly argues that equity must be addressed throughout design, evaluation, implementation, and monitoring rather than appended after product development [30].

For psychosocial research, the implications extend beyond subgroup performance. Structural disadvantage affects whether data are generated in the first place, which devices or services are accessible, what language appears in records, whether social needs are disclosed, and how behavioral signals should be interpreted. Fairness is therefore partly a measurement problem, partly a resource and power problem, and partly a decision problem.

The horizon proposed here is not an algorithm that has achieved abstract fairness. It is a research ecosystem in which representation, context, differential missingness, accessibility, trust, and downstream allocation are treated as components of validity. This requires avoiding demographic essentialism: group membership should never be used as a shortcut for an individual's psychological state, motivation, preference, resilience, or trust.

Social and structural determinants as dynamic data

Social and structural determinants are often represented in health data as static covariates. Psychosocial care requires a more dynamic interpretation because housing, employment, caregiving, financial strain, transportation, discrimination, social support, and access to services can change over time and modify both behavior and the feasibility of treatment.

Evidence on exchanging sensitive social-determinant information shows that collection is not simply a technical interoperability problem. Patients and clinicians identify transparency about purpose, trusting relationships, staff and time capacity, and usable technical infrastructure as conditions affecting whether information should be collected and shared [31]. Digital-health readiness likewise includes more than device access: preliminary measurement work distinguishes technical readiness from concerns about the quality of digitally mediated care [32].

These findings suggest that missing social data cannot automatically be classified as missing information. Non-disclosure may reflect privacy, stigma, distrust, unclear benefit, or concern about how information will be used. Likewise, identifying a social need does not imply that an organization has resources to address it.

We therefore propose context-with-provenance as a future data principle: social information should retain when it was obtained, why it was collected, how it was reported or inferred, what level of confidence surrounds it, and whether an actionable service pathway exists. This is analytically different from increasing the number of social variables included in prediction models.

Human-centered AI and relational care

Human-centered AI in psychosocial healthcare should not be reduced to an attractive interface or satisfaction measurement. Participatory work with adolescents in rural South Africa demonstrates that generative AI can support co-design processes while simultaneously revealing cultural bias and situations in which human facilitation remains necessary [33]. Digital mental-health implementation scholarship similarly emphasizes guided-care structures, integration with health systems, stakeholder coordination, and evidence standards rather than stand-alone technological deployment [34].

Patient acceptance provides a further signal. Experimental survey evidence indicates that trust and preference for AI-assisted medical encounters can vary with clinician presence, system performance, representativeness of training data, and governance arrangements [35]. These factors imply that the social meaning of AI is partly determined by how responsibility is distributed. The same model may be interpreted differently when it advises a clinician, communicates directly with a patient, or appears to replace a human decision maker.

Relational care consequently requires outcomes that conventional model evaluation rarely captures: whether patients feel heard, whether explanations are understandable, whether uncertainty is disclosed, whether disagreement can be expressed, whether escalation to a person is possible, and whether technology preserves rather than obscures accountability. These outcomes should remain distinct from clinical efficacy and system efficiency.

The proposed human-centered horizon is therefore relationally bounded augmentation. AI should be evaluated according to the role it occupies within an encounter, the consequences of error, patient preferences, cultural context, and the availability of human repair when communication or inference fails. Human involvement is not automatically beneficial, nor is automation automatically dehumanizing. The research problem is to identify which functions can be

computationally supported without losing the relationships, interpretive judgment, accessibility, and accountability on which psychosocial care depends.

Hybrid human–digital therapeutic relationships

Hybrid care should not be defined simply by simultaneous use of clinicians and digital tools. Its central research question is allocation of relational responsibility. Short-term chatbot studies support continued investigation but do not establish equivalence with sustained therapeutic relationships [13]. Trust in AI-enabled care also depends on institutional context and prior experiences [18]. Participatory design shows that human facilitation may remain necessary when automated interaction encounters cultural ambiguity or contextual complexity [33].

