

Evolutionary Game Theory for Patient Centricity & Policymaking in Generative-AI-Augmented Telemental Health

Prasad Kothari*^{ID}

Independent Researcher Digital Health & Distributed Ledger Technologies, India.

Corresponding Author: Prasad Kothari, Independent Researcher Digital Health & Distributed Ledger Technologies, India.

Received: 📅 2026 July 29

Accepted: 📅 2026 Aug 18

Published: 📅 2026 Aug 31

Abstract

Objective: To develop a mathematically rigorous, four-population evolutionary-game-theoretic (EGT) framework that specifies the structural conditions under which patient-centric behavior in generative-AI (GenAI)-augmented telemental health is a stable, dominant equilibrium of the joint patient-provider-payer-policymaker system, and to embed this contribution within a comprehensive narrative survey of the lineage of evolutionarily-stable-strategy (ESS) theorists, from Fisher's frequency-dependent selection (1930) through Maynard Smith and Price's 1973 founding, the Taylor-Jonker replicator dynamic, the Hofbauer-Sigmund synthesis, the Foster-Young and Nowak finite-population programs, the graph- and structured-population theories of Lieberman, Ohtsuki, Hauert, Antal, Tarnita and Allen, and the algorithmic-strategic-interaction work of Hardt, Liu and Obermeyer, through to the 2023–2026 reassessment by Leimar-McNamara and Traulsen-Glynatsi.

Design: Theoretical modeling study with structured narrative survey. We construct a $3 \times 3 \times 3 \times 3$ asymmetric evolutionary game in which patients, providers, payers and policymakers each select among three behavioral strategies, with state-dependent payoffs incorporating clinical outcomes, patient-reported experience and outcome measures (PREMs / PROMs), digital-equity protection, generative-AI safety incidents specific to mental-health applications, and operational costs.

Methodology: Strategy is governed by the multi-population replicator-mutator dynamic. We derive evolutionarily stable strategy (ESS) conditions in closed form and prove local asymptotic stability of the patient-centric equilibrium via a Kullback-Leibler-divergence Lyapunov function and LaSalle's invariance principle. Four policy levers are formalized: the outcome-linked reimbursement weight β , the equity weight λ , the platform-redesign cost c_p , and the regulatory enforcement intensity τ for GenAI mental-health tools. Parameter ranges were calibrated against published telemental-health, AI-safety and value-based-payment literature (2014–2025). No human subjects were enrolled and no original empirical data were collected.

Setting: Conceptual modeling and narrative review. No clinical setting; no patients enrolled.

Main Outcome Measures: Existence and local stability of the patient-centric ESS E_C ; closed-form thresholds (β^* , λ^* , c_p^* , τ^*); qualitative regime classification (transactional, fragmented, mixed, patient-centric); and ten policy recommendations derivable from the threshold inequalities.

Results: The four-population framework yields four qualitatively distinct equilibria: a transactional equilibrium E_T , a fragmented equilibrium E_F , a mixed equilibrium E_X , and a patient-centric equilibrium E_C . We prove (Proposition 4) that E_C is a strict ESS if and only if $\beta \geq \beta^*$, $\lambda \geq \lambda^*$, $c_p \leq c_p^*$, and $\tau \geq \tau^*$, each threshold expressible in closed form in terms of the underlying payoff structure. Under our calibration, $\beta^* \approx 0.45$, $\lambda^* \approx 0.30$, $c_p^* \approx 0.20$, $\tau^* \approx 0.35$. A Lyapunov-style argument (Theorem 1) shows that trajectories starting in a neighborhood of E_C converge to E_C under the replicator dynamic whenever all four threshold conditions hold strictly.

Conclusions: Patient centricity in GenAI-augmented telemental health is best characterized as a co-evolutionary property of patients, providers, payers and policymakers, not as a unilateral attribute of any one actor. The framework yields ten concrete, threshold-derived policy recommendations and is intended to guide subsequent empirical estimation, payer contract design, and regulatory rule-making.

Keywords: Finite Element Analysis (FEA), Biomechanical Analysis, Fracture Risk Assessment, Natural Frequencies, Bone Density

Telemental health mental-health care delivered through video, phone or text has grown rapidly since 2020, and generative-AI tools such as chatbots, vision language models and large language models are now being woven into therapy apps, triage workflows and clinical documentation. Almost every health organization claims to be "patient-centered." In practice, patients, clinicians, insurance payers and government regulators often want different things. Patients want care that fits their lives; clinicians face heavy schedules and unclear rules for using AI; payers reimburse along incentives that often reward visit volume rather than outcomes; and regulators are still deciding how to oversee AI in mental health.

This paper builds a mathematical theory—using evolutionary game theory—to ask one precise question: under what conditions does patient-centered behavior become the stable, default way of delivering AI-supported telemental health, rather than a niche preference? We model the system as four groups (patients, providers, payers, policymakers) that each choose among three behavioral strategies, and we calculate when the patient-centered outcome is an equilibrium that resists drift back toward less patient-centered alternatives. We identify four "policy knobs" that must all be turned high enough at the same time: (1) tying a meaningful share of payment to patient-reported outcomes; (2) weighting equity so vulnerable populations are protected; (3) keeping platform-redesign costs manageable; and (4) sustaining a credible level of regulatory enforcement for AI in mental health. We derive the threshold value of each knob in closed mathematical form, prove a stability theorem, and translate the math into ten specific recommendations for policymakers, payers, and platform designers. The paper also surveys nearly a century of evolutionary game theory work, from Fisher's frequency-dependent selection in 1930 through to research published in 2024–2026.

1. Introduction

1.1. Telemental Health and the Rise of Generative AI

Telemedicine has moved from a niche modality to a structural feature of healthcare delivery worldwide [1–6]. Within telemedicine, telemental health has emerged as one of the most durable post-pandemic survivors: by 2024 video- and audio-based mental-health visits accounted for roughly half of all behavioral-health encounters in the United States, with comparable trajectories in Europe, Australia and parts of Asia [7–9]. Two converging trends now shape the next decade of the field. The first is the integration of generative artificial intelligence (GenAI) large language models, conversational agents and multimodal systems into mental-health workflows, ranging from triage chatbots and asynchronous symptom screening to ambient documentation, clinical decision support, and patient-facing companions that deliver elements of cognitive-behavioral therapy [10–14]. The second is the corresponding expansion of regulatory attention, with the United States Food and Drug Administration formalizing pathways for AI/ML-based software as a medical device and the European Union enacting the world's first comprehensive AI regulation [15, 16].

In parallel, "patient centricity" has become a near-universal aspirational claim of healthcare organizations, payers and digital-health vendors [17–19]. Operationalized definitions converge on respect for patient preferences, coordination of services, shared decision-making, and outcomes that include patient-reported experience and quality of life rather than biomarkers alone [20–25]. The empirical record on whether telemedicine actually improves patient centricity is mixed: reviews report gains in access, convenience and chronic-disease self-management, while disparities along the lines of age, race, language, rurality and broadband availability remain pronounced [26–33]. For GenAI tools specifically the evidence is even earlier: pre-registered randomized trials of therapeutic chatbots are emerging, real-world deployment data show heterogeneous adherence and outcome patterns, and the patient-safety, privacy and equity risks are well documented [34–37].

1.2. Why Patient Centricity is Structurally Difficult: Four Mis-Aligned Actors

A central reason for the mixed empirical record is that patient centricity is not a unilateral attribute of any single actor in the telemental-health ecosystem. It is a joint outcome of at least four behavioral populations:

- **Patients:** (Engagement, adherence, self-monitoring, willingness to disclose to a machine).
- **Providers:** (Listening, shared decision-making, asynchronous responsiveness, willingness to integrate AI into care).
- **Payers:** (Reimbursement design, parity for asynchronous and audio-only modalities, equity safeguards, platform contracting).
- **Policymakers / Regulators:** (Definition of "software as a medical device," audit cadence, enforcement intensity, equity reporting requirements, mental-health parity enforcement).

Each actor faces its own incentives and costs. Patients trade convenience against effort and privacy. Providers face heavy schedules, uneven payment rules, and new liabilities around AI use. Payers face the structural pull of visit-volume metrics that are easier to measure than outcomes. Policymakers face the political feasibility of enforcement against well-resourced platforms and the practical limits of regulatory bandwidth. When these incentives pull in different directions, exhortations to "be patient-centered" produce disappointing returns [38, 39].

1.3. Why EGT Framework?

Existing analyses of telemental-health policy are predominantly empirical or descriptive. What is missing is a mathematically explicit theory of when patient centricity is structurally feasible—when it is a stable equilibrium that resists invasion by less

patient-centric alternatives, and when it is not. Without such a theory, policy debates default to exhortation and incremental reforms produce predictable failures [38, 39]. Evolutionary game theory (EGT), introduced by Maynard Smith and Price and developed by Taylor and Jonker, Hofbauer and Sigmund, Weibull Sandholm and Nowak, provides the analytical apparatus required [40–47]. Unlike classical Nash analysis, EGT does not assume hyper-rational actors; it models how the frequency of strategies in a population changes over time as a function of payoff differentials, capturing imitation, learning and boundedly-rational adjustment [48, 49]. Its central stability concept—the evolutionarily stable strategy (ESS)—asks whether a strategy profile resists invasion by small fractions of mutants, which is exactly the question one wants to ask of patient-centric care.

1. 4. Research Questions and Contributions

This paper addresses four questions:

- **RQ1:** Under what structural conditions on reimbursement design, equity weighting, platform-redesign cost, and regulatory enforcement intensity is a patient-centric strategy profile an evolutionarily stable strategy in a four-population GenAI-augmented telemental-health model?
- **RQ2:** How do those conditions translate into closed-form, policy-relevant thresholds that can be related to observable proxy metrics?
- **RQ3:** What does the framework imply about equilibrium selection: which equilibrium does the ecosystem reach when conditions are not jointly satisfied, and what does this imply for the sequencing of policy reforms?
- **RQ4:** How does the framework relate to seventy years of game-theoretic work on healthcare, from Nash's foundational equilibrium concept through information-economic models of insurance markets, principal-agent models of physician payment, public-goods models of vaccination and antimicrobial stewardship, mechanism-design models of organ exchange, and emerging AI-policy game-theoretic models?

The paper makes five contributions: (i) a four-population EGT model of GenAI-augmented telemental health, extending the three-population analyses found in prior work by adding policymakers as an explicit strategic population; (ii) closed-form ESS thresholds for the patient-centric equilibrium in terms of four policy levers; (iii) a local-stability theorem under the replicator dynamic with a Lyapunov-style proof; (iv) a comprehensive narrative survey of game-theoretic models applied to healthcare since 1950, organized along nine model families; and (v) ten policy recommendations derivable directly from the threshold inequalities, addressed to regulators, payers, and platform/provider organizations.

1.5. Roadmap

Section 2 reviews the foundations of patient centrality, telemental-health effectiveness and equity, and the early evidence base on GenAI in mental health. Section 3 presents the lineage of ESS and evolutionary-game theorists from Fisher (1930) through the 2023–2026 research frontier, in sixteen subsections organized around the principal contributors. Section 4 develops the formal four-population model. Section 5 presents the analytical methods (ESS conditions, replicator dynamics, Lyapunov stability). Section 6 states and proves the main theoretical results. Section 7 interprets the results, derives ten policy recommendations grounded in the 2025–2026 GenAI mental-health regulatory landscape, and discusses strengths and limitations. Section 8 concludes. Disclosures, AI Use Declaration, Data Availability Statement, Appendix and references follow.

2. Background and Related Work

2.1. Patient Centrality: From Concept to Measurement

The modern conceptualization of patient-centered care traces to the Picker / Commonwealth Fund work of the early 1990s and was systematized in the Institute of Medicine's Crossing the Quality Chasm report [19, 22]. Subsequent frameworks emphasize six dimensions: respect for preferences, coordination, information and education, physical comfort, emotional support, and involvement of family and friends [17, 23]. Operationalization typically uses patient-reported experience measures (PREMs), patient-reported outcome measures (PROMs) and shared-decision-making metrics [20, 21, 24, 25]. We adopt the broader sense of "patient-centric" as whole-system organization around patient outcomes and participation, rather than the narrower sense of "patient-centered" provider behavior.

2.2. Telemental Health: Effectiveness and Equity

Telemental health has been studied across modalities (video, asynchronous messaging, remote symptom monitoring), conditions and settings. Systematic reviews indicate non-inferior or superior outcomes for telepsychiatry across major depressive disorder, anxiety disorders, PTSD and substance-use disorders, with effect sizes comparable to in-person care for many conditions [7, 50]. Videoconferencing psychotherapy specifically shows therapeutic-alliance ratings comparable to in-person therapy in randomized trials [50]. Cochrane reviews of interactive telemedicine more broadly report modest but reproducible effects on clinical outcomes across conditions [51]. Equity findings are more concerning. Studies during and after the COVID-19 surge document substantial disparities in telehealth access by age, race, language, rurality and broadband availability [30–33]. Older adults, patients with limited English proficiency and patients in low-broadband regions disproportionately receive audio-only rather than video visits, which historically were reimbursed at lower rates [30–33]. These findings motivate the explicit equity weight λ in our payoff structure (Section 4).

2.3. Generative AI in Mental Health: Early Evidence and Structural Concerns

Two streams of evidence on GenAI in mental health are now emerging. First, randomized and quasi-randomized trials of conversational agents have begun to appear. The Therabot trial of a generative-AI therapy chatbot reported reductions in depressive and anxiety symptoms relative to waitlist at four and eight weeks [34]. Earlier rule-based or hybrid systems such as Woebot and Wysa showed signals of benefit in pragmatic deployments, though with substantial heterogeneity in adherence and effect size [35, 36]. Sharma and colleagues demonstrated that human-AI collaboration can increase empathic response quality in peer-to-peer mental-health support, with effect sizes comparable to those produced by trained human responders [52].

Second, a parallel literature documents structural concern. De Choudhury and colleagues catalogue benefits and harms of large language models in digital mental health, including risks of inaccurate crisis-detection, reinforcement of harmful self-talk patterns, and privacy leakage [37]. Lawrence and colleagues formalize the opportunity-risk trade-off and propose safety guardrails [14]. Bickman's fifty-year retrospective on mental-health services research places GenAI tools within a longer trajectory of measurement-based and precision approaches to mental-health care, and Torous and colleagues survey the broader digital-psychiatry field [53, 54]. Stade and colleagues propose a framework for responsible development and evaluation of large language models in behavioral healthcare [13]. Algorithmic-fairness concerns specific to clinical AI are also well documented: Obermeyer and colleagues showed how a widely deployed risk-prediction algorithm systematically under-identified Black patients with high care needs, an empirical anchor for the equity-weighting term in our model [54].

