

Designing AI That Patients Trust: Moral Foundations for Human-Centered Healthcare

WHITE PAPER

By Richa Nevatia^{1*}, Lanyi Zhu^{1*}, Emma Wigdahl², Alla Zarifyan², Andrew Kucheriavy²

**These authors contributed equally*

¹ University of Washington Communication Leadership

² Intechnic



Contents

01 Executive Summary

02 Introduction

03 The Haidt Moral Dimensions

05 Methods

07 Findings: What Do Patients Want?

16 Recommendations

20 Long-Term Strategic Value

23 Conclusion

24 Author's Note

25 About the Authors

Executive Summary

As generative AI becomes increasingly embedded in healthcare delivery, understanding how patients perceive and judge its use is vital. This white paper explores the moral triggers and mitigators that shape patient acceptance of AI through a national survey of 275 U.S. patients, grounded in social psychologist Jonathan Haidt's moral foundations theory.

Key findings reveal that 38.5% of U.S. patients would reject any use of AI in their healthcare outright, while the remaining 61.5% are highly sensitive to issues of privacy, control, data management and expert oversight. Patients are most

supportive of AI when it offers tangible benefits, complies with HIPAA, allows data opt-outs, and includes clear validation by medical professionals.

Importantly, moral judgments highly engage five of the six foundational dimensions and extend beyond safety concerns to include autonomy, fairness, and transparency. The paper provides evidence-based recommendations spanning product design, communications, and organizational strategy for developing AI systems that align with human values, demonstrating that ethical implementation is not only a compliance necessity but a long-term competitive advantage.



Introduction

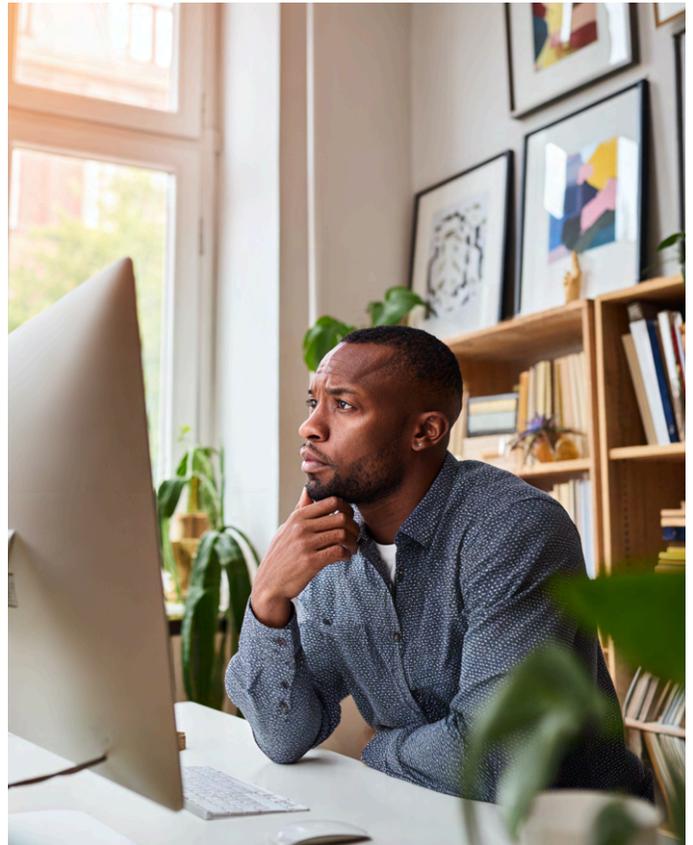
Generative AI is gaining traction in healthcare, with its increasing usage in patient portals, personalized treatment plans, and adherence tools. The notion that AI could free clinicians from electronic health record (EHR) busy work, improve patient adherence to treatments, and give clinicians more time with patients holds strong appeal. So does the promise of greater efficiency, sharper diagnostics, and better outcomes.

Across the industry, there's growing hope that artificial intelligence might make care more human, not less.

But these gains won't come automatically. They depend on trust, and trust begins with understanding how people experience, interpret, and judge the presence of AI in their healthcare journey.

This white paper explores what drives patients to accept or reject AI in healthcare. We don't focus on any specific product or platform. Instead, we examine the underlying moral reflexes that shape patients' reactions.