A hybrid relationship should therefore specify when technology informs, communicates, supports, or acts; when a clinician reviews its output; how disagreement is handled; and who becomes responsible when automated interpretation fails. Health-system integration and guided digital-care models provide one implementation foundation [34], while patient acceptance varies according to clinician presence and governance arrangements [35]. The proposed research priority is not maximum automation but testable allocation of human and computational functions.

Workforce transformation and psychologically informed systems

AI may redistribute psychosocial work rather than simply remove it. Electronic draft-response systems have not automatically produced communication-time savings [17], whereas ambient documentation technology has been associated with improved documentation experience and reduced perceived burnout in observational implementation research [36]. These findings concern different tasks and should not be generalized into a single productivity claim.

Workforce research should therefore measure hidden review work, cognitive burden, responsibility for error correction, emotional labor, professional autonomy, and effects on clinician–patient attention. Psychologically informed systems should optimize not only task completion but also whether technology preserves capacity for listening, interpretation, uncertainty management, and relational repair.

Cross-horizon convergence

The major horizons become scientifically consequential when they interact. Continuous behavioral measurement can inform adaptive support; adaptive systems can use AI-generated content; equity conditions shape which behavioral data exist; trust influences engagement; and organizational infrastructure determines whether any of these capabilities become sustainable care.

These interactions remain proposed convergence, not an established causal architecture. A technically successful

component can still fail at another level: accurate sensing may lack psychological specificity, effective messaging may reduce trust in some contexts, or equitable model performance may be undermined by unequal access.

Table 3 distinguishes the principal boundary conditions under which cross-horizon convergence may strengthen or break down.

Table 3. Boundary conditions governing convergence of computational, behavioral, relational, and equity-oriented psychosocial care

Convergence condition	Processes that become coupled	Boundary that can interrupt convergence	What must remain analytically separate	Proposed validation need
Sensing informs adaptive support	Behavioral measurement and intervention timing	Context-dependent meaning of sensed change	Observable behavior and psychological state	Within-person construct and decision validation
AI supports clinical communication	Computational generation and relational care	Error, omission, unclear accountability	Fluency and clinical correctness	Accuracy, escalation, trust, and downstream action
Social context enters digital models	Structural information and prediction	Selective disclosure and unequal data generation	Social exposure and individual preference	Provenance-aware missingness and subgroup validation
Human oversight accompanies automation	Model output and professional judgment	Hidden workload or ambiguous responsibility	Automation and clinical accountability	Workflow, burden, and error-recovery evaluation
Personalization expands across populations	Adaptive care and equity science	Unequal access, representation, language, or digital readiness	Individual tailoring and population fairness	Transportability plus accessibility testing

The convergence itself is proposed; the listed boundaries reflect evidence-supported concerns that require direct testing.

The cross-level dependencies are represented in **Figure 2** because future psychosocial systems are more accurately understood as interacting fields than as a sequence of technologies.