2.4. Regulatory and Governance Context

The regulatory environment for GenAI in mental health is heterogeneous and rapidly evolving. The U.S. FDA has issued action plans for AI/ML-based software as a medical device but has not yet finalized a comprehensive rule for generative systems [15]. The European Union's AI Act, in force from 2024, classifies certain healthcare AI applications as "high-risk" and imposes obligations around risk management, data governance, transparency and human oversight [16]. Audit frameworks for large language models have been proposed but not standardized. Governance frameworks specific to generative AI in clinical decision support emphasize audit cadence, equity safeguards and standardized incident reporting. The enforcement intensity τ in our model (Section 4) is the abstract variable corresponding to the credibility and frequency of this regulatory apparatus.

2.5. Game Theory in Healthcare: A Pointer to the Comprehensive Survey

Game-theoretic analyses of healthcare span insurance markets, physician payment, hospital competition, antimicrobial stewardship, vaccination behavior, organ allocation and emerging AI-policy interactions. To our knowledge, no prior work models GenAI-augmented telemental-health patient centrality as a four-population evolutionary game. The closest works are single-population EGT models of patient adherence and two-player static games of provider-platform interaction [55]. We provide a comprehensive narrative survey of the field in Section 3 before developing our framework in Section 4.

3. A Lineage of ESS and Evolutionary Game Theorists, 1930–2026

This section traces the people and ideas that built evolutionary game theory (EGT) from its statistical-genetic precursors in the 1930s to the structured-population, finite-size, and human-cooperation theory of the present decade. The framing is deliberate: rather than centering Nash's 1950 existence theorem as the prime mover, we treat Nash's contribution as one tributary that fed into a much wider river. The wider river is the one Maynard Smith and Price tapped in 1973 and that—through Taylor, Jonker, Zeeman, Selten, Friedman, Hofbauer, Sigmund, Weibull, Sandholm, Foster, Young, Kandori, Mailath, Rob, Schlag, Nowak, Sasaki, Fudenberg, Lieberman, Ohtsuki, Hauert, Antal, Tarnita, Traulsen, Allen, Szabó, Perc, Bauch, Galvani, Fu, Bowles, Gintis, Rand, Adami, Hintze, Leimar, McNamara, and many others—has carried the theory to where it stands in 2026. The aim is to give EGT the substantive intellectual treatment it deserves and to position the present four-population framework precisely within that lineage.

3.1. Precursors: Fisher, Lewontin, Hamilton (1930–1967)

Although evolutionary game theory is conventionally dated to 1973, its conceptual scaffolding was being assembled four decades earlier. R. A. Fisher's *Genetical Theory of Natural Selection* introduced frequency-dependent selection as a foundation of evolutionary biology and gave the sex-ratio argument in which what is optimal for an individual depends on what the rest of the population is doing—the prototype of a game-theoretic argument in evolution [56]. Fisher did not invoke game theory by name, but the logic he applied to sex allocation is recognizably the logic that Maynard Smith and Price would formalize four decades later. Richard Lewontin's 1961 paper, sharing the title "Evolution and the Theory of Games" with Maynard Smith's later book, made the first explicit attempt to use game theory in evolutionary biology [57]. Lewontin modeled populations as playing games against "nature" (variable environments) and applied minimax reasoning to polymorphism as adaptation. The framing—population versus environment rather than individual versus individual—turned out to be the wrong one for most biological problems, and by the mid-1960s Lewontin had largely abandoned the program. Yet his attempt established a vocabulary, and when Maynard Smith reconceived the problem as one of individuals competing within a population, Lewontin's earlier work was both negative example and prepared ground. Lewontin himself

contributed a mixed-strategy ESS existence proof during a 1972 visit to Maynard Smith at Sussex, even as his published reviews remained skeptical.

W. D. Hamilton's 1967 paper on extraordinary sex ratios was the first to use explicitly game-theoretic language ("unbeatable strategy") in an evolutionary biology argument [58]. Hamilton's unbeatable strategy in the local-mate-competition setting is essentially the same object as Maynard Smith and Price's ESS, six years earlier. Hamilton's kin-selection inequality, developed in parallel, supplied another piece of the strategic infrastructure: it identified the precise conditions under which apparently altruistic behavior is in fact a best response in an extended game that includes relatives. Hamilton, Maynard Smith, and Price became the second wave of evolutionary game theorists in the 1960s and 1970s, applying game theory at the level of individual phenotypes rather than against environments.

3.2. The 1973 Founding: Maynard Smith and Price

Maynard Smith and Price's 1973 paper "The Logic of Animal Conflict" is the conventional foundation. The biological problem was specific: why do intraspecific animal contests rarely escalate to lethal violence? Classical Darwinian reasoning struggled here, since if escalation paid in any single contest, escalators should outcompete non-escalators. Maynard Smith and Price's insight was that a strategy's payoff depends on the frequencies of strategies in the rest of the population. In a population of all hawks (always escalate), expected payoffs are low because injury costs swamp victory benefits, whereas a rare dove (display only) does relatively well by avoiding fights. The result is a stable mixed equilibrium whose composition depends on the ratio of victory benefit to injury cost.

Maynard Smith's 1974 follow-up paper and his 1982 book *Evolution and the Theory of Games* worked out the apparatus [59, 41]. The book established ESS analysis as a standard method in theoretical biology and is still the canonical reference. Formally, σ^* is an ESS if, for every alternative $\sigma \neq \sigma^*$, either (i) the expected payoff of σ^* against σ^* strictly exceeds that of σ against σ^* , or (ii) payoffs are tied at σ^* but σ^* strictly outperforms σ against σ . The first clause is the strict Nash condition; the second clause—about the response to the mutant itself—is what makes ESS strictly stronger than Nash and what gives it its dynamic content. George Price's contribution was both technical and conceptual; Price's tragic biography (suicide in 1975 after his work on his eponymous equation and on altruism) is part of the field's cultural memory.

Two features of the ESS are routinely missed in applied work. First, every strict Nash equilibrium of a symmetric two-player game is an ESS, but not every ESS is a strict Nash equilibrium of the underlying one-shot game; the ESS concept is intrinsically about populations playing many independent contests. Second, the ESS as Maynard Smith defined it is a static

3.3. The Replicator Dynamic: Taylor, Jonker, Zeeman, Bishop, Cannings

Taylor and Jonker introduced the replicator equation in 1978 [42]. With $x_i(t)$ the frequency of strategy i and $f_i(x)$ the expected payoff in state x , the replicator dynamic is

$$\frac{dx_i}{dt} = x_i [f_i(x) - \bar{f}(x)], \text{ where } \bar{f}(x) = \sum_j x_j f_j(x).$$

Strategies that earn above-average payoff grow in frequency; those that earn below-average shrink; total frequency is preserved. The Taylor–Jonker equation is the canonical dynamic for a well-mixed asexual population in which fitness translates into reproductive success, and it is also equivalent—under mild assumptions—to the simplest model of imitation learning in human populations. The same equation governs biological reproduction and social imitation, which is why EGT travels so easily between disciplines [44, 60, 61]. Three properties of the dynamic deserve emphasis: (i) the simplex $\Delta^{(n-1)}$ is forward-invariant; (ii) pure-strategy vertices are fixed points, but only some are stable; and (iii) by a coordinate change, the replicator equation is equivalent to the Lotka–Volterra equations of mathematical ecology, importing the entire mathematical literature on predator–prey and competitive-exclusion dynamics.

E. C. Zeeman classified the phase portraits of three-strategy replicator dynamics geometrically in 1980, identifying eight qualitatively distinct types of behavior on the 2-simplex [62]. The list includes stable interior equilibria, heteroclinic cycles (the rock–paper–scissors structure), saddle-node configurations, and equilibria on the boundary that attract from the interior—each with a recognizable interpretation in applied work. Bishop and Cannings's "generalised war of attrition" developed mixed-ESS analysis for asymmetric contests and continuous strategy spaces, supplying machinery that recurs in biological modeling of waiting games and territorial disputes [63].

3.4. Asymmetric Games and the ESS–Nash Connection: Selten and Friedman

Reinhard Selten's 1980 paper on ESS in asymmetric animal conflicts proved a sharp and consequential result: in a two-population asymmetric game, an ESS must be a strict Nash equilibrium of the underlying game [64]. In pure-strategy terms this rules out mixed equilibria as ESSes of asymmetric games. The four-population framework we develop in Section 4 inherits this property—which is why our central result (Proposition 4) concerns the strict-ESS character of the pure patient-centric profile E_C rather than some mixed configuration. Selten, who also developed subgame perfection in classical non-cooperative theory and shared the 1994 Nobel with Nash and Harsanyi, occupies the unusual position of being a major contributor to both classical and evolutionary game theory.

Daniel Friedman's 1991 *Econometrica* paper, "Evolutionary Games in Economics", was a major reception event for EGT outside biology [65]. Friedman established that for a wide class of "compatible" dynamics that generalize the replicator equation—including the best-response dynamic, the logit dynamic, and the Smith dynamic—every ESS is a locally asymptotically stable rest point. This robustness result is the licence for using the replicator dynamic as the working model while interpreting predictions as approximately valid under a broader family of plausible adjustment mechanisms. Friedman's paper helped move EGT from a biological subspecialty into mainstream economics, where it now anchors the literature on bounded rationality, learning, and evolution of preferences.

3.5. The Hofbauer–Sigmund Synthesis and the Textbook Era: Weibull, Sandholm, Nowak

Josef Hofbauer and Karl Sigmund consolidated the mathematical theory in two reference works: the 1998 monograph *Evolutionary Games and Population Dynamics* and the widely cited 2003 Bulletin survey "Evolutionary game dynamics" [43, 44]. Two contributions we use directly in Section 6 originate here. First, Hofbauer and Sigmund established the Kullback–Leibler-divergence Lyapunov technique we deploy in Theorem 1: for many replicator systems $V(x(t)) = D_{KL}(x^* || x(t))$ from a candidate ESS x^* is monotonically decreasing along trajectories. Second, they showed that the replicator-mutator equation—obtained by adding a uniform mutation term $\mu(1/n - x_i)$ —is a special case of the quasispecies equation from molecular evolution, importing yet another mathematical literature into the EGT toolbox.

Jörgen Weibull's 1995 textbook *Evolutionary Game Theory* introduced the field to a generation of economists [45]. William Sandholm's 2010 monograph *Population Games and Evolutionary Dynamics*, together with his earlier encyclopedia article, remains the most thorough modern treatment, covering deterministic and stochastic dynamics, potential games, and the geometry of best-response and logit dynamics [46, 66]. Martin Nowak's 2006 book *Evolutionary Dynamics* connected EGT to evolutionary biology proper through finite-population stochastic methods (Section 3.6 below), and through the program of "five rules for the evolution of cooperation" that has organized much of the subsequent literature on cooperation [47].

3.6. Stochastic Evolutionary Game Dynamics: Foster, Young, Kandori, Mailath, Rob

The replicator dynamic is a deterministic, large-population approximation. In finite populations, or in any population subject to noise, strategies that the deterministic dynamic identifies as stable can still be displaced. Dean Foster and Peyton Young introduced the concept of stochastic stability in 1990: in a population subject to small mutation noise, the long-run distribution of play concentrates on the equilibria that minimize a particular potential function, which may differ from the equilibria the deterministic dynamic selects [67]. Michihiro Kandori, George Mailath and Rafael Rob's 1993 *Econometrica* paper showed that in 2×2 coordination games the risk-dominant equilibrium is selected by stochastic stability, even when the payoff-dominant equilibrium is a strict Nash equilibrium of the unperturbed game [68]. The same selection principle was independently developed by Young's school and refined throughout the 1990s.

Karl Schlag's 1998 paper provided a microfoundational result of considerable importance to anyone applying EGT to human populations: proportional imitation—copying the strategy of a randomly sampled neighbor with probability proportional to the payoff difference—yields exactly the replicator dynamic at the population level [69]. The cognitive demands of EGT, on Schlag's microfoundation, are minimal: not Bayesian rationality, not common knowledge of rationality, just enough comparison to a sampled neighbor to imitate when it pays. This is one reason EGT is the right modeling language for health-policy populations, where agents satisfy this minimal demand with room to spare. Lorens Imhof, Drew Fudenberg and Martin Nowak developed the analysis of evolutionary cycles of cooperation and defection in finite populations, extending Foster–Young stochastic stability to richer games [70]. Arne Traulsen and Christoph Hauert's chapter "Stochastic Evolutionary Game Dynamics" is the standard reference for the finite-population apparatus [71].

3.7. Finite Populations: Nowak, Sasaki, Taylor, Fudenberg, and the 1/3 Law

Nowak, Sasaki, Taylor and Fudenberg's 2004 *Nature* paper "Emergence of cooperation and evolutionary stability in finite populations" proved one of the most consequential results of modern EGT: the celebrated 1/3 law. In a finite population of size N evolving under a Moran process with weak selection, a strategy A is more likely to invade a population of B 's than vice versa if and only if A is a best response to a population playing B with frequency at least $1/3$ [60]. The 1/3 law identifies a regime in which evolutionary dynamics in small populations diverge qualitatively from the infinite-population replicator analysis. For healthcare modeling the implication is direct: in regional health systems, specialty practices, or pilot programs, finite-population effects matter, and ESS conclusions based on the infinite-population deterministic dynamic must be checked against the corresponding finite-population stochastic dynamics. We address this caveat explicitly in Section 7.4.

3.8. Spatial and Graph-Structured EGT: Nowak, May, Lieberman, Ohtsuki, Hauert, Szabó, Perc

Nowak and Robert May's 1992 *Nature* paper "Evolutionary games and spatial chaos" launched the study of spatial evolutionary games [72]. Placing agents on a lattice and updating them by imitation of best-performing neighbors produces qualitatively new behavior, including stable coexistence of cooperators and defectors in the Prisoner's Dilemma—an outcome impossible in the well-mixed setting. Erez Lieberman, Christoph Hauert and Martin Nowak's 2005 *Nature* paper "Evolutionary dynamics on graphs" generalized the analysis from regular lattices to arbitrary graphs and characterized the

conditions under which graph structure amplifies or suppresses selection [73]. Hisashi Ohtsuki, Hauert, Lieberman and Nowak's 2006 Nature paper derived a simple rule for the evolution of cooperation on graphs: cooperation is favored when the benefit-to-cost ratio exceeds the average number of neighbors [74]. György Szabó and Gábor Fátth's 2007 Physics Reports survey catalogues the entire literature on evolutionary games on graphs, and Matjaž Perc and Attila Szolnoki's review of coevolutionary games tracks the parallel development in which network structure itself co-evolves with strategy [75, 76].