The goal is to guide the design of AI systems that reflect patient values from the outset, before disappointment, resistance, or regulation force a reactive course correction.



How can we anticipate those reflexes? What kind of framework can help us understand the early signals of discomfort or approval, before the technology is fully deployed?

To explore that question, we applied Jonathan Haidt's (2012) moral foundations theory, a cross-cultural framework that identifies six core dimensions of fast, intuitive moral judgment. This model has helped explain reactions to everything from public health to politics to environmental ethics. Used here, it provides a way to foresee which scenarios of AI will resonate with patients and which are likely to raise red flags before the consequences play out in practice.

6 Dimensions of the Haidt Moral Foundations Theory

Haidt's (2012)¹ moral foundations theory identifies six instinctive dimensions that shape how people make rapid moral judgments. These dimensions help explain

both what sparks moral outrage and what might defuse it. Across nearly any issue, human reactions typically draw on at least one of these foundations:



Care vs Harm

This foundation reflects our deep-rooted instinct to nurture and protect. It makes us alert to signs of suffering and need, driving strong reactions to cruelty, neglect, or indifference. Compassion, empathy, and concern for others live here.



Fairness vs Cheating

People quickly sense when the rules of fair exchange are broken. This foundation fuels anger at freeloaders, cheaters, or manipulators and reinforces the belief that effort and reward should align.



Loyalty vs Betrayal

Humans are social beings, wired to form alliances and protect their groups. Loyalty creates cohesion and inspires sacrifice for the common good. Betrayal (both real or perceived) can provoke intense backlash.



Authority vs Subversion

Respect for legitimate authority helps societies maintain order and structure. This foundation supports hierarchies, traditions, and shared expectations about roles. When authority is undermined or systems break down, people often feel uneasy or resistant.

¹ Jonathan Haidt, *The Righteous Mind: Why Good People Are Divided by Politics and Religion* (New York: Vintage Books, 2012).



Sanctity vs Degradation

What began as a sensitivity to physical contamination has evolved into a concern for moral and spiritual purity. This foundation drives visceral reactions (often disgust) toward perceived corruption, defilement, or violations of sacred values.



Liberty vs Oppression

This dimension reflects a basic desire for autonomy and freedom from control. People resist coercion and deeply value the right to make their own choices, especially in personal domains like health.

People resist coercion and deeply value the right to make their own choices, especially in personal domains like health.

This framework guided our approach to studying AI in healthcare. Rather than focusing solely on technical performance, we designed scenarios that tapped into these deeper moral instincts. By ensuring our survey engaged all six foundations, we captured a broader spectrum of human reactions, not just concerns about safety or accuracy, but gut-level responses about control, fairness, and trust.

Knowing which scenarios patients consider “not okay” versus “okay if” offers practical guidance for how AI should be designed, introduced, communicated, and governed. These insights can inform everything from product features to interface language, help

desk scripts, PR strategy, UX content, and even regulatory frameworks.

Ultimately, our goal is to help healthcare providers and technology developers design AI systems that earn and sustain patient trust by safeguarding privacy, reinforcing medical expertise, and aligning with both ethical principles and business goals.

Methods

Before launching the main survey, we conducted two pilot rounds using convenience and snowball sampling. Participants included graduate students in the Communications Leadership master's program at the University of Washington and members of their broader networks. These pilots helped us refine the scenarios to be clear, intuitive, and understandable across a wide range of English fluency levels.

To further improve real-world relevance, we partnered with [Intechnic](#), a leader in patient-experience design that integrates human-centered research with cutting-edge AI methodologies. Their user and market researchers edited the scenarios for clarity, tone, and plausibility, ensuring they would feel realistic and relevant to everyday healthcare contexts.

In the final survey, respondents were asked to imagine an artificial intelligence (AI) system that provides health information. They then reviewed 33 ways that such a system could be designed, used, or behave in real-world settings. For each scenario, they indicated how much it would increase or decrease their support for using AI in healthcare.

Responses were recorded on a 7-point Likert scale, ranging from "Greatly decrease" to "Greatly increase" support. The survey also included screening questions, AI usage history, and demographic items to provide context for responses.