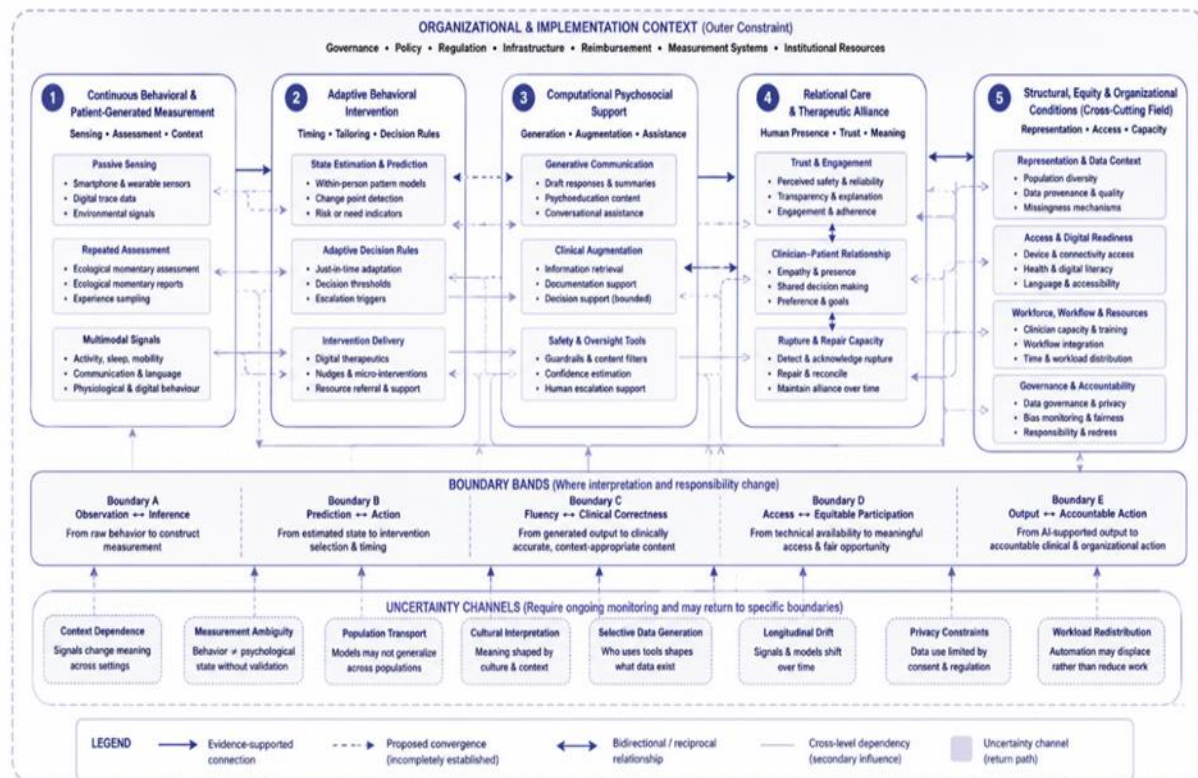


Figure 2. Convergence boundaries among continuous behavioral intelligence, adaptive intervention, relational AI, equity conditions, and health-system implementation

Near-term horizon: 1–3 years

The near-term horizon is most defensibly characterized by bounded augmentation: clinician-supervised communication tools, improved documentation support, more explicit dataset transparency, equity auditing, and increasingly structured governance. These developments are already observable, but their broader psychosocial value remains conditional on accuracy, trust, accessibility, workflow integration, and patient choice. The horizon represents evidence proximity, not a prediction that adoption will be uniform.

Medium-term horizon: 3–5 years

A medium-term possibility is tighter linkage between longitudinal behavioral sensing and adaptive intervention. Individualized behavioral models demonstrate technical feasibility [20], while adaptive-intervention science provides methods for time-varying decisions [24]. The unresolved step is causal and interpretive: systems must show that sensed changes represent meaningful states and that modifying intervention at those moments improves relevant

proximal processes without imposing excessive burden or inequity.

Longer-term horizon: 5–10 years

Longer-term possibilities include more continuous psychosocial assessment, multimodal state estimation, and hybrid systems that coordinate automated support with human care. These remain high-uncertainty trajectories. Their development will depend on technical reliability but also on regulation, cultural acceptance, institutional resources, payment structures, data governance, and changing expectations of human relationships in care. No probability or inevitability is assigned to these futures.

Risks, weak signals, and potential discontinuities

Current risks include privacy failures, inaccurate generated content, biased proxies, unequal representation, surveillance burden, cultural mismatch, and redistribution of clinical workload. Weak signals concern more extensive automation of psychosocial interpretation and increasingly autonomous intervention selection. Potential discontinuities would arise if

computational systems began reorganizing where therapeutic responsibility resides rather than merely assisting existing care.

These categories require different evidentiary treatment. Demonstrated bias or privacy failure is not equivalent to

a speculative future risk. **Table 4** therefore separates mechanism, context, evidence maturity, and the condition that would make each signal clinically consequential.