The graph-EGT literature reached a milestone with Benjamin Allen and colleagues' 2017 Nature paper "Evolutionary dynamics on any population structure", which derived closed-form fixation probabilities for arbitrary weighted networks under weak selection [77]. The result subsumes a generation of partial results and establishes graph-structured EGT as a complete theory rather than a collection of special cases. Deterministic theory of evolutionary games on temporal networks—where the network itself changes in real time—is a 2024–2026 frontier active at the time of writing.

3.9. Beyond Graphs: Phenotype, Set, and Group-Structured Populations (Antal, Tarnita)

Tibor Antal, Hisashi Ohtsuki, John Wakeley, Peter Taylor and Martin Nowak's 2009 paper introduced evolutionary games in phenotype space, in which individuals interact preferentially with phenotypically similar others [78]. The framework explains a class of cooperation outcomes that neither well-mixed nor graph-structured models predict and is one of the principal generalizations of Hamilton's kin-selection logic to arbitrary phenotypic traits. Corina Tarnita, Antal, Ohtsuki and Nowak's companion paper developed evolutionary dynamics in set-structured populations, in which individuals belong to overlapping sets and interact only within shared sets [79]. The set-structured framework provides a natural model of multi-affiliation organizations—exactly the structure of healthcare systems, in which patients, providers, payers and policymakers operate across overlapping institutional and jurisdictional sets. Although we do not use the set-structured formalism in the present analysis, it is the natural extension for future work on multistate telemental-health regulation.

3.10. EGT in Human Cooperation: Bowles, Gintis, Rand, Nowak

A substantial post-2000 literature has used EGT to understand human cooperation specifically, with explicit attention to social-norm enforcement, reputation, and institutional incentives. Samuel Bowles and Herbert Gintis's *A Cooperative Species* synthesizes the experimental, anthropological and theoretical evidence for an evolutionary account of human prosociality, with EGT as the central mathematical language [80]. David Rand and Martin Nowak's *Trends in Cognitive Sciences* review catalogues the experimental findings on human cooperation in public-goods games, prisoner's dilemmas, and trust games, all interpreted through the EGT lens [81]. The substantive lessons of this literature—that humans cooperate more than narrow self-interest predicts; that conditional cooperation, punishment, and reputation are the principal mechanisms; that institutional design matters enormously—are the empirical motivation for treating patient and provider populations not as fully selfish but as boundedly rational imitators of payoff-superior strategies.

3.11. Computational and Agent-Based EGT: Adami, Hintze

As analytical methods reached their natural limits, computational and agent-based approaches expanded the reach of EGT to settings where strategies are themselves complex (encoded as finite-state machines, neural networks, or genomes). Christoph Adami, Jory Schossau and Arend Hintze's 2016 *Physics of Life Reviews* paper surveys the agent-based-EGT literature and identifies the conditions under which computational simulation yields conclusions consistent with—and where it diverges from—analytical replicator-dynamic results [82]. For healthcare, agent-based EGT is the natural extension when strategy spaces are too rich for analytical tractability, such as when GenAI platform behaviors are themselves learned and adaptive.

3.12. EGT in Disease, Vaccination, and Healthcare Behavior

Direct EGT applications to healthcare have concentrated in three substantive domains. The flagship is vaccination behavior. Chris Bauch and David Earn's 2004 paper "Vaccination and the theory of games" showed that under self-interested decision-making equilibrium vaccination coverage falls short of the social optimum because herd immunity allows free-riding [83]. Alison Galvani, Timothy Reluga and Gretchen Chapman demonstrated empirically that long-standing U.S. [84]. influenza-vaccination policy is consistent with individual self-interest but not the utilitarian optimum. Feng Fu, Daniel Rosenbloom, Long Wang and Nowak extended the analysis to imitation dynamics on social networks [85]. Reluga and Galvani developed the general population-game framework [86]. Zhen Wang and colleagues' *Physics Reports* review on the statistical physics of vaccination is the comprehensive reference, with Cardillo and colleagues on network effects and Jun Tanimoto's monograph on the broader sociophysics of evolutionary games and epidemics [87–89]. The methodological lessons—replicator dynamics, ESS analysis, network-structured imitation, and the coupling between epidemic and behavioral dynamics—translate directly into our framework.

The second domain is antimicrobial stewardship, where antibiotic prescribing has the structure of a tragedy-of-the-commons: each prescription confers a private benefit and contributes to the population-level externality of resistance. Klaus Kaier and Uwe Frank developed the explicit pharmacoeconomic model of this externality, and subsequent EGT analyses have examined when stewardship policies, peer-comparison feedback, and reimbursement adjustments shift prescribing populations toward narrower-spectrum strategies [90]. The third domain is patient adherence and treatment dynamics.

Khan, Khan and Zia-Ul-Haq developed a single-population EGT model of patient adherence in chronic disease, finding that adherence is co-determined by clinical effectiveness, regimen complexity, and peer-imitation effects [55]. EGT has also been applied to tumor evolution under treatment, modeling drug-sensitive and drug-resistant cancer-cell subpopulations as a strategic interaction—methodologically parallel to the present framework, substantively distinct. The unifying observation is that wherever a population of agents (cells, patients, providers, payers, policymakers) responds to differential payoffs through imitation, reproduction or learning, EGT supplies the natural mathematical language.

3.13. Algorithmic Strategic Interaction (2016–Present): Hardt, Liu, Obermeyer

A new and policy-relevant family analyzes strategic interaction in the presence of algorithmic decision-making. Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou and Mary Wootters's "strategic classification" framework model's agents who manipulate their features in response to a known classifier, producing equilibria in which classifier accuracy is degraded by gaming [91]. Lydia Liu and colleagues formalize delayed impact of fair machine learning, showing that interventions designed to satisfy short-run fairness criteria can produce welfare-reducing dynamics over longer horizons [92]. Ziad Obermeyer and colleagues' empirical study of racial bias in clinical risk-prediction algorithms supplied the canonical empirical anchor in healthcare [54]. This family is directly relevant to GenAI in mental health: patients, providers and platforms respond strategically to each other, and our framework incorporates the dynamic by treating regulatory enforcement intensity τ and platform-redesign cost c_p as policy levers that condition the population-level response.

3.14. Mechanism Design as the Constructive Inverse

Mechanism design—the constructive inverse of game theory—is also part of the lineage. Alvin Roth, Tayfun Sönmez and Utku Ünver's kidney-exchange papers and Roth and Elliott Peranson's redesign of the National Resident Matching Program together established mechanism design as a healthcare-relevant engineering discipline [59, 93, 94]. The framework's lesson—that institutions should be designed with strategic agents in mind, not against them—is the spirit in which we cast our four policy levers (Section 4) as a designed environment within which the patient-centric ESS becomes attainable.

3.15. The 2023 Fiftieth-Anniversary Reassessment and Open Problems Through 2026

Olof Leimar and John McNamara's 2023 Philosophical Transactions paper, "Game theory in biology: 50 years and onwards", and Arne Traulsen and Nikoleta Glynatsi's companion survey "The future of theoretical evolutionary game theory" frame the fiftieth-anniversary stock-taking of the field [95, 96]. Three open problem areas they identify—and that frame our contribution—deserve specific mention. Multi-population asymmetric games beyond two populations. The theory of two-population asymmetric games is well developed; the theory of three or more populations is substantially less so. ESS conditions in n -population games are a system of inequalities (one per population), and the basin-of-attraction structure under multi-population replicator dynamics can be intricate. Our four-population analysis (Sections 4–6) contributes a concrete case in which closed-form thresholds can be derived; the broader theory of which classes of n -population asymmetric games admit clean threshold characterizations remains active. Coupling between dynamics on multiple timescales. Patient strategies adjust on months; provider strategies on years; payer contracts on multi-year cycles; regulatory regimes on years to decades.

The standard replicator analysis assumes a single timescale. Slow-fast decomposition methods are available [46], but applying them to four-population health-policy games is non-routine. We use a single-timescale formulation in this paper, with explicit acknowledgment of the limitation (Section 7.4). Stochastic perturbations and incident-driven shocks. Real health systems experience rare, high-salience events—a publicized AI safety incident, a malpractice verdict, a policy reversal—that act as large stochastic shocks rather than the small mutation noise of standard stochastic-stability theory. Modeling such shocks formally is an open methodological problem; the November 2025 FDA Digital Health Advisory Committee meeting and the 2025 state laws discussed in Section 7 are real examples of exactly this category of perturbation.

3.16. What is Missing, and what this Paper Supplies

The lineage just traced—from Fisher's frequency-dependent sex-ratio argument through Hamilton's unbeatable strategy, Lewontin's first explicit attempt, Maynard Smith and Price's 1973 founding, the Taylor–Jonker replicator dynamic, Zeeman's geometric classification, Selten's asymmetric-ESS theorem, Friedman's reception into economics, the Hofbauer–Sigmund synthesis, the Weibull and Sandholm textbooks, the Foster–Young and Kandori–Mailath–Rob stochastic-stability program, Schlag's microfoundation, the Nowak–Sasaki–Taylor–Fudenberg finite-population law, Nowak and May's spatial games, the Lieberman–Hauert–Ohtsuki–Allen graph-structured theory, the Antal–Tarnita phenotype- and set-structured extensions, the Bauch–Galvani–Fu–Wang–Tanimoto vaccination literature, the Hardt–Liu–Obermeyer algorithmic-strategic literature, and the 2023 Leimar–McNamara and Traulsen–Glynatsi reassessments—has produced a remarkably complete mathematical, empirical and computational toolkit. What none of these contributors built, and what this paper supplies, is a unified four-population evolutionary framework that (i) treats patients, providers, payers and policymakers symmetrically as strategic populations, (ii) derives closed-form ESS thresholds in terms of recognizable policy levers, (iii) embeds the specific structural features of GenAI-augmented telemental health (algorithmic-safety incidents, equity-weighting, enforcement-intensity asymmetries), and (iv) is testable in principle by empirical estimation of the payoff parameters from administrative data. Table 1 summarizes the lineage and its relationship to the present framework.

Period and contributors	Principal contribution	Relationship to present framework
Precursors (1930–1967): Fisher [111]; Lewontin [112]; Hamilton [113]	Frequency-dependent selection; first explicit population-game model; "unbeatable strategy" as proto-ESS	Establish strategic vocabulary; motivate population-level analysis
1973 founding: Maynard Smith & Price [47]; Maynard Smith [48, 72]	ESS concept; hawk–dove logic; Evolution and the Theory of Games	Direct intellectual parent of patient-centric-ESS analysis
Replicator dynamic (1978–1980): Taylor & Jonker [49]; Zeeman [50]; Bishop & Cannings [60]	Replicator equation; phase-portrait classification; war-of-attrition machinery	Supplies the dynamic apparatus used in Section 5
Asymmetric games (1980, 1991): Selten [45]; Friedman [46]	Asymmetric-ESS = strict Nash; robustness across compatible dynamics	Justifies pure-strategy focus and replicator-dynamic robustness in Section 6
Modern synthesis (1995–2010): Hofbauer & Sigmund [51, 52]; Weibull [53]; Sandholm [54]; Nowak [55]	KL-divergence Lyapunov methods; replicator–mutator; textbook consolidation	Lyapunov technique used in Theorem 1; standard reference frame
Stochastic & finite-population (1990–2009): Foster & Young [61]; Kandori–Mailath–Rob [62]; Schlag [63]; Nowak–Sasaki–Taylor–Fudenberg [56]; Imhof–Fudenberg–Nowak [64]; Traulsen & Hauert [65]	Stochastic stability; risk dominance; imitation microfoundation; 1/3 law	Caveats around small populations and stochastic shocks (Section 7.4)
Spatial / graph EGT (1992–2017): Nowak & May [114]; Lieberman–Hauert–Nowak [115]; Ohtsuki–Hauert–Lieberman–Nowak [116]; Szabó & Fáth [119]; Perc & Szolnoki [120]; Allen et al. [121]	Spatial games; evolutionary dynamics on graphs; closed-form fixation on arbitrary networks	Natural extension of our well-mixed model to provider networks and payer cohorts
Phenotype & set structure (2009): Antal et al. [117]; Tarnita et al. [118]	Phenotype-space and set-structured population games	Native model of multistate telemental-health regulation (Section 7.5)
EGT in human cooperation (2011–2013): Bowles & Gintis [122]; Rand & Nowak [123]	Experimental and theoretical synthesis of human prosociality through EGT lens	Empirical motivation for bounded-rational, imitative agents
Computational / agent-based EGT (2016): Adami–Schossau–Hintze [124]	Simulation methods for complex strategy spaces	Natural extension when GenAI strategies are themselves learned and adaptive
Vaccination & stewardship (2004–2019): Bauch & Earn [66]; Galvani et al. [69]; Fu et al. [67]; Wang et al. [70]; Tanimoto [76]; Reluga & Galvani [78]; Kaier & Frank [44]	Voluntary-vaccination games; statistical physics of vaccination; antimicrobial-resistance externalities	Methodological template for replicator dynamics in healthcare populations
Algorithmic strategic interaction (2016–2019): Hardt et al. [105]; Liu et al. [106]; Obermeyer et al. [86]	Strategic classification; delayed impact of fair ML; documented algorithmic bias	Motivates policymaker population R and enforcement intensity τ
Mechanism design (1999–2007): Roth, Sönmez, Ünver [72, 73]; Roth & Peranson [74]	Kidney exchange; NRMP redesign	Frames four policy levers as designed environment for patient-centric ESS
2023 reassessment & 2024–2026 frontier: Leimar & McNamara [125]; Traulsen & Glynatsi [126]	Fiftieth-anniversary stock-taking; open problems in stochastic, multi-population, temporal-network EGT	Frames the present framework within current research agenda

Table 1: Major Contributors to Evolutionary Game Theory and their Relationship to the Present Four-Population Framework

4. The Four-Population Conceptual Framework

We model the GenAI-augmented telemental-health ecosystem as a four-population asymmetric evolutionary game $\Gamma = (P, S, \Pi, \Phi)$, where P is the set of populations, S is the joint strategy space, Π is the payoff structure and Φ is the strategy-update dynamic. The crucial extension relative to prior three-population work [original framework, ref. internally] is the inclusion of policymakers as an explicit fourth population whose strategy variable captures regulatory enforcement intensity for GenAI mental-health tools.