We recruited participants through SurveyMonkey's professional sample platform. The initial sample included 520 U.S. adults who are willing to take the

33

scenarios in which an AI health system could be designed, used, or behave in real-world settings, rated on a 7-point Likert scale

-3

Greatly
Decrease

-2

Moderately
Decrease

-1

Slightly
Decrease

0

No
change

1

Slightly
Increase

2

Moderately
Increase

3

Greatly
Increase

survey. A total of 486 of them are qualified by confirming they had used any U.S. healthcare service, such as doctor visits, hospitals, clinics, or telehealth, within the last two years.

Of those, 187 (38.5%) respondents stated they would refuse to use any form of AI in their healthcare. Once that position was recorded, they were respectfully removed from the rest of the survey. This follows a common practice in user research: participants who categorically reject a concept are unlikely to provide meaningful feedback on specific implementation scenarios. We focused the remainder of the survey on those who were at least open to conditional use.

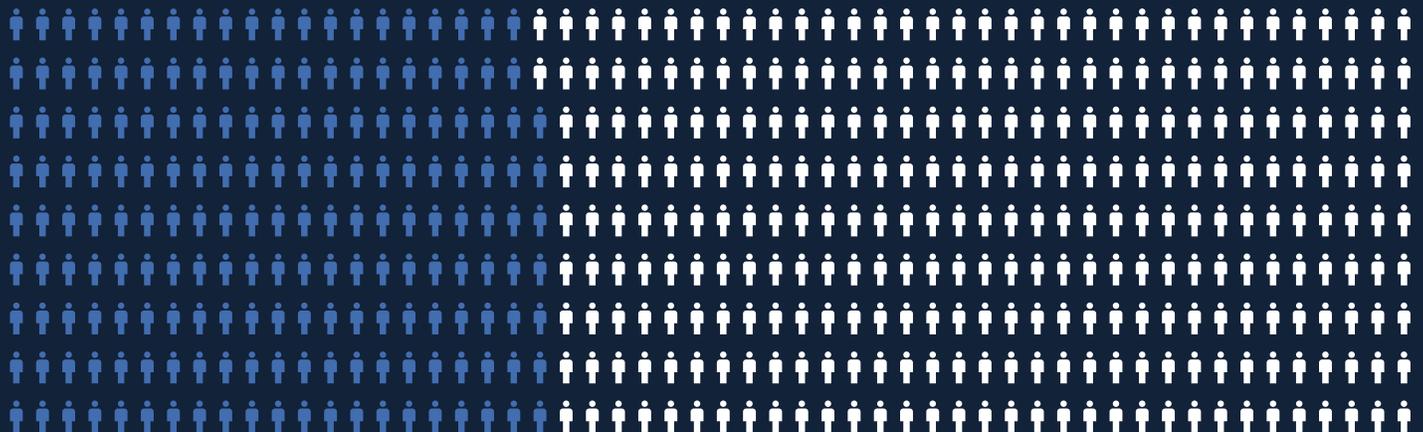
An additional 24 responses were excluded due to poor data quality, such as

completion times under 2 minutes or selecting identical responses across all scenarios. The final sample included 275 thoughtful, engaged U.S. respondents.

This approach allowed us to focus our analysis on the most strategically important segment: patients who are hesitant, cautious, or curious about healthcare AI, not just enthusiasts or categorical rejectors. This is precisely the audience where thoughtful implementation can make the difference between acceptance and resistance.

Our findings provide insight into how support for AI in healthcare can be built, shaped, or lost through specific design and communication choices.

187 out of 486 participants who have had a U.S. health service within the past 2 years would refuse to use AI in their healthcare. *(These participants were dismissed from the remainder of the study)*



Findings: What do Patients Want?

1. A Significant Minority Still Rejects AI Entirely

An early screening question asked respondents to agree or disagree with the statement: **"I would refuse to use any form of artificial intelligence (AI) involved in my healthcare."**

A full 38.5% of respondents answered true. This is both striking and strategically important. It suggests that patient rejection of AI is not niche or fringe, but mainstream.

Rather than dismiss those who reject AI in healthcare, healthcare leaders should see it as a design and messaging challenge, shaped by transparency, education, and trust-building.

2. Non-Rejectors Still Hold Low Confidence in Using AI

What, then, about the 61.5% who remain open to AI? Their attitudes reveal both opportunity and complexity.