Table 4. Mechanism–context crosswalk for risks, weak signals, and potential discontinuities in future psychosocial healthcare

Mechanism of concern	Context amplifying risk	Current evidentiary status	Condition requiring escalation
Signal or boundary case — Behavioral surveillance			
Continuous data collection exceeds clinically necessary interpretation	Low privacy control or unclear data use	Observed concern with established privacy precedent	Persistent collection without transparent purpose or meaningful consent
Signal or boundary case — Generative communication error			
Fluent output obscures omission or incorrect information	Complex or high-risk clinical communication	Directly observed limitation	Output influences action without verification or escalation
Signal or boundary case — Algorithmic inequity			
Proxies or datasets encode unequal care histories	Structurally unequal access and underrepresentation	Directly demonstrated in health algorithms	Differential allocation or harm persists after deployment
Signal or boundary case — Relational displacement			
Automated interaction substitutes for needed human interpretation	High vulnerability, uncertainty, conflict, or cultural mismatch	Conditional / incompletely established	Human repair or escalation becomes unavailable
Signal or boundary case — Autonomous psychosocial decision systems			
Prediction is converted directly into intervention selection	Continuous sensing combined with insufficient oversight	Weak signal / prospective discontinuity	Systems act on inferred states without validated decision pathways

Ethical and governance challenges

Future governance must connect privacy, fairness, transparency, responsibility, and relational accountability. Bias cannot be solved solely by balancing datasets, because inequity may originate in proxies, access, measurement, institutional practices, or downstream resource allocation. Governance should therefore specify intended use, population boundaries, human oversight, error escalation, data provenance, and responsibility for consequential decisions. Patient trust should be treated as an outcome of governance practice rather than assumed from consent or technical accuracy.

Research infrastructure and methodological priorities

Priority infrastructure includes longitudinal and externally validated datasets, transparent documentation of data provenance, multimodal measurement standards, prospective evaluation of human-AI workflows, and participatory research capable of identifying cultural and accessibility failures. Adaptive interventions require

designs that separate state prediction from intervention effect. Equity analyses require examination of who is absent as well as subgroup performance among those observed. Future evidence should increasingly connect technical, behavioral, relational, organizational, and clinical endpoints without collapsing them.

A future psychosocial healthcare research agenda

The proposed agenda centers on five questions: when digital behavior validly represents changing psychosocial states; when adaptive intervention improves outcomes rather than merely increasing personalization; which AI functions preserve or enhance relational care; how structural conditions alter data generation and treatment opportunity; and which organizational arrangements make human-digital systems accountable and sustainable.

Research should test these questions longitudinally, comparatively, and across culturally and structurally different settings. The principal unit of progress should

be a validated relationship among measurement, meaning, action, context, and consequence, not technological capability alone.

Strengths and limitations of the horizon scan

This review deliberately integrates evidence types that conventional effectiveness reviews often keep separate, allowing technical, psychological, relational, equity, and organizational signals to be examined together. That breadth also limits inference. Horizon selection is interpretive, evidence maturity differs markedly among domains, and rapidly changing technologies can outpace published evaluation. The targeted scan was not designed to estimate publication prevalence, pooled effectiveness, or probability of future adoption. Proposed horizons therefore require prospective testing rather than retrospective confirmation.