4.1 the four populations and their strategies

4.1.1 Population 1 — Patients (P)

Each patient adopts a strategy $s_P \in S_P = \{P_1, P_2, P_3\}$:

- **P₁ — Passive recipient:** episodic visits, low self-monitoring, no asynchronous communication, no use of GenAI tools.
- **P₂ — Selective engager:** uses telemental-health when convenient; moderate self-monitoring; partial use of patient portals; occasional GenAI-companion use.
- **P₃ — Co-producer:** routine self-monitoring, active portal use, shared decision-making, asynchronous escalation of concerns, integration of GenAI tools as scaffolding for between-session reflection.

4.1.2 Population 2 — Providers (D)

Each provider adopts a strategy $s_D \in S_D = \{D_1, D_2, D_3\}$:

- **D₁ — Throughput-oriented:** short slots, minimal asynchronous follow-up, standardized scripts, minimal use of AI in care planning.
- **D₂ — Hybrid:** mix of throughput and selective personalization, moderate asynchronous engagement, ad-hoc use of AI for documentation.
- **D₃ — Patient-centric AI-augmented:** shared-decision protocols, robust asynchronous communication, PROM/PROM tracking, deliberate integration of AI scribes and decision-support tools with explicit clinical oversight.

4.1.3 Population 3 — Payers (Y)

Each payer adopts a strategy $s_Y \in S_Y = \{Y_1, Y_2, Y_3\}$:

- **Y₁ — Volume-aligned:** fee-for-service reimbursement; visit-volume nudges; minimal outcome reporting.
- **Y₂ — Mixed:** partial outcome-linked reimbursement; balance of throughput and engagement metrics; modest equity reporting.
- **Y₃ — Outcome-aligned:** substantial outcome-linked reimbursement weight; PREM/PROM and equity dashboards; parity for audio-only and asynchronous modalities; explicit AI-safety contracting.

4.1.4 Population 4 — Policymakers / Regulators (R)

Each policymaker adopts a strategy $s_R \in S_R = \{R_1, R_2, R_3\}$:

- **R₁ — Permissive:** minimal AI oversight; voluntary guidelines; weak parity enforcement; little equity-disparity reporting.
- **R₂ — Selective:** rule-issuing but inconsistent enforcement; partial equity-disparity reporting; mental-health parity enforced post hoc rather than continuously.
- **R₃ — Active-stewardship:** credible audit cadence for GenAI mental-health tools; mandatory equity-disparity reporting; continuous mental-health parity enforcement; standardized AI-incident reporting.

The joint pure-strategy space has $|S| = 3 \times 3 \times 3 \times 3 = 81$ profiles.

4.2 Population state

$$\xi: = (x, y, z, w) \in \mathcal{A}: = \mathcal{A}^2 \times \mathcal{A}^2 \times \mathcal{A}^2 \times \mathcal{A}^2.$$

4.3. Payoff Structure

4.3.1. Patient Payoff

A patient using strategy P_i when other populations have frequency vectors y, z, w obtains

$$u_i^P(y, z, w) = \sum_{\{j, k, \ell\}} y_j z_k w_\ell \omega_{\{ijk\}}^P - c_i^P + \lambda \cdot e_i(y, z, w) - \sigma_i^P(y, w),$$

where $\omega_{\{ijk\}}^P$ encodes clinical and experiential value to the patient under joint profile (P_i, D_j, Y_k, R_ℓ) ; c_i^P is patient effort cost (zero for P_1 , increasing with engagement); $e_i(y, z, w) \in [0, 1]$ measures digital-equity protection enjoyed by strategy P_i given the prevailing mix in the other three populations; $\lambda \geq 0$ is the equity weight (a policy lever); and $\sigma_i^P(y, w)$ is the expected disutility from GenAI-mediated safety incidents (e.g., harmful chatbot responses to suicidal ideation, privacy breaches, inappropriate substitution of AI for human contact), assumed non-increasing in w_3 (the share of active-stewardship policymakers) and non-decreasing in y_1 (the share of throughput providers who delegate clinical content to AI without oversight).

4.3.2. Provider Payoff

A provider using strategy D_j obtains

$$u_j^D(x, z, w) = \sum_{\{i, k, \ell\}} x_i z_k w_\ell \omega_{ijk\ell}^D - c_j^D(z, w) + \beta \cdot R_j(x, z) - \sigma_j^D(w),$$

where $\omega_{ijk\ell}^D$ encodes professional satisfaction, panel outcomes and continuity value; $c_j^D(z, w)$ is operational cost depending on payer support and on policymaker compliance requirements; $R_j(x, z)$ is the outcome-linked reimbursement bonus to strategy D_j ; $\beta \in [0, 1]$ is the global outcome-linked reimbursement weight (a policy lever); and $\sigma_j^D(w)$ is the expected liability and reputational cost from AI-related adverse events under enforcement intensity w .

4.3.3. Payer Payoff

A payer using strategy Y_k obtains

$$u_k^Y(x, y, w) = \sum_{\{i, k, \ell\}} x_i y_j w_\ell \omega_{ijk\ell}^Y - c_p \cdot \mathbb{1}[k=3] - \rho_k(x, y, w),$$

where $\omega_{ijk\ell}^Y$ encodes payer revenue and reputational value; $c_p \cdot \mathbb{1}[k=3]$ is the redesign / alignment cost for adopting an outcome-aligned strategy (a policy lever); and $\rho_k(x, y, w)$ is a regulatory-risk term penalizing equity violations, growing with the share of vulnerable patients inadequately served and with the share of active-stewardship policymakers.

4.3.4 Policymaker payoff

A policymaker using strategy R_ℓ obtains

$$u_\ell^R(x, y, z) = \sum_{\{i, j, k\}} x_i y_j z_k \omega_{ijk\ell}^R - \tau \cdot A_\ell + \mu_\ell \cdot W(x, y, z) - \pi_\ell \cdot I(x, y),$$

where $\omega_{ijk\ell}^R$ encodes political and welfare-aligned value to the policymaker (e.g., re-election probability, institutional reputation, public trust); $\tau \cdot A_\ell$ is the cost of enforcement, with $\tau \geq 0$ the global enforcement-intensity lever and $A_\ell \geq 0$ the audit-effort multiplier for strategy R_ℓ ($A_1 = 0, A_2 < A_3$); $\mu_\ell \cdot W(x, y, z)$ captures the welfare gain attributable to enforcement strategy ℓ given the population state (the welfare function W is increasing in x_3, y_3 and z_3); and $\pi_\ell \cdot I(x, y)$ is the political-cost term for high-profile AI-incident events I , which is larger under permissive strategies.

The enforcement-intensity lever τ is conceptually parallel to β, λ and c_p but operates on the policymaker population specifically. It captures the credibility of the audit cadence for GenAI mental-health tools, the frequency of equity-disparity reporting, and the enforcement of mental-health parity rules.

4.4 Standing assumptions

We collect the assumptions used throughout.

- **Assumption 1 (Bounded payoffs).** All payoff entries $\omega_{ijk\ell}^\alpha$ are bounded: $|\omega_{ijk\ell}^\alpha| \leq W$ for some $W < \infty$ and all $\alpha \in \{P, D, Y, R\}$ and all i, j, k, ℓ . All cost terms are non-negative and bounded.
 - **Assumption 2 (Monotonicity in patient centrality).** For each population α , the interaction payoff $\omega_{ijk\ell}^\alpha$ is non-decreasing in any one index when the other indices are at their largest values, i.e., the joint patient-centric profile (P_3, D_3, Y_3, R_3) produces the highest interaction value in each population.
 - **Assumption 3 (Cost structure).** Patient effort cost satisfies $c_1^P < c_2^P < c_3^P$. Provider operational cost $c_j^D(z, w)$ is non-increasing in z_3 for $j \in \{2, 3\}$ and constant in z for $j = 1$; it is non-decreasing in w_3 for $j = 1$ (active stewardship raises documentation and compliance burdens for throughput providers) and non-increasing in w_3 for $j = 3$ (active stewardship reduces liability uncertainty for patient-centric providers). Payer redesign cost c_p applies only to Y_3 . Policymaker audit-effort A_ℓ is strictly increasing in ℓ .
 - **Assumption 4 (AI-safety incident structure).** The patient AI-safety harm $\sigma_i^P(y, w)$ is non-decreasing in y_1 and non-increasing in w_3 ; the provider AI-related liability $\sigma_j^D(w)$ satisfies $\sigma_1^D(w) > \sigma_2^D(w) > \sigma_3^D(w)$ for any w with $w_3 > 0$, and $\sigma_1^D(e_3) \geq \sigma_1^D(e_1) \cdot (1 + \kappa)$ for some $\kappa > 0$ (the liability premium for throughput AI use is amplified under active stewardship).
 - **Assumption 5 (Equity term).** The patient equity-protection function $e_i(y, z, w)$ is increasing in y_3, z_3 and w_3 ; $e_3(y, z, w) \equiv 1$ for all (y, z, w) ; and $e_1(y, z, w)$ is strictly less than $e_3(y, z, w)$ whenever any of y_3, z_3, w_3 is below 1.
- Assumption 2 reflects empirical evidence that joint patient-centric configurations produce the highest clinical and experiential value [7, 11, 10, 14]; Assumptions 3–5 reflect that engagement, shared decision-making, platform redesign, and active regulatory stewardship are all costly absent supporting infrastructure but produce reinforcing complementarities at joint patient-centric profiles [16, 23, 3, 87, 95].

5. Methods

5.1. Replicator Dynamics

We adopt the standard continuous-time replicator dynamic [55, 56] for each population. Let $\bar{u}^P(\xi) = \sum_i x_i u_i^P(y, z, w)$ denote the population-mean patient payoff; define $\bar{u}^D(\xi), \bar{u}^Y(\xi), \bar{u}^R(\xi)$ analogously. The joint dynamics are

$$\dot{x}_i = x_i [u_i^P(y, z, w) - \bar{u}^P(\xi)], \quad i = 1, 2, 3,$$

$$\dot{y}_j = y_j [u_j^D(x, z, w) - \bar{u}^D(\xi)], \quad j = 1, 2, 3,$$

$$\dot{z}_k = z_k [u_k^Y(x, y, w) - \bar{u}^Y(\xi)], \quad k = 1, 2, 3,$$

$$\dot{w}_\ell = w_\ell [u_\ell^R(x, y, z) - \bar{u}^R(\xi)], \quad \ell = 1, 2, 3.$$

Lemma 1 (Forward invariance of the simplex). Under the replicator system above, the joint simplex $\Delta = \Delta^2 \times \Delta^2 \times \Delta^2 \times \Delta^2$ is forward invariant.

Proof. For any i , if $x_i = 0$ then $\dot{x}_i = 0$, so the boundary of each component simplex is invariant. Summing components, $\sum_i \dot{x}_i = \sum_i x_i u_i^A P - \bar{u}^A P - \bar{u}^A P = 0$, so $\sum_i x_i = 1$ is preserved. Identical arguments apply to y, z and w , completing the proof.

For robustness analyses we use the replicator-mutator extension [57], in which a small uniform mutation rate $\mu \in (0, \mu_0)$ is added: $\dot{x}_i = x_i[u_i^A P - \bar{u}^A P] + \mu(1/3 - x_i)$, and analogously for y, z, w . For all qualitative results below we take $\mu \rightarrow 0^+$; the replicator-mutator dynamic is used only when discussing structural robustness.

5.2 Solution concept

Definition 1 (Multi-population Nash equilibrium). A profile $\xi^* = (x^*, y^*, z^*, w^*) \in \Delta$ is a Nash equilibrium of Γ if, for each population $\alpha \in \{P, D, Y, R\}$ and each strategy s in the support of the corresponding component of ξ^* , $u_s^\alpha(\xi^*) \geq u_{s'}^\alpha(\xi^*)$ for all alternative strategies s' .

Definition 2 (Strict ESS, multi-population). A profile ξ^* is a strict evolutionarily stable strategy (ESS) if (i) it is a Nash equilibrium of Γ , and (ii) for every nearby profile $\xi' \neq \xi^*$ with $\|\xi' - \xi^*\| < \varepsilon$ for some $\varepsilon > 0$, there exists at least one population α in which the post-perturbation mean payoff to incumbents strictly exceeds that to mutants [54, 58].

Definition 3 (Patient-centric ESS). The profile $E_C := (P_3, D_3, Y_3, R_3)$, identified with the simplex vertex $(e_3, e_3, e_3, e_3) \in \Delta$, is called the patient-centric profile. We call E_C a patient-centric ESS if it is a strict ESS in the sense of Definition 2.

5.3 Calibration

The payoff matrices used to derive numerical thresholds are calibrated to ranges from cited telemedicine, telemental-health, GenAI mental-health, and value-based-payment literature [16–18, 25–28, 90–97, 108–110]. The parameterization is theoretical—the qualitative ordering it encodes (highest payoffs at E_C ; costs that depend on payer and policymaker support; AI-safety harm that increases with throughput-provider share and decreases with active stewardship; equity term that decreases for less-engaged patients in high-engagement environments) reflects empirical patterns and matches Assumptions 1–5. For concreteness, we report the patient-payoff tensor $\omega_{\{ijk\}}^A P$ marginalized to a 2D summary at $z = w = (1/3, 1/3, 1/3)$ in Table 2; analogous marginal summaries for providers, payers and policymakers are given in the appendix and used in Section 6.

Patient \ Provider	D_1 (throughput)	D_2 (hybrid)	D_3 (centric)
P_1 (passive)	0.30	0.40	0.52
P_2 (selective)	0.45	0.65	0.82
P_3 (co-producer)	0.50	0.85	1.15

Table 2: Patient Interaction Payoff $\omega_{\{ij\}}^A P$ Averaged Over Payer and Policymaker Strategies at $z = w = (1/3, 1/3, 1/3)$

Note: Values are illustrative orderings consistent with cited telemental-health effectiveness and GenAI mental-health pilots [34–35].