Among those who did not reject AI outright, most had **some exposure to generative AI**, with 71% reporting prior use (Figure 1). Yet only 27% felt **very confident** in using it (Figure 2). This confidence gap is crucial: even among those open to AI, hesitation and lack of experience runs deep.

Have you ever personally used generative AI tools (like ChatGPT, DALL-E) or similar systems that create content using AI?

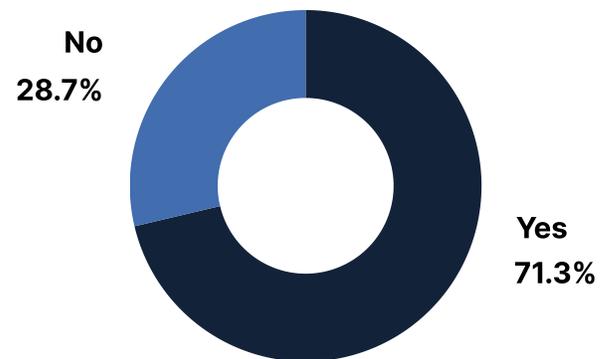


Figure 1. Use of generative AI among non-rejectors (US patient sample n=275, June 2025).

How confident are you using generative AI?

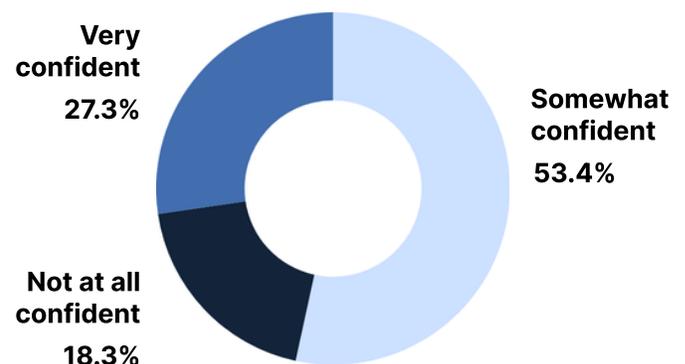


Figure 2. AI confidence among non-rejectors (US patient sample n=275, June 2025).

These findings reinforce the need for patient-centered onboarding and messaging. Confidence is a trust challenge, not just a technical challenge.

3. Scenario Testing Reveals Clear Patterns of Support and Resistance

Respondents reviewed **33 real-world AI design** scenarios and rated how each would affect their support for AI in healthcare, on a 7 point scale of "Greatly decrease," "Moderately decrease," "Slightly decrease," "No change," "Slightly increase," "Moderately increase" and "Greatly increase" (Figure 3). These scenarios ranged from data privacy and consent mechanics to expert validation and marketing usage.

Notably, **confidence in AI use correlated strongly ($p < 0.001$) with reactions to these scenarios**. To illustrate the correlation with more details, we plotted the mean ratings for two groups: those who are very confident and those who are not at all confident in using AI in a single figure.

Users with low confidence were more easily triggered by problematic designs. But importantly, both low- and high-confidence users agreed on which features made AI feel acceptable. We also color-coded all the design scenarios according to the moral dimensions for a richer understanding of the data in terms of its correlation to our grounding theory in Haidt (2012).