Conclusion

The future of psychosocial healthcare research is unlikely to be defined by artificial intelligence, behavioral sensing, precision intervention, equity science, or human-centered design in isolation. Its more consequential frontier lies in how these domains interact. Progress will require simultaneous validation of what technologies measure, what their outputs mean, when action is justified, who can participate equitably, how relationships are affected, and where responsibility resides. The most valuable future systems will therefore be those whose computational capabilities remain empirically grounded, psychologically interpretable, relationally accountable, structurally fair, and clinically governable.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*. 2021;20(3):318-35. doi:10.1002/wps.20883
2. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: Systematic review. *J Med Internet Res*. 2019;21(7):e12869. doi:10.2196/12869
3. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: A timely opportunity to consider purpose, quality, and safety. *NPJ Digit Med*. 2019;2:88. doi:10.1038/s41746-019-0166-1
4. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Ann Behav Med*. 2018;52(6):446-62. doi:10.1007/s12160-016-9830-8
5. Hines P, Yu LH, Guy RH, Brand A, Papaluca-Amati M. Scanning the horizon: A systematic literature review of methodologies. *BMJ Open*. 2019;9(5):e026764. doi:10.1136/bmjopen-2018-026764
6. Garcia Gonzalez-Moral S, Beyer FR, Oyewole AO, Richmond C, Wainwright L, Craig D. Looking at the fringes of MedTech innovation: A mapping review of horizon scanning and foresight methods. *BMJ Open*. 2023;13(9):e073730. doi:10.1136/bmjopen-2023-073730
7. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: How informatics interventions can worsen inequality. *J Am Med Inform Assoc*. 2018;25(8):1080-8. doi:10.1093/jamia/ocy052
8. Huckvale K, Torous J, Larsen ME. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA Netw Open*. 2019;2(4):e192542. doi:10.1001/jamanetworkopen.2019.2542
9. Hua Y, Na H, Li Z, Liu F, Fang X, Clifton D, et al. A scoping review of large language models for generative tasks in mental health care. *NPJ Digit Med*. 2025;8(1):230. doi:10.1038/s41746-025-01611-4
10. Wang X, Zhou Y, Zhou G. The application and ethical implication of generative AI in mental health: Systematic review. *JMIR Ment Health*. 2025;12:e70610. doi:10.2196/70610
11. Wang L, Bhanushali T, Huang Z, Yang J, Badami S, Hightow-Weidman L. Evaluating generative AI in

- mental health: Systematic review of capabilities and limitations. *JMIR Ment Health*. 2025;12:e70014. doi:10.2196/70014
12. Zhang Q, Zhang R, Xiong Y, Sui Y, Tong C, Lin FH. Generative AI mental health chatbots as therapeutic tools: Systematic review and meta-analysis of their role in reducing mental health issues. *J Med Internet Res*. 2025;27:e78238. doi:10.2196/78238
 13. Campellone TR, Flom M, Montgomery RM, Bullard L, Pirner MC, Pavez A, et al. Safety and user experience of a generative artificial intelligence digital mental health intervention: Exploratory randomized controlled trial. *J Med Internet Res*. 2025;27:e67365. doi:10.2196/67365
 14. Tong ACY, Wong KTY, Chung WWT, Mak WWS. Effectiveness of topic-based chatbots on mental health self-care and mental well-being: Randomized controlled trial. *J Med Internet Res*. 2025;27:e70436. doi:10.2196/70436
 15. Zaretsky J, Kim JM, Baskharoun S, Zhao Y, Austrian J, Aphinyanaphongs Y, et al. Generative artificial intelligence to transform inpatient discharge summaries to patient-friendly language and format. *JAMA Netw Open*. 2024;7(3):e240357. doi:10.1001/jamanetworkopen.2024.0357
 16. Small WR, Wiesenfeld B, Brandfield-Harvey B, Jonassen Z, Mandal S, Stevens ER, et al. Large language model-based responses to patients' in-basket messages. *JAMA Netw Open*. 2024;7(7):e2422399. doi:10.1001/jamanetworkopen.2024.22399
 17. Tai-Seale M, Baxter SL, Vaida F, Walker A, Sitapati AM, Osborne C, et al. AI-generated draft replies integrated into health records and physicians' electronic communication. *JAMA Netw Open*. 2024;7(4):e246565. doi:10.1001/jamanetworkopen.2024.6565
 18. Nong P, Platt J. Patients' trust in health systems to use artificial intelligence. *JAMA Netw Open*. 2025;8(2):e2460628. doi:10.1001/jamanetworkopen.2024.60628
 19. Akre S, Seok D, Douglas C, Aguilera A, Carini S, Dunn J, et al. Advancing digital sensing in mental health research. *NPJ Digit Med*. 2024;7(1):362. doi:10.1038/s41746-024-01343-x
 20. Balliu B, Douglas C, Seok D, Shenhav L, Wu Y, Chatzopoulou D, et al. Personalized mood prediction from patterns of behavior collected with smartphones. *NPJ Digit Med*. 2024;7(1):49. doi:10.1038/s41746-024-01035-6
 21. Beltrán JM, Jacob Y, Mehta MM, Hossain T, Adams A, Fontaine S, et al. Digital measures of activity and motivation impact depression and anxiety in the real world. *NPJ Digit Med*. 2025;8(1):268. doi:10.1038/s41746-025-01669-0
 22. Ringwald WR, King G, Vize CE, Wright AGC. Passive smartphone sensors for detecting psychopathology. *JAMA Netw Open*. 2025;8(7):e2519047. doi:10.1001/jamanetworkopen.2025.19047
 23. Grimm B, Yilmam P, Talbot B, Larsen L. Multimodal machine learning for video based single question mental health assessment. *NPJ Digit Med*. 2026;9(1):65. doi:10.1038/s41746-025-02242-5
 24. Nahum-Shani I, Murphy SA. Just-in-time adaptive interventions: Where are we now and what is next? *Annu Rev Psychol*. 2026;77(1):679-703. doi:10.1146/annurev-psych-121024-044244
 25. von Lützow U, Neuendorf NL, Scherr S. Effectiveness of just-in-time adaptive interventions for improving mental health and psychological well-being: A systematic review and meta-analysis. *BMJ Ment Health*. 2025;28(1):e301641. doi:10.1136/bmjment-2025-301641
 26. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-53. doi:10.1126/science.aax2342
 27. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—Reconsidering the use of race correction in clinical algorithms. *N Engl J Med*. 2020;383(9):874-82. doi:10.1056/NEJMms2004740
 28. Alderman JE, Palmer J, Laws E, McCradden MD, Ordish J, Ghassemi M, et al. Tackling algorithmic bias and promoting transparency in health datasets: The STANDING Together consensus recommendations. *Lancet Digit Health*. 2025;7(1):e64-e88. doi:10.1016/S2589-7500(24)00224-3
 29. Mihan A, Pandey A, Van Spall HGC. Mitigating the risk of artificial intelligence bias in cardiovascular care. *Lancet Digit Health*. 2024;6(10):e749-e754. doi:10.1016/S2589-7500(24)00155-9
 30. Robinson A, Flom M, Forman-Hoffman VL, Histon T, Levy M, Darcy A, et al. Equity in digital mental health interventions in the United States: Where to