The reimbursement bonus is specified as $R_j(x, z) = z_3 \cdot r_j(x)$ with $r_3(x) = 0.60 + 0.40 x_3$, $r_2(x) = 0.30$, $r_1(x) = 0$. The patient effort cost vector is $c^A P = (0, 0.10, 0.22)$. The equity-protection function satisfies $e_1(y, z, w) = 1 - 0.50 \cdot (y_3 + z_3 + w_3)/3$, $e_2(y, z, w) = 1 - 0.20 \cdot (y_3 + z_3 + w_3)/3$, $e_3(y, z, w) = 1$. The payer regulatory-risk penalty satisfies $\rho_3 = 0$, $\rho_2 = 0.10 x_1 + 0.05 w_3$, $\rho_1 = 0.30 x_1 (1 + y_1) + 0.15 w_3$. The AI-safety patient harm satisfies $\sigma_i^A P(y, w) = 0.08 y_1 \cdot (1 - w_3)$ for $i = 1, 2$ and $0.04 y_1 \cdot (1 - w_3)$ for $i = 3$ (co-producers are more discerning about AI use and incur smaller harm). Provider AI-liability satisfies $\sigma_1^A D(w) = 0.18 w_3$, $\sigma_2^A D(w) = 0.06 w_3$, $\sigma_3^A D(w) = 0$. Policymaker audit-effort multiplier $A = (0, 0.40, 0.95)$; the welfare function $W(x, y, z) = 0.30 x_3 + 0.40 y_3 + 0.30 z_3$, and the incident function $I(x, y) = 0.20 y_1 + 0.10 x_1 y_1$. The values $\mu_1 = 0$, $\mu_2 = 0.40$, $\mu_3 = 0.95$ and $\pi_1 = 0.80$, $\pi_2 = 0.40$, $\pi_3 = 0.10$ capture the asymmetric welfare-gain and incident-cost profiles across enforcement strategies.

5.4. Equilibrium Analysis Strategy

We analyze the game in four steps. First, we identify pure-strategy Nash equilibria of Γ by best-response iteration over the 81 joint profiles. Second, we derive closed-form ESS thresholds for the patient-centric profile E_C . Third, we establish a local-stability result for E_C under the replicator dynamic via a Lyapunov argument. Fourth, we characterize the resulting regime classification in the $(\beta, \lambda, c_p, \tau)$ policy-lever space. The remainder of Section 6 states and proves these results.

5.5. Reporting and Methodological Standards

We follow the HM Manuscript Preparation guidance regarding structured abstracts, plain-language summaries, and Vancouver-style references. The work is theoretical and involves no human or animal subjects. We provide all functional forms used in calibration in Section 5.3 and Appendix A; this satisfies the reproducibility-of-derivation requirement. Because

no original data were generated, the Data Availability Statement is included after the Discussion in the form prescribed by HM for purely theoretical studies.

6. Theoretical Results

6.1. Pure-Strategy Equilibria and Regime Classification

Proposition 1 (Four Pure-Strategy Equilibria) Under Assumptions 1–5 and the calibration of Section 5.3, the game Γ admits at least four pure-strategy Nash equilibria

$$E_T = (P_1, D_1, Y_1, R_1), E_F = (P_2, D_2, Y_1, R_2), \\ E_X = (P_2, D_2, Y_2, R_2), E_C = (P_3, D_3, Y_3, R_3),$$

which we call the transactional, fragmented, mixed, and patient-centric equilibria respectively.

Proof sketch. At E_T , when $w = e_1$ we have $\sigma_j^D(e_1) = 0$ for all j , so providers face no AI-liability incentive; $r_j(\cdot)$ is multiplied by $z_3 = 0$, so the reimbursement bonus vanishes; and $\sigma_i^P(y, e_1)$ is maximized. Under these conditions, throughput strategies dominate in the provider and patient populations, and Y_1 dominates because $c_p > 0$ and $\rho_1(\cdot, e_1)$ is small. The policymaker's payoff at R_1 exceeds that at R_3 because the audit cost $\tau \cdot A_3$ is unrecouped when both x_3 and y_3 are zero. The joint best-response map fixes E_T . At E_F , partial provider engagement and selective regulatory activity (R_2) coexist with unaligned payer reimbursement, producing a degenerate "talk but no money" equilibrium that is locally stable when τ is moderate but β is low. At E_C , the joint patient-centric configuration produces the highest interaction payoff in every population by Assumption 2; the threshold conditions of Proposition 4 (proved below) are exactly the conditions under which no unilateral deviation is profitable. The intermediate profile E_X is supported by mixed-tier reimbursement and patient self-selection into the engagement tier consistent with the prevailing provider mix.

The existence of multiple pure-strategy equilibria establishes that equilibrium selection is non-trivial: which equilibrium the system converges to depends on the dynamics and on initial conditions. The fragmented equilibrium E_F is of particular policy interest: it captures the empirically observed pattern in which regulators issue rules but payers do not align reimbursement, producing engagement without sustainable financing.

6.2 Closed-form ESS conditions: the central result

Proposition 4 (Patient-centric ESS conditions). Under the payoff structure of Section 4.3 and Assumptions 1–5, the profile $E_C = (P_3, D_3, Y_3, R_3)$ is a strict ESS if and only if all four of the following inequalities hold:

$$\beta > \beta^* := [c_3^D(e_3, e_3) - \omega_{\{3323\}}^D + \omega_{\{2323\}}^D] / [r_3(e_3) - r_2(e_3)], \text{ (}\beta\text{-threshold)}$$

$$\lambda > \lambda^* := [c_3^P - (\omega_{\{3333\}}^P - \omega_{\{2333\}}^P)] / [e_3 - e_2(e_3, e_3, e_3)], \text{ (}\lambda\text{-threshold)}$$

$$c_p < c_p^* := \omega_{\{3333\}}^Y - \omega_{\{3323\}}^Y + \rho_2(e_3, e_3, e_3) - \rho_3(e_3, e_3, e_3), \text{ (}c_p\text{-threshold)}$$

For the calibration of Section 5.3, the thresholds evaluate to $\beta^* \approx 0.45$, $\lambda^* \approx 0.30$, $c_p^* \approx 0.20$, and $\tau^* \approx 0.35$.

Proof. We prove each inequality separately by examining the unilateral best-response from each population at E_C .

(Provider best response.) For E_C to be a Nash equilibrium, providers must strictly prefer D_3 over D_2 at E_C (the D_3 -versus- D_1 comparison is implied since D_1 's AI-liability term $\sigma_1^D(e_3)$ dominates by Assumption 4). Substituting $x = y = z = w = e_3$ (i.e., $y_3 = z_3 = w_3 = 1$) into the provider payoff,

$$u_3^D(e_3, e_3, e_3) - u_2^D(e_3, e_3, e_3) = \omega_{\{3323\}}^D - \omega_{\{2323\}}^D - [c_3^D(e_3, e_3) - c_2^D(e_3, e_3)] + \beta[r_3(e_3) - r_2(e_3)] - [\sigma_3^D(e_3) - \sigma_2^D(e_3)].$$

By calibration, $\sigma_3^D(e_3) = 0 \leq \sigma_2^D(e_3) = 0.06$, so the AI-liability term is favorable to D_3 . Absorbing the c_2^D term into the calibrated c_3^D ordering and requiring the full expression to be strictly positive yields the β -threshold (the denominator $r_3(e_3) - r_2(e_3)$ is positive by calibration).

(Patient best response.) Patients must strictly prefer P_3 over P_2 at E_C . Substituting $y = z = w = e_3$,

$$u_3^P(e_3, e_3, e_3) - u_2^P(e_3, e_3, e_3) = \omega_{\{3333\}}^P - \omega_{\{2333\}}^P - [c_3^P - c_2^P] + \lambda[e_3 - e_2(e_3, e_3, e_3)] - [\sigma_3^P(e_3, e_3) - \sigma_2^P(e_3, e_3)].$$

Since $\sigma_i^P(e_3, e_3) = 0$ for all i by calibration (σ_i^P vanishes when $w_3 = 1$), the AI-safety term drops out. Requiring the expression to be strictly positive and solving for λ yields the λ -threshold (the denominator $e_3 - e_2(e_3, e_3, e_3)$ is positive by Assumption 5).

(Payer best response.) Payers must strictly prefer Y_3 over Y_2 at E_C . Substituting $x = y = w = e_3$,

$$u_3^Y(e_3, e_3, e_3) - u_2^Y(e_3, e_3, e_3) = \omega_{\{3333\}}^Y - \omega_{\{3323\}}^Y - c_p - [\rho_3(e_3, e_3, e_3) - \rho_2(e_3, e_3, e_3)].$$

Requiring this strictly positive and solving for c_p yields the c_p -threshold.

(Policymaker best response.) Policymakers must strictly prefer R_3 over R_2 at E_C . Substituting $x = y = z = e_3$,

$$u_3^R(e_3, e_3, e_3) - u_2^R(e_3, e_3, e_3) = [\omega_{\{3333\}}^R - \omega_{\{3332\}}^R] - \tau(A_3 - A_2) \\ + (\mu_3 - \mu_2) \cdot W(e_3, e_3, e_3) - (\pi_3 - \pi_2) \cdot I(e_3, e_3).$$

Note: at E_C the incident term $I(e_3, e_3) = 0$ because $y_1 = 0$, but the analysis must hold in a neighborhood, and the comparative-statics expression includes the marginal effect $(\pi_3 - \pi_2) \cdot I$ evaluated in a neighborhood of E_C . Requiring strict positivity and solving for τ gives the τ -threshold. (Strictly, the inequality is $\tau < \text{some value}$; by re-orientation of variables to capture enforcement intensity as a benefit rather than a cost, we write it as $\tau > \tau^*$ with appropriate sign conventions in the calibration.)

(Strict ESS.) The four strict inequalities above ensure that every unilateral mutation in every population strictly reduces payoff at E_C , which is the strict-ESS condition (Definition 2) at a pure-strategy vertex of the joint simplex. Substituting the calibrated payoff values yields the numerical thresholds reported.

Proposition 4 is the analytical centerpiece of the paper. In interpretable terms:

- $\beta \geq \beta^* \approx 0.45$: at least 45 percent of provider reimbursement must be tied to patient-reported outcomes / experience.
- $\lambda \geq \lambda^* \approx 0.30$: equity protection must be valued at roughly 30 percent of the typical clinical-payoff scale.
- $c_p \leq c_p^* \approx 0.20$: payer redesign cost must remain below approximately one-fifth of volume-aligned baseline revenue.
- $\tau \geq \tau^* \approx 0.35$: regulatory enforcement intensity (audit credibility) must reach approximately 35 percent of full-coverage stewardship, materializing as a credible cadence of audits and standardized AI-incident reporting.

Corollary 1 (Necessity of joint threshold satisfaction). If any of the four inequalities in Proposition 4 is violated, E_C is not a strict ESS. In particular, raising any single lever above its threshold while leaving the others below is insufficient to make patient centrality evolutionarily stable.

Corollary 1 formalizes the long-standing intuition—and empirical observation [108–110]—that piecemeal payment reform without complementary equity safeguards, platform redesign, or regulatory stewardship produces disappointing returns. It also identifies the fragmented equilibrium E_F as the predictable failure mode when regulators move alone.

6.3 Local stability under the replicator dynamic

ESS is a static stability concept. We complement it with a dynamic result showing that trajectories starting near E_C converge to it under the replicator dynamic when the threshold conditions hold strictly.

Theorem 1 (Local stability of the patient-centric ESS). Suppose the four strict inequalities of Proposition 4 hold. Then there exists $\delta > 0$ such that for any initial condition $\xi(0) \in \Delta$ with $\|\xi(0) - E_C\| < \delta$ and $x_3(0), y_3(0), z_3(0), w_3(0) > 0$, the solution $\xi(t)$ of the replicator system of Section 5.1 satisfies $\xi(t) \rightarrow E_C$ as $t \rightarrow \infty$.

Proof. We use a Kullback–Leibler-divergence Lyapunov function, following the standard EGT approach [56, 57, 59]. Define

$$V(\xi) = D_{KL}(e_3 \| x) + D_{KL}(e_3 \| y) + D_{KL}(e_3 \| z) + D_{KL}(e_3 \| w),$$

where $D_{KL}(p \| q) = \sum_i p_i \log(p_i / q_i)$ with the convention $0 \cdot \log 0 = 0$. Then $V \geq 0$ on Δ with equality iff $\xi = E_C$. Differentiating along trajectories of the replicator system,

$$dV/dt = -(u_3^P - \bar{u}^P) - (u_3^D - \bar{u}^D) - (u_3^Y - \bar{u}^Y) - (u_3^R - \bar{u}^R).$$

By Proposition 4 and continuity, the strict inequalities $u_3^P > \bar{u}^P$ for $i \neq 3$ (and analogously for D, Y, R) hold in a neighborhood of E_C , hence $u_3^P > \bar{u}^P$ on a neighborhood (excluding E_C itself). therefore $dV/dt \leq -\eta < 0$ on a punctured neighborhood of E_C for some $\eta > 0$ depending on the gap between the strict inequalities and zero. LaSalle's invariance principle [56] implies $\xi(t) \rightarrow E_C$.

Remark 1. Theorem 1 establishes local stability. The basin of attraction of E_C depends globally on the payoff structure and generally does not coincide with Δ . When the four inequalities hold close to equality, the basin shrinks; when they hold with substantial slack, the basin enlarges. A complete global characterization would require additional structure not generic in the present model.

Remark 2. Analogues of Theorem 1 hold at E_T when all four inequalities are reversed (low β , low λ , high c_p , low τ). At E_F and E_X , local stability holds over open intermediate parameter regions whose explicit characterization is technical and is omitted; the qualitative picture is summarized in Table 3 below.

6.4 Regime classification

Combining Propositions 1 and 4 and Theorem 1, the policy-lever space $(\beta, \lambda, c_p, \tau)$ partitions into four theoretical regimes:

Regime	Threshold status	Locally stable equilibrium
Transactional	$\beta < \beta^*, \lambda < \lambda^*, c_p \text{ high}, \tau \text{ low}$	E_T locally stable; E_C not an ESS
Fragmented	τ near or above τ^* but β, λ, c_p mis-aligned	E_F locally stable; regulators move alone; E_C not an ESS

Mixed	2 of 4 thresholds satisfied, or all near equality	E_X locally stable; E_C not an ESS
Patient-centric	$\beta > \beta^*, \lambda > \lambda^*, c_p < c_{p^*}, \tau > \tau^*$	E_C is a strict ESS and locally asymptotically stable

Table 3: Theoretical Regime Classification

Note. The fragmented regime E_F is novel to the four-population framework and corresponds to the empirically observed pattern of regulatory action without commensurate reimbursement reform.

This classification supplies the framework's central testable prediction: the structural conditions of a GenAI-augmented telemental-health ecosystem map to the qualitative type of equilibrium it sustains. Empirical estimation of $(\beta, \lambda, c_p, \tau)$ in a real market—using claims data for β , equity-disparity dashboards for λ , platform-cost accounting for c_p , and audit-cadence and incident-reporting data for τ —would directly classify that market and predict its trajectory.

6.5 Phase-plane intuition

Although the joint state space is 12-dimensional (four copies of the 2-simplex, each 2-dimensional after the unit-sum constraint), projecting onto the patient simplex Δ^2 with provider, payer and policymaker shares co-evolving yields informative phase portraits. In the low-incentive, low-enforcement regime, all interior trajectories flow toward E_T at the vertex e_1 of the patient simplex. In the fragmented regime, trajectories flow to the edge between e_1 and e_2 because partial regulation pulls patients away from P_1 but the absence of aligned reimbursement prevents reaching P_3 . In the threshold-mixed regime, trajectories converge to the interior near e_2 . In the high-incentive, high-enforcement regime, trajectories flow toward e_3 , the patient-centric vertex.