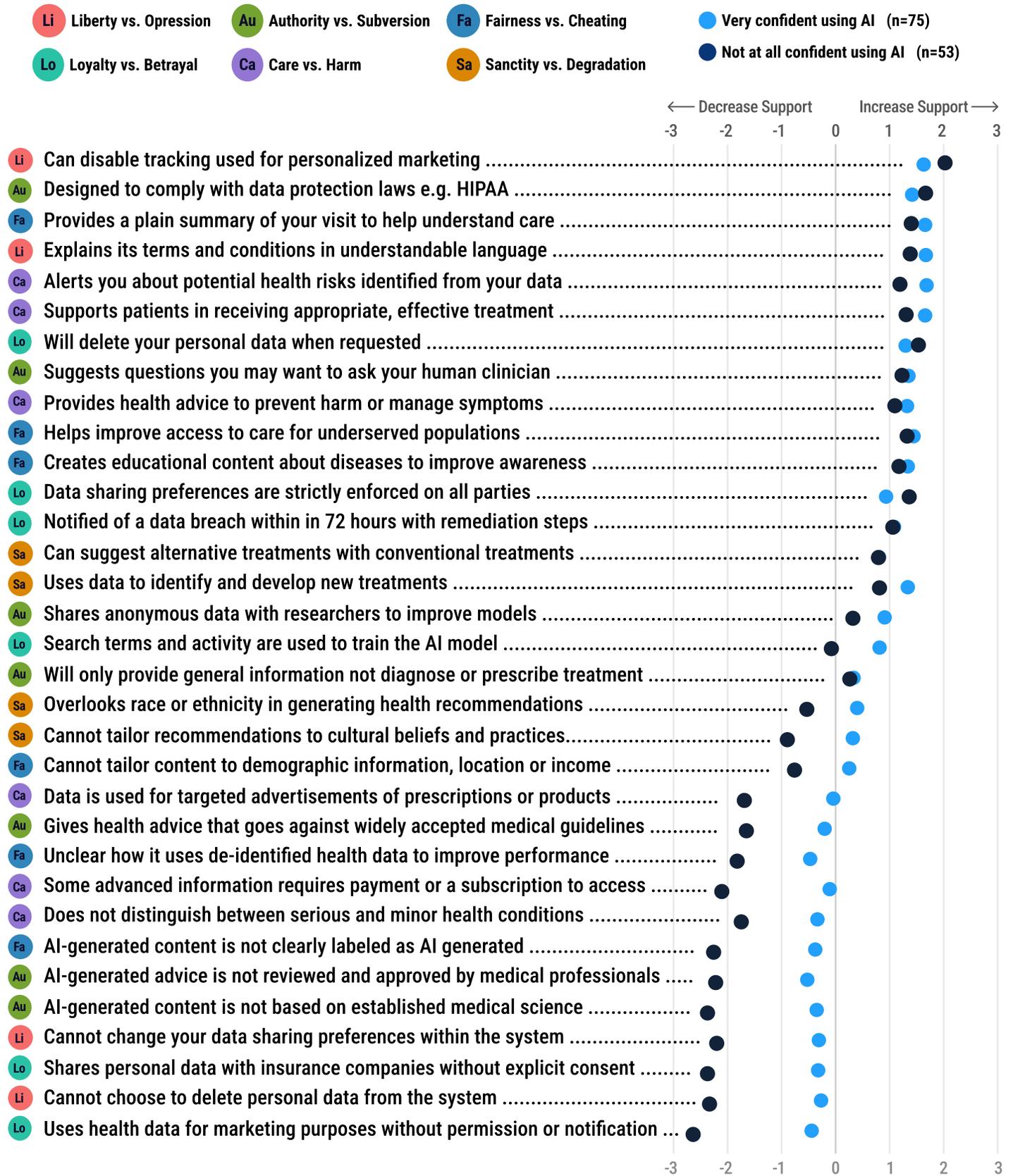


Figure 3. Healthcare AI scenarios most increasing support (top) to most decreasing support (bottom) among all respondents (US patient sample n=275, June 2025). They are broken out by AI confidence and dimensions of moral judgments (Haidt, 2012).

To understand what drives these moral reflexes, we analyzed the scenarios that most strongly influenced support.

4. What Triggers Rejection vs. What Builds Support

We conducted a focused analysis of the **top 7 most triggering** and **top 7 most mitigating** scenarios (Figure 4). To better showcase such insights, we aggregated the bottom 2 point scale (Greatly Decrease and Moderately Decrease) and the top 2 point scale (Greatly Increase and Moderately Increase). The results reveal a consistent pattern:

Most Triggering Scenarios:

- Use of health data for targeted marketing
- Inability to delete personal data
- Sharing data with unauthorized third parties
- Displaying information not approved by medical professionals

Most Mitigating Scenarios:

- Control over marketing and deletion settings
- HIPAA compliance and legal transparency
- Clear benefits like better treatment plans or visit summaries
- Visible expert oversight