- next? J Med Internet Res. 2024;26:e59939. doi:10.2196/59939
31. DesRoches CM, Wachenheim D, Garcia A, Harcourt K, Henry J, Shah R, et al. Clinician and patient perspectives on the exchange of sensitive social determinants of health information. JAMA Netw Open. 2024;7(10):e2444376. doi:10.1001/jamanetworkopen.2024.44376
 32. Rising KL, Guth A, Gentsch AT, Martin Gonzalez K, Hass R, Shughart L, et al. Development and preliminary validation of a screener for digital health readiness. JAMA Netw Open. 2024;7(9):e2432718. doi:10.1001/jamanetworkopen.2024.32718
 33. Dallison S, Moffett B, Makhubela P, Nkuna T, Pozuelo JR, van Heerden A, et al. Using generative AI to co-design digital mental health interventions with adolescents in rural South Africa: Qualitative thematic analysis of participatory workshops. J Med Internet Res. 2025;27:e73535. doi:10.2196/73535
 34. Mohr DC, Azocar F, Bertagnolli A, Choudhury T, Chrisp P, Frank R, et al. Banbury Forum consensus statement on the path forward for digital mental health treatment. Psychiatr Serv. 2021;72(6):677-83. doi:10.1176/appi.ps.202000561
 35. Bracic A, Spector-Bagdady K, Towle S, Zhang R, James CA, Price WN 2nd. Factors for patient trust and acceptance of medical artificial intelligence. JAMA Netw Open. 2026;9(3):e260815. doi:10.1001/jamanetworkopen.2026.0815
 36. You JG, Dbouk RH, Landman A, Ting DY, Dutta S, Wang JC, et al. Ambient documentation technology in clinician experience of documentation burden and burnout. JAMA Netw Open. 2025;8(8):e2528056. doi:10.1001/jamanetworkopen.2025.28056