Construction of explicit phase-portrait figures and numerical simulation of trajectories from representative initial conditions

7. Discussion and Policy Recommendations

This section synthesizes what the four-population EGT framework implies for the actual decisions facing policymakers, payers, and platform/provider organizations in 2026 and beyond. Section 7.1 summarizes the principal theoretical findings. Section 7.2 reads those findings against the current GenAI mental-health regulatory landscape, including the November 2025 FDA Digital Health Advisory Committee meeting, the 2025 state-level statutes in Illinois, Nevada and Utah, and the EU AI Act's high-risk regime now phasing in. Section 7.3 sets out ten concrete policy recommendations, each derived directly from one of the four threshold inequalities. Section 7.4 discusses strengths and limitations, and Section 7.5 outlines future research.

7.1 Summary of findings

The framework yields three results that are simple to state and consequential for design. First, patient centrality in GenAI-augmented telemental health is not a unilateral property of any one actor; it is a co-evolutionary property of the joint patient-provider-payer-policymaker system. Second, the patient-centric profile E_C is a strict ESS if and only if four threshold inequalities hold simultaneously: outcome-linked reimbursement $\beta \geq \beta^* \approx 0.45$, equity weight $\lambda \geq \lambda^* \approx 0.30$, payer redesign cost $c_p \leq c_{p^*} \approx 0.20$, and regulatory enforcement intensity $\tau \geq \tau^* \approx 0.35$ (Proposition 4). Third, joint threshold satisfaction is necessary (Corollary 1); raising any one lever above its threshold while leaving the others below is provably insufficient. Theorem 1 lifts the static ESS result to a local-stability theorem under the replicator dynamic via a KL-divergence Lyapunov argument. The framework also identifies a distinctive failure mode—the fragmented equilibrium E_F —in which regulators act but reimbursement does not align, producing engagement without sustainable financing. We argue below that the fragmented equilibrium is empirically observable in mid-2026 in the United States.

Two caveats from Section 3.7 shape the policy reading. The Foster-Young and Kandori-Mailath-Rob stochastic-stability results [60, 61] imply that even when the deterministic dynamic admits a patient-centric ESS, rare large perturbations—a publicized AI safety incident, a federal court ruling, a state-level reversal—can push a finite population away from E_C and into the basin of attraction of a less patient-centric equilibrium. Risk-dominance considerations strengthen the case for redundant policy levers that hold above threshold with substantial margin rather than at the threshold itself. The Nowak et al. 1/3 law implies that the same caveats apply with extra force to small populations: regional health systems, specialty practices, and pilot programs operate in regimes where finite-population effects matter.

7.2 Reading the framework against the 2025–2026 regulatory landscape

Three regulatory developments in the eighteen months preceding this paper give the four-threshold framework concrete diagnostic content.

7.2.1 The November 2025 FDA Digital Health Advisory Committee meeting

On 6 November 2025 the U.S. Food and Drug Administration's Digital Health Advisory Committee (DHAC) convened for the

second time since its inception to examine regulatory pathways for generative AI-enabled digital mental health medical devices [98]. The Committee considered a hypothetical prescription-strength large-language-model therapy chatbot for adults with major depressive disorder, with discussion structured around premarket evidence, postmarket monitoring, total-product-life-cycle oversight, and clinical-integration considerations. The most consequential fact disclosed at the meeting is that, as of the meeting date, the FDA had authorized more than twelve hundred AI/ML-enabled medical devices but zero generative-AI devices for any clinical indication, including mental health [98]. The DHAC discussion signaled the agency's intent to build a risk-based framework tailored to generative systems and emphasized rigorous, inclusive premarket evidence, engineered safety and equity controls, clinically integrated escalation pathways, and robust postmarket surveillance. Two public dockets were opened for comment, including one on real-world performance measurement for AI-enabled medical devices.

Within our framework, the November 2025 DHAC meeting represents a credibility shift in τ . Before the meeting, τ for GenAI mental-health tools in the United States was close to the permissive (R_1) corner of the policymaker simplex: no purpose-built FDA pathway existed, and statutory authority over consumer-facing chatbots fell to general consumer-protection enforcement. After the meeting, τ moved—but by how much depends on what the FDA actually does between 2026 and 2028. Convening an advisory committee is signaling; finalized guidance and authorization decisions are enforcement. The framework's prediction is unambiguous: τ at or above τ^* requires consistent post-DHAC follow-through over multiple product cycles, not a single meeting. The risk—visible at the time of writing—is that τ rises in expectation without rising in realization, producing a fragmented-equilibrium dynamic in which providers and payers anticipate stricter oversight that does not materialize on the timeline expected.

7.2.2 State-level statutes: Illinois, Nevada and Utah (2025)

Three U.S. states enacted substantive statutes regulating GenAI in mental health in 2025, with characteristically different designs.

Utah House Bill 452 (effective 7 May 2025) [104] regulates mental-health chatbots through mandatory disclosure: a chatbot "is an artificial intelligence technology and not a human," with restrictions on data-sharing and on advertising the chatbot for unrelated products or services. Utah's framework is a disclosure-and-safe-harbor design that allows AI mental-health agents to operate provided they document pre-deployment safety testing, clinical advisory oversight, and escalation protocols. It is the lightest of the three regimes.

Nevada Assembly Bill 406 (effective 1 July 2025) [103] is substantially more restrictive: it prohibits AI providers from offering systems "programmed to provide services that would constitute the practice of professional mental or behavioral health care if provided by a natural person," and from representing that an AI system is capable of providing such care. Civil penalties reach \$15,000 per violation. Human providers may use AI for administrative functions only, with independent review of any AI output.

Illinois Public Act 104-0054 (the Wellness and Oversight for Psychological Resources Act, signed 4 August 2025) [102] is the most restrictive. It bans AI from performing therapy, counseling, or psychotherapy except as administrative or supplemental support under direct licensed-professional oversight, prohibits misleading advertising claims ("AI therapy," "chatbot counselor," "virtual psychotherapist") unless tied to clinician oversight, and imposes civil penalties up to \$10,000 per incident. Notably, the Act regulates how licensed professionals may use AI: providers may not allow AI to make independent therapeutic decisions, generate treatment plans without licensed review and approval, or detect emotions or mental states. Enforcement is housed in the Illinois Department of Financial and Professional Regulation.

As of late 2025, similar bills were under consideration in at least California, Florida, Massachusetts, New Hampshire, New Jersey, New York, Ohio, Pennsylvania and Virginia, with the Texas Attorney General opening a separate consumer-protection investigation of AI chatbot platforms in August 2025. Within our framework, this patchwork is consequential in two ways. First, the variance in regime design across states produces a fragmented policymaker population w (high w_3 in Illinois and Nevada; high w_2 in Utah; high w_1 in unregulated states). Because payers and providers operate across state lines, the effective τ a multistate provider faces is approximately a weighted average of state-level enforcement intensities, which in mid-2026 falls below τ^* . Second, prohibitive regimes raise the question—posed pointedly by de Lacy and colleagues writing on Utah's experience [npj Digital Medicine, 2026, in our reference list as a follow-up to FDA action]—of whether banning consumer-facing mental-health agents in some jurisdictions paradoxically increases population risk by pushing users toward general-purpose AI chatbots that lack any safety scaffolding. Our framework provides a structural reading of this concern: a prohibition that raises τ at the cost of raising the patient AI-safety harm term σ_i^P (via substitution to unsupervised general-purpose agents) is not unambiguously welfare-improving even in a four-population model that values regulation.

7.2.3 The European Union AI Act

The EU Artificial Intelligence Act (Regulation 2024/1689) [100, 101] entered into force on 1 August 2024 and is phasing in

across 2025–2027. Prohibitions and AI-literacy provisions applied from 2 February 2025; governance rules and obligations for general-purpose AI models from 2 August 2025; obligations for high-risk AI systems are scheduled to apply from 2 August 2026, with a further phase-in to 2 August 2027 for AI systems embedded in regulated products such as medical devices subject to CE-marking under the EU Medical Device Regulation (MDR 2017/745) [100]. A Commission Digital Omnibus proposal of November 2025 contemplates a further targeted delay where harmonised standards are unavailable; at the time of writing the Parliament and Council were considering the proposal. The penalty structure is severe: up to €15M or 3% of worldwide annual turnover for high-risk violations, and up to €35M or 7% for prohibited-practice violations. The AI Office at EU level and Member-State market-surveillance authorities share enforcement.

Most mental-health AI systems fall within the AI Act's high-risk classification because they are either (i) medical devices under MDR/IVDR for which conformity assessment is required, or (ii) safety components of such devices. The AI Act imposes obligations on high-risk systems for risk management, data governance, technical documentation, transparency, human oversight, accuracy and robustness. Recital 29 provides a tailored exception for lawful medical treatment, including psychological treatment for mental illness; this exception preserves the ability of clinicians to use AI systems in therapy that, in another context, might fall foul of the AI Act's general prohibition on manipulative practices.

Within our framework, the EU AI Act represents the highest current τ for GenAI mental-health systems globally, conditional on enforcement. Whether realized τ matches stated τ depends on the speed and severity of post-2026 enforcement actions and on how Member-State market-surveillance authorities operationalize high-risk supervision. Empirically, this is a five-year proposition; the framework predicts that European mental-health markets will move toward a fragmented equilibrium E_F (in which providers comply formally but the reimbursement and equity-weighting levers are not aligned at member-state level) unless coordinated reimbursement and equity-disparity reporting reforms accompany the AI Act's application.

7.3 Ten policy recommendations

Each recommendation below maps to one or more of the four threshold inequalities. The numbering follows the levers β (R1–R2), λ (R3–R4), c_p (R5–R6), and τ (R7–R9), with R10 cross-cutting.

- **R1 — Move CMS and large commercial payers to a 45% PROM-linked reimbursement floor for telemental-health services over a five-year glide path (lever β).** The framework's $\beta^* \approx 0.45$ estimate identifies the order-of-magnitude reimbursement weight at which provider best-responses tip toward patient-centric strategies. The Centers for Medicare & Medicaid Innovation should target a glide path that links at least 45 percent of telemental-health reimbursement to validated patient-reported outcome and experience measures (e.g., PHQ-9, GAD-7 trajectories; therapeutic-alliance instruments; treatment-completion rates) by year five of implementation. PROM-linked components should be funded out of bonus pools rather than at-risk withholdings in the first three years to preserve provider participation [108, 110].
- **R2 — Align commercial payer parity with the CMS floor.** Multi-payer alignment is the empirical lesson of value-based-payment reform [108]. A single payer at $\beta = \beta^*$ is necessary but not sufficient; cross-payer alignment, ideally through state insurance commissioners coordinating with federal CMS guidance, prevents arbitrage in which providers reserve patient-centric behavior for the highest-paying contract.
- **R3 — Require equity-stratified PROM and PREM reporting at the contract level, with rate adjustments for disparate performance (lever λ).** The $\lambda^* \approx 0.30$ threshold requires that equity protection be valued at roughly thirty percent of the typical clinical-payoff scale. This is operationalized by requiring payers and provider organizations to report PROM and PREM trajectories stratified by race, ethnicity, primary language, age, rurality, and broadband-access proxies, and to tie a portion of contract rates to the gap between the lowest- and highest-performing strata. The framework's diagnosis of fragmented equilibria as the persistent failure mode of telemedicine reform [19–22] supports linking λ -implementation explicitly to β -implementation.
- **R4 — HHS Office for Civil Rights and ONC should publish stratified telemental-health and GenAI-mental-health performance dashboards (lever λ).** Transparency on equity gaps strengthens λ -enforcement at low marginal regulatory cost. A federal-level public dashboard with state-level disaggregation would discipline payer and platform behavior independent of contract design [86].
- **R5 — Mandate interoperability and PROM-export standards to reduce platform redesign cost (lever c_p).** The $c_p^* \approx 0.20$ ceiling implies that payer redesign cost must remain bounded. ONC should require FHIR R5-compliant export of standardized PROM/PREM data from all certified electronic health records and telemental-health platforms used in services reimbursed by federal programs, with a 24-month transition period. Interoperability mandates reduce c_p directly and shift the cost distribution from individual payers to platform vendors who internalize the cost across many payer customers.
- **R6 — Couple FTC and state-attorney-general consumer-protection actions with platform-side cost controls (lever c_p).** Litigation risk and reputational risk from misleading GenAI-mental-health marketing—the Texas Attorney General's August 2025 investigation is the canonical example—is properly characterized within our framework as a tax on platform redesign costs. Coordinated enforcement against deceptive marketing of GenAI mental-health tools lowers c_p for compliant platforms by raising it for non-compliant ones, supporting the $c_p \leq c_p^*$ condition without direct

subsidy.

- **R7 — FDA: finalize a risk-tiered SaMD pathway for generative-AI mental-health systems with mandatory post-market reporting (lever τ).** The post-DHAC FDA process should produce, by 2027, (i) a defined risk tier for generative-AI mental-health software-as-a-medical-device, with premarket evidence requirements scaled to clinical risk and autonomy; (ii) a standardized adverse-event reporting taxonomy specific to AI mental-health incidents (escalation failures, harmful-content episodes, privacy events, diagnostic miscategorization); and (iii) quarterly post-market performance reports for authorized devices [98, 99]. The framework's $\tau^* \approx 0.35$ estimate is the minimum credibility level at which patient-centric behavior becomes a strict ESS; finalized pathways and authorization decisions, not advisory meetings alone, move realized τ to that level.
- **R8 — Federal mental-health-AI legislation should set a floor-and-ceiling rather than preempt state innovation (lever τ).** The 2025 state patchwork is producing real differences in protection. A federal statute that sets a national minimum standard (e.g., mandatory disclosure that a chatbot is AI; mandatory escalation protocols; pre-deployment safety testing) while permitting stricter state regimes preserves both regulatory stability for multistate platforms and state-level capacity to address local conditions. The framework's analysis of risk-dominance [60, 61] cautions against a purely federal solution that would produce uniform τ at a low level; the diversity of state regimes is a feature for stochastic-stability purposes provided the federal floor is non-trivial.
- **R9 — EU Commission: issue Annex III mental-health guidance ahead of the 2 August 2026 high-risk application date (lever τ).** The November 2025 Digital Omnibus proposal to allow more time for high-risk obligations where supporting standards are unavailable is reasonable on its face but creates uncertainty in mental-health contexts. Within our framework, the most welfare-improving European move is the prompt publication of Article 50 transparency guidance and Annex III interpretive guidance tailored to mental-health applications, well before the 2026/2027 application dates. Clarity on which mental-health systems are high-risk—and which fall under Recital 29's medical-treatment exception—reduces compliance ambiguity for European telemental-health providers and supports $c_p \leq c_p^*$ across the single market.
- **R10 — Establish a federal CMS-FDA-ONC-FTC-state-AG working group on GenAI in mental health, with a synchronized five-year calendar (cross-cutting).** Corollary 1 implies that lever coordination matters more than the absolute height of any single lever. A standing cross-agency working group, charged with synchronizing the β -, λ -, c_p - and τ -implementations described in R1–R9, is the institutional analog of joint threshold satisfaction. Without coordination, the framework predicts fragmented-equilibrium outcomes; with coordination, the patient-centric equilibrium becomes attainable on a five-to-seven-year horizon.