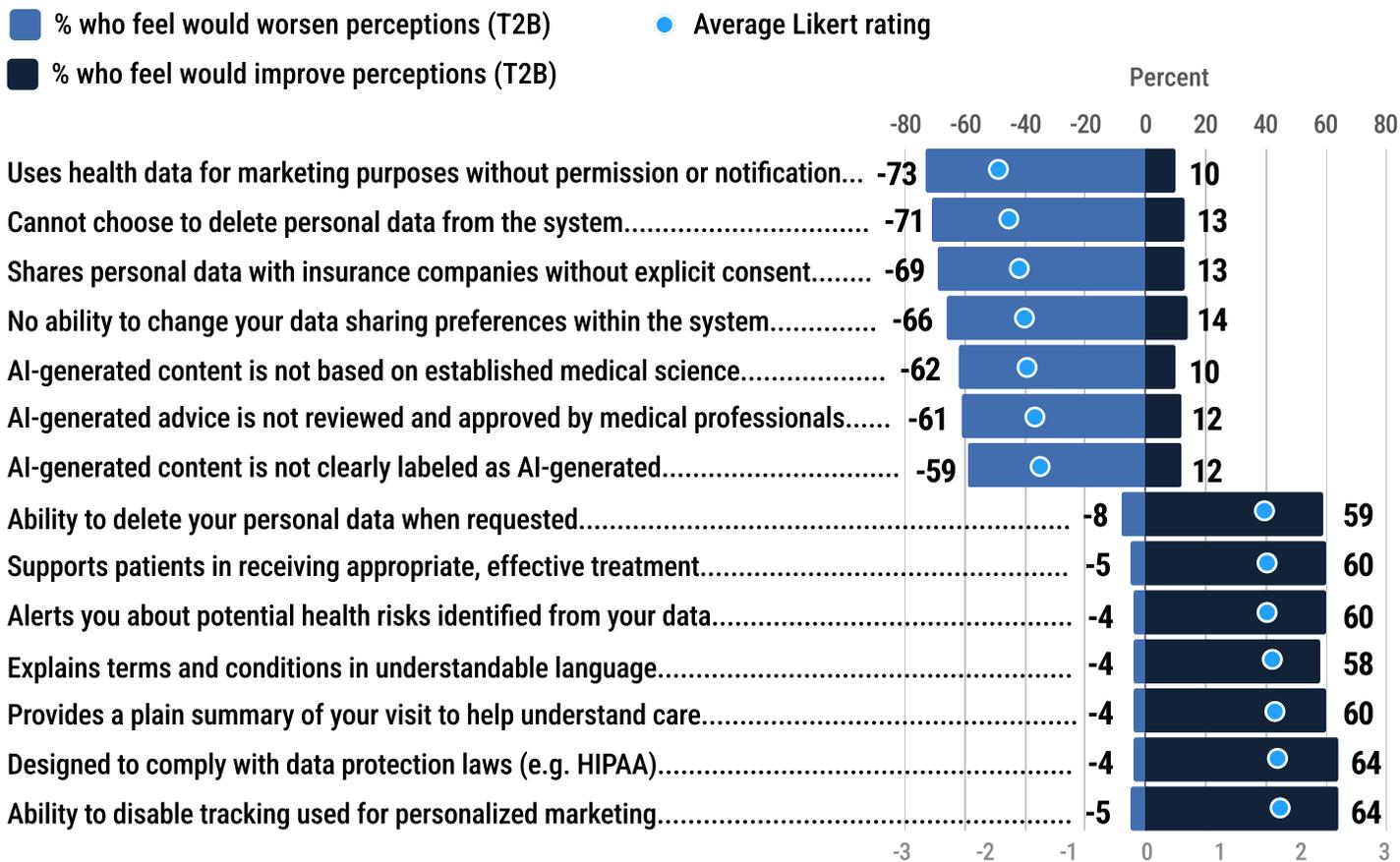


Figure 4. Top 7 most triggering (top) and most mitigating (bottom) scenarios affecting support for AI use in healthcare (US patient sample n=275, June 2025).

These patterns reveal a clear roadmap for implementation: patients want control over their data, transparency about AI use, visible expert involvement, and tangible benefits that improve their care experience.

Transparency is foundational to user trust in AI-driven healthcare systems.

5. Moral Foundations Shape Patient Judgments

We mapped responses to the six moral dimensions outlined in Haidt's framework (Figure 5). We initially expected Care/Harm to