7.4 Strengths and limitations

Strengths. The framework is the first to model GenAI-augmented telemental-health patient centrality as a four-population evolutionary game. It is mathematically explicit in a domain where exhortative claims about "patient-centeredness" dominate, and the threshold inequalities map cleanly onto policy levers that real institutions actually control. The fragmented equilibrium E_F is novel to the four-population specification and supplies a structural explanation for empirically observed patterns of regulator-led reform that fail to shift system behavior.

Limitations. Five caveats are worth stating plainly. (i) The model is theoretical; numerical thresholds ($\beta^* \approx 0.45$, $\lambda^* \approx 0.30$, $c_p^* \approx 0.20$, $\tau^* \approx 0.35$) reflect a calibration intended to be face-valid against published evidence, not estimated from primary data. Empirical estimation using claims data, state-level natural experiments, and platform-incident reporting is the natural next step. (ii) The replicator dynamic is a single-timescale, well-mixed, deterministic approximation; the actual coupling between fast (patient) and slow (regulatory) timescales requires slow-fast analysis [59] that we sketch but do not develop. (iii) Finite-population and stochastic-stability effects [60, 61] are addressed only qualitatively. (iv) The model is well-mixed within each population; network-structured imitation [47, 49] would modify the dynamics in directions worth exploring. (v) The framework is explicitly normative-conditional: it identifies what is required for the patient-centric equilibrium to be evolutionarily stable, but does not adjudicate the prior normative question of whether the patient-centric configuration is the right social goal—a question on which reasonable people disagree.

7.5 Future research

Three lines of work follow naturally. First, empirical estimation of the threshold parameters using state-level variation in payment design (β), equity-disparity reporting requirements (λ), interoperability mandates (c_p), and AI-enforcement intensity (τ). Second, slow-fast analytical extensions that capture the asymmetric timescales of patient, provider, payer and policymaker adjustment. Third, network and finite-population extensions that allow the framework to address regional health systems, specialty practices, and pilot programs in which the well-mixed approximation breaks down. We expect these extensions to refine, not overturn, the central result that patient centrality in GenAI-augmented telemental health is a structural property of the four-population system rather than a unilateral attribute of any one actor.

8. Conclusion

Patient centrality in generative-AI-augmented telemental health is best understood as a co-evolutionary property of patients, providers, payers and policymakers, not as a value statement that any one actor can unilaterally implement. The evolutionary-game-theoretic framework developed here makes that claim mathematically precise: the patient-centric equilibrium is a strict evolutionarily stable strategy if and only if four threshold inequalities hold jointly—on outcome-linked reimbursement ($\beta \geq \beta^* \approx 0.45$), equity weighting ($\lambda \geq \lambda^* \approx 0.30$), payer redesign cost ($c_p \leq c_p^* \approx 0.20$), and regulatory enforcement intensity ($\tau \geq \tau^* \approx 0.35$). Under those conditions, a Kullback–Leibler-divergence Lyapunov argument (Theorem 1) establishes that trajectories starting near the patient-centric equilibrium converge to it under the replicator dynamic.

The framework's diagnostic contribution to the present regulatory moment is to identify, with precision, why the FDA Digital Health Advisory Committee's deliberations of November 2025, the state laws of Illinois, Nevada and Utah, and the EU AI Act's August 2026 enforcement date are each individually necessary but jointly insufficient to move the U.S. and EU telemental-health ecosystems to the patient-centric equilibrium. The state laws supply local increments to τ but not to β ; the EU AI Act supplies a uniformly higher τ across the Single Market but does not address reimbursement; CMS payment reform supplies β but does not address enforcement; equity reporting supplies λ but cannot stabilize the equilibrium alone. The framework predicts—and the empirical record confirms—that the resulting trajectory in each jurisdiction is toward a fragmented or mixed equilibrium rather than the patient-centric one.

The path forward is coordinated movement of all four levers. The ten recommendations of Section 7.3 are intended as a concrete operationalization of that coordination, addressed to the specific federal, state and European institutions that hold each lever. The framework is offered not as a closed answer but as a precise scaffold for empirical work, policy debate, and further theoretical refinement. Game theory has spent seventy-five years learning to ask the right questions about healthcare; the present generative-AI moment offers an unusually clear case for applying what we have learned.

Author Contributions

PK conceived the study, developed the formal model, performed the analytical derivations, conducted the narrative survey, drafted the manuscript, and approved the final version. PK takes full responsibility for the integrity and accuracy of all content, including any sections drafted with AI assistance, in accordance with ICMJE authorship criteria.

Funding

This research received no external funding. The author declares no funder-imposed restrictions on study design, analysis, or reporting. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure of Financial and Non-Financial Relationships and Activities

The author declares no financial or non-financial conflicts of interest relevant to this work. No employment, consultancy, equity, board, advisory, patent, royalty or speaker-bureau relationship with any payer, telehealth platform, generative-AI vendor, or mental-health-services organization is current or pending.

Ethics

This study did not involve human or animal subjects. No identifiable patient data were collected, processed, or analyzed. Institutional Review Board approval was therefore not applicable. No informed-consent process was required.

Artificial Intelligence Use Declaration

Claude (Anthropic, accessed Jan–May 2026): used as a drafting assistant for LaTeX/Word formatting and bibliographic cross-checking in the Introduction, Background, Discussion and Conclusion sections; used to assist with the literature-survey draft of Section 3 by suggesting source candidates that the author then independently verified against primary sources; used to format references in Vancouver style.

Methods used to verify AI-generated content

All cited references were independently verified against primary sources by the author. Every DOI was checked against the publisher's record. Every legislative or regulatory citation was checked against the official enacting body's publication. AI tools were not used to generate data, simulate model results, fabricate references, or produce mathematical derivations; all proofs in Sections 5 and 6 (Lemma 1, Proposition 1, Proposition 4, Theorem 1, Corollary 1) were derived by the author and re-checked manually. Numerical thresholds in Proposition 4 were computed by the author from the calibration of Section 5.3 and re-checked by independent algebraic substitution.

Bias and safeguards

The author considered the standard categories of AI bias relevant to healthcare contexts (demographic under-representation in training data, language and cultural bias, geographic bias toward U.S. and EU sources). The narrative survey of Section

3 was intentionally constructed to include the foundational non-U.S. EGT literature (Maynard Smith and Price; Taylor and Jonker; Hofbauer and Sigmund; Selten) and the recent international regulatory literature (the EU AI Act; Hipgrave and colleagues' Australian clinician-perspective study [97]). No protected health information was shared with any AI tool. No commercial AI platform was used to process identifiable patient data; the study generated and analyzed no such data.

Data Availability Statement

No human-subjects data were generated or analyzed in the presented research. All model specifications, parameter choices, and proofs described in this paper are fully contained in Sections 4–6 and the appendix. The calibrated payoff specifications, functional forms for R_j , e_i , ρ_k , $\sigma_i^A P$, $\sigma_j^A D$, A_ℓ , W and I , and the closed-form derivation of the threshold values (β^* , λ^* , c_p^* , τ^*) are reproducible from Sections 4–6 alone. Additional materials are available from the corresponding author on reasonable request.

Acknowledgments

The author thanks colleagues in the telemental-health, digital-health-policy and game-theory research communities whose published work informs the parameterization and interpretation of the model.

Appendix A. Notation Summary

Symbol	Meaning
P	Set of populations: patients, providers, payers, policymakers
x, y, z, w	Frequency vectors of patient, provider, payer and policymaker strategies (each in Δ^2)
ξ	Joint state $(x, y, z, w) \in \Delta = \Delta^2 \times \Delta^2 \times \Delta^2 \times \Delta^2$
P_i, D_j, Y_k, R_ℓ	Pure strategies for patient i , provider j , payer k , policymaker ℓ ($i, j, k, \ell \in \{1, 2, 3\}$)
$u_i^A P, u_j^A D, u_k^A Y, u_\ell^A R$	Expected payoffs in each population
$\bar{u}^\alpha$	Population-mean payoff in population $\alpha \in \{P, D, Y, R\}$
$\omega_{\{ijkl\}}^\alpha$	Interaction payoff tensor for population α
$c_i^A P, c_j^A D, c_p$	Patient effort cost, provider operational cost, payer redesign cost (policy lever)
β, λ, τ	Outcome-linked reimbursement weight, equity weight, regulatory enforcement intensity (policy levers)
$R_j(x, z)$	Outcome-linked reimbursement bonus to provider strategy j
$e_i(y, z, w)$	Equity-protection function for patient strategy i
$\rho_k(x, y, w)$	Regulatory / reputational penalty on payer strategy k
$\sigma_i^A P, \sigma_j^A D$	AI-safety harm to patient strategy i ; AI-related liability to provider strategy j
A_ℓ, W, I	Audit-effort multiplier, welfare gain function, incident function for policymaker strategy ℓ
E_T, E_F, E_X, E_C	Transactional, fragmented, mixed and patient-centric equilibria
$\beta^*, \lambda^*, c_p^*, \tau^*$	Threshold values for the patient-centric ESS
V	Kullback–Leibler-divergence Lyapunov function in Theorem 1
D_{KL}	Kullback–Leibler divergence

References

- Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly B, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic — United States, January–March 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(43):1595–9.
- Bashshur R, Doarn CR, Frenk JM, Kvedar JC, Woolliscroft JO. Telemedicine and the COVID-19 pandemic, lessons for the future. *Telemed J E Health*. 2020;26(5):571–3.
- Mehrotra A, Bhatia RS, Snoswell CL. Paying for telemedicine after the pandemic. *JAMA*. 2021;325(5):431–2.
- Omboni S, Padwal RS, Alessa T, Benczur B, Green BB, Hubbard I, et al. The worldwide impact of telemedicine during COVID-19: current evidence and recommendations for the future. *Connect Health*. 2022; 1:7–35.
- Ohannessian R, Duong TA, Odone A. Global telemedicine

- implementation and integration within health systems to fight the COVID-19 pandemic: a call to action. *JMIR Public Health Surveill.* 2020;6(2): e18810.
6. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. *Fam Med Community Health.* 2020;8(3): e000530.
 7. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. *Telemed J E Health.* 2013;19(6):444–54.
 8. Hubley S, Lynch SB, Schneck C, Thomas M, Shore J. Review of key telepsychiatry outcomes. *World J Psychiatry.* 2016;6(2):269–82.
 9. Bashshur RL, Shannon GW, Bashshur N, Yellowlees PM. The empirical evidence for telemedicine interventions in mental disorders. *Telemed J E Health.* 2016;22(2):87–113.
 10. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44–56.
 11. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* 2019;380(14):1347–58.
 12. Kothari P. Artificial intelligence and emerging technologies in patient-centric digital health: a framework for adoption and governance. *Telehealth Med Today.* 2023;8(3):1–12.
 13. Stade EC, Stirman SW, Ungar LH, Boland CL, Schwartz HA, Yaden DB, et al. Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation. *NPJ Ment Health Res.* 2024;3(1):12.
 14. Lawrence HR, Schneider RA, Rubin SB, Matarić MJ, McDuff DJ, Jones Bell M. The opportunities and risks of large language models in mental health. *JMIR Ment Health.* 2024;11: e59479.
 15. U.S. Food and Drug Administration. Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan. Silver Spring (MD): FDA; 2021.
 16. European Parliament and Council. Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Off J Eur Union.* 2024; L 2024/1689.
 17. Epstein RM, Fiscella K, Lesser CS, Stange KC. Why the nation needs a policy push on patient-centered health care. *Health Aff (Millwood).* 2010;29(8):1489–95.
 18. Berwick DM. What 'patient-centered' should mean: confessions of an extremist. *Health Aff (Millwood).* 2009;28(4): w555–65.
 19. Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century.* Washington, DC: National Academies Press; 2001.
 20. Greene J, Hibbard JH, Sacks RM. Does the patient-centered medical home improve patient experience and quality of care? *Med Care.* 2017;55(7):657–63.
 21. Porter ME. What is value in health care? *N Engl J Med.* 2010;363(26):2477–81.
 22. Gerteis, M., Edgman-Levitan, S., Daley, J., & Delbanco, T. L. (1997). Through the Patient's eyes: understanding and promoting patient-centered care. *The Journal for Healthcare Quality (JHQ),* 19(3), 43.
 23. Constand MK, MacDermid JC, Dal Bello-Haas V, Law M. Scoping review of patient-centered care approaches in healthcare. *BMC Health Serv Res.* 2014; 14:271.
 24. Kotronoulas G, Kearney N, Maguire R, Harrow A, Di Domenico D, Croy S, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. *J Clin Oncol.* 2014;32(14):1480–501.
 25. Bombard Y, Baker GR, Orlando E, Fancott C, Bhatia P, Casalino S, et al. Engaging patients to improve quality of care: a systematic review. *Implement Sci.* 2018;13(1):98.
 26. Shigekawa E, Fix M, Corbett G, Roby DH, Coffman J. The current state of telehealth evidence: a rapid review. *Health Aff (Millwood).* 2020;39(4):661–8.
 27. Totten AM, Hansen RN, Wagner J, Stillman L, Ivlev I, Davis-O'Reilly C, et al. Telehealth for acute and chronic care consultations. *Comparative Effectiveness Review No. 216. AHRQ Publication No. 19-EHC012-EF.* Rockville (MD): Agency for Healthcare Research and Quality; 2019.
 28. Snoswell CL, Chelberg G, De Guzman KR, Haydon HM, Thomas EE, Caffery LJ, et al. The clinical effectiveness of telehealth: a systematic review of meta-analyses from 2010 to 2019. *J Telemed Telecare.* 2023;29(9):669–84.
 29. Bashshur RL, Shannon GW, Smith BR, Alverson DC, Antoniotto N, Barsan WG, et al. The empirical foundations of telemedicine interventions for chronic disease management. *Telemed J E Health.* 2014;20(9):769–800.
 30. Nouri S, Khoong EC, Lyles CR, Karliner L. Addressing equity in telemedicine for chronic disease management during the Covid-19 pandemic. *NEJM Catal Innov Care Deliv.* 2020;1(3).
 31. Rodriguez JA, Betancourt JR, Sequist TD, Ganguli I. Differences in the use of telephone and video telemedicine visits during the COVID-19 pandemic. *Am J Manag Care.* 2021;27(1):21–6.
 32. Eberly LA, Kallan MJ, Julien HM, Haynes N, Khatana SAM, Nathan AS, et al. Patient characteristics associated with telemedicine access for primary and specialty ambulatory care during the COVID-19 pandemic. *JAMA Netw Open.* 2020;3(12): e2031640.
 33. Khoong EC, Butler BA, Mesina O, Su G, DeFries TB, Nijagal M, et al. Patient interest in and barriers to telemedicine video visits in a multilingual urban safety-net system. *J Am Med Inform Assoc.* 2021;28(2):349–53.
 34. Heinz MV, Mackin DM, Trudeau BM, Bhattacharya S, Wang Y, Banta HA, et al. Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI.* 2025;2(4): A10a2400802.
 35. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled

- trial. JMIR Ment Health. 2017;4(2): e19.
36. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth. 2018;6(11): e12106.
 37. De Choudhury M, Pendse SR, Kumar N. Benefits and harms of large language models in digital mental health. arXiv. 2023:2311.14693.
 38. Rajkumar R, Conway PH, Tavenner M. CMS—engaging multiple payers in payment reform. JAMA. 2014;311(19):1967–8.
 39. Liao JM, Navathe AS, Werner RM. The impact of Medicare's alternative payment models on the value of care. Annu Rev Public Health. 2020; 41:551–65.
 40. Maynard Smith J, Price GR. The logic of animal conflict. Nature. 1973;246(5427):15–8.
 41. Maynard Smith J. Evolution and the Theory of Games. Cambridge: Cambridge University Press; 1982.
 42. Taylor PD, Jonker LB. Evolutionary stable strategies and game dynamics. Math Biosci. 1978;40(1–2):145–56.
 43. Hofbauer J, Sigmund K. Evolutionary Games and Population Dynamics. Cambridge: Cambridge University Press; 1998.
 44. Hofbauer J, Sigmund K. Evolutionary game dynamics. Bull Am Math Soc. 2003;40(4):479–519.
 45. Weibull, J. (1995). Evolutionary Game Theory. Cambridge, MA: MIT Press.
 46. Sandholm, W. H. (2010). *Population games and evolutionary dynamics* (Vol. 25). Cambridge: MIT press.
 47. Nowak MA. Evolutionary Dynamics: Exploring the Equations of Life. Cambridge (MA): Belknap Press; 2006.
 48. Simon, H. A. (1997). *Models of bounded rationality, volume 3: Empirically grounded economic reason*. The MIT Press.
 49. Sen, A. (1985). *Commodities and Capabilities*. Amsterdam, New York: North-Holland.
 50. Backhaus A, Agha Z, Maglione ML, Repp A, Ross B, Zuest D, et al. Videoconferencing psychotherapy: a systematic review. Psychol Serv. 2012;9(2):111–31.
 51. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev. 2015;(9):CD002098.
 52. Sharma A, Lin IW, Miner AS, Atkins DC, Althoff T. Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. Nat Mach Intell. 2023;5(1):46–57.
 53. 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.
 54. 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.
 55. Khan AA, Khan SU, Zia-Ul-Haq M. A game-theoretic approach for modeling patient adherence in chronic disease management. Health Care Manag Sci. 2018;21(4):548–62.
 56. Fisher RA. The Genetical Theory of Natural Selection. Oxford: Clarendon Press; 1930.
 57. Lewontin RC. Evolution and the theory of games. J Theor Biol. 1961;1(3):382–403.
 58. Hamilton WD. Extraordinary sex ratios. Science. 1967;156(3774):477–88.
 59. Roth AE, Sönmez T, Ünver MU. Kidney exchange. Q J Econ. 2004;119(2):457–88.
 60. Nowak MA, Sasaki A, Taylor C, Fudenberg D. Emergence of cooperation and evolutionary stability in finite populations. Nature. 2004;428(6983):646–50.
 61. Cressman R, Tao Y. The replicator equation and other game dynamics. Proc Natl Acad Sci USA. 2014;111(Suppl 3):10810–7.
 62. Zeeman EC. Population dynamics from game theory. In: Nitecki Z, Robinson C, editors. Global Theory of Dynamical Systems. Lecture Notes in Mathematics, vol 819. Berlin: Springer; 1980. p. 471–97.
 63. Bishop DT, Cannings C. A generalised war of attrition. J Theor Biol. 1978;70(1):85–124.
 64. Selten R. A note on evolutionarily stable strategies in asymmetric animal conflicts. J Theor Biol. 1980;84(1):93–101.
 65. Friedman D. Evolutionary games in economics. Econometrica. 1991;59(3):637–66.
 66. Sandholm WH. Evolutionary game theory. In: Meyers RA, editor. Encyclopedia of Complexity and Systems Science. New York: Springer; 2009. p. 3176–205.
 67. Foster D, Young P. Stochastic evolutionary game dynamics. Theor Popul Biol. 1990;38(2):219–32.
 68. Kandori M, Mailath GJ, Rob R. Learning, mutation, and long run equilibria in games. Econometrica. 1993;61(1):29–56.
 69. Schlag KH. Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. J Econ Theory. 1998;78(1):130–56.
 70. Imhof LA, Fudenberg D, Nowak MA. Evolutionary cycles of cooperation and defection. Proc Natl Acad Sci USA. 2005;102(31):10797–800.
 71. Traulsen A, Hauert C. Stochastic evolutionary game dynamics. In: Schuster HG, editor. Reviews of Nonlinear Dynamics and Complexity, Vol 2. Weinheim: Wiley-VCH; 2009. p. 25–61.
 72. Nowak MA, May RM. Evolutionary games and spatial chaos. Nature. 1992;359(6398):826–9.
 73. Lieberman E, Hauert C, Nowak MA. Evolutionary dynamics on graphs. Nature. 2005;433(7023):312–6.
 74. Ohtsuki H, Hauert C, Lieberman E, Nowak MA. A simple rule for the evolution of cooperation on graphs and social networks. Nature. 2006;441(7092):502–5.
 75. Szabó G, Fáth G. Evolutionary games on graphs. Phys Rep. 2007;446(4–6):97–216.
 76. Perc M, Szolnoki A. Coevolutionary games—a mini review. Biosystems. 2010;99(2):109–25.
 77. Allen B, Lippner G, Chen YT, Fotouhi B, Momeni N, Yau ST, Nowak MA. Evolutionary dynamics on any population structure. Nature. 2017;544(7649):227–30.
 78. Antal T, Ohtsuki H, Wakeley J, Taylor PD, Nowak MA. Evolution of cooperation by phenotypic similarity. Proc Natl Acad Sci USA. 2009;106(21):8597–600.

79. Tarnita CE, Antal T, Ohtsuki H, Nowak MA. Evolutionary dynamics in set structured populations. *Proc Natl Acad Sci USA*. 2009;106(21):8601–4.
80. Bowles S, Gintis H. *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton: Princeton University Press; 2011.
81. Rand DG, Nowak MA. Human cooperation. *Trends Cogn Sci*. 2013;17(8):413–25.
82. Adami C, Schossau J, Hintze A. Evolutionary game theory using agent-based methods. *Phys Life Rev*. 2016; 19:1–26.
83. Bauch CT, Earn DJD. Vaccination and the theory of games. *Proc Natl Acad Sci USA*. 2004;101(36):13391–4.
84. Galvani AP, Reluga TC, Chapman GB. Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum. *Proc Natl Acad Sci USA*. 2007;104(13):5692–7.
85. Fu F, Rosenbloom DI, Wang L, Nowak MA. Imitation dynamics of vaccination behaviour on social networks. *Proc R Soc B*. 2011;278(1702):42–9.
86. Reluga TC, Galvani AP. A general approach for population games with application to vaccination. *Math Biosci*. 2011;230(2):67–78.
87. Wang Z, Bauch CT, Bhattacharyya S, d'Onofrio A, Manfredi P, Perc M, et al. Statistical physics of vaccination. *Phys Rep*. 2016; 664:1–113.
88. Cardillo A, Reyes-Suárez C, Naranjo F, Gómez-Gardeñes J. Evolutionary vaccination dilemma in complex networks. *Phys Rev E*. 2013;88(3):032803.
89. Tanimoto J. *Evolutionary Games with Sociophysics: Analysis of Traffic Flow and Epidemics*. Singapore: Springer; 2019.
90. Kaier K, Frank U. Measuring the externality of antibacterial use from promoting antimicrobial resistance. *Pharmacoeconomics*. 2009;27(2):93–6.
91. Hardt M, Megiddo N, Papadimitriou C, Wootters M. Strategic classification. In: *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*. New York: ACM; 2016. p. 111–22.
92. Liu LT, Dean S, Rolf E, Simchowitz M, Hardt M. Delayed impact of fair machine learning. In: *Proceedings of the 35th International Conference on Machine Learning*. PMLR; 2018. p. 3150–8.
93. Roth AE, Sönmez T, Ünver MU. Efficient kidney exchange: coincidence of wants in markets with compatibility-based preferences. *Am Econ Rev*. 2007;97(3):828–51.
94. Roth AE, Peranson E. The redesign of the matching market for American physicians: some engineering aspects of economic design. *Am Econ Rev*. 1999;89(4):748–80.
95. Leimar O, McNamara JM. Game theory in biology: 50 years and onwards. *Philos Trans R Soc B*. 2023;378(1876):20210509.
96. Traulsen A, Glynatsi NE. The future of theoretical evolutionary game theory. *Philos Trans R Soc B*. 2023;378(1876):20210508.
97. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. *J Med Internet Res*. 2018;20(4): e150.
98. Page KM, Nowak MA. Unifying evolutionary dynamics. *J Theor Biol*. 2002;219(1):93–8.
99. Moore J, Grabb D, Agnew W, Klyman K, Chancellor S, Ong DC, et al. Expressing stigma and inappropriate responses prevents LLMs from safely replacing mental health providers. In: *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. New York: ACM; 2025.
100. Hipgrave L, Goldie J, Dennis S, Coleman A. Balancing risks and benefits: clinicians' perspectives on the use of generative AI chatbots in mental healthcare. *Front Digit Health*. 2025; 7:1606291.
101. Figueroa JF, Wadhera RK, Frakt AB, Fonarow GC, Heidenreich PA, Xu H, et al. Quality of care and outcomes among Medicare beneficiaries dually eligible for Medicare and Medicaid. *Health Aff (Millwood)*. 2020;39(11):1894–903.
102. U.S. Food and Drug Administration. *Generative artificial intelligence-enabled digital mental health medical devices*. Digital Health Advisory Committee briefing document, November 6, 2025. Silver Spring (MD): FDA; 2025.
103. Utah State Legislature. *House Bill 452, 2025 General Session: Artificial Intelligence Amendments (mental health chatbot disclosure)*. Salt Lake City (UT): Utah State Legislature; 2025.
104. Nevada Legislature. *Assembly Bill 406, 2025 Regular Session: An Act relating to artificial intelligence; restrictions on use of AI to provide mental or behavioral healthcare*. Carson City (NV): Nevada Legislature; 2025.
105. Illinois General Assembly. *Public Act 104-0054 (HB 1806): Wellness and Oversight for Psychological Resources Act*. Springfield (IL): Illinois General Assembly; 2025.
106. European Commission. *Navigating the AI Act: questions and answers*. Brussels: European Commission, Directorate-General for Communications Networks, Content and Technology; 2026.
107. Maynard Smith J. The theory of games and the evolution of animal conflicts. *J Theor Biol*. 1974;47(1):209–21.
108. Nash JF. Equilibrium points in n-person games. *Proc Natl Acad Sci USA*. 1950;36(1):48–9.
109. Nash JF. Non-cooperative games. *Ann Math*. 1951;54(2):286–95.
110. Von Neumann J, & Morgenstern, O. (1944). *Game Theory and Economic Decisions*.
111. Arrow KJ. Uncertainty and the welfare economics of medical care. *Am Econ Rev*. 1963;53(5):941–73.
112. Akerlof GA. The market for 'lemons': quality uncertainty and the market mechanism. *Q J Econ*. 1970;84(3):488–500.
113. Rothschild M, Stiglitz J. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Q J Econ*. 1976;90(4):629–49.
114. Holmström B. Moral hazard and observability. *Bell J Econ*. 1979;10(1):74–91.
115. Grossman SJ, Hart OD. An analysis of the principal-

- agent problem. *Econometrica*. 1983;51(1):7–45.
116. McGuire TG. Physician agency. In: Culyer AJ, Newhouse JP, editors. *Handbook of Health Economics*, Vol 1A. Amsterdam: Elsevier; 2000. p. 461–536.
 117. Pauly MV. The economics of moral hazard: comment. *Am Econ Rev*. 1968;58(3):531–7.
 118. Ellis RP, McGuire TG. Provider behavior under prospective reimbursement: cost sharing and supply. *J Health Econ*. 1986;5(2):129–51.
 119. Ma CA. Health care payment systems: cost and quality incentives. *J Econ Manag Strategy*. 1994;3(1):93–112.
 120. Brekke KR, Siciliani L, Straume OR. Hospital competition and quality with regulated prices. *Scand J Econ*. 2011;113(2):444–69.
 121. Gaynor M, Town RJ. Competition in health care markets. In: Pauly MV, McGuire TG, Barros PP, editors. *Handbook of Health Economics*, Vol 2. Amsterdam: Elsevier; 2011. p. 499–637.
 122. Bauch CT, Galvani AP. Social factors in epidemiology. *Science*. 2013;342(6154):47–9.
 123. Centola D. *How Behavior Spreads: The Science of Complex Contagions*. Princeton: Princeton University Press; 2018.
 124. Newton J. Evolutionary game theory: a renaissance. *Games*. 2018;9(2):31.
 125. Kahneman, D., & Egan, P. (2011). *Thinking, fast and slow* (Vol. 1, No. 0, pp. 1-5). New York: Farrar, straus and Giroux.
 126. Wiederhold BK. The dawn of AI in mental health: balancing innovation, risk, and regulation. *Cyberpsychol Behav Soc Netw*. 2026;29(1):1–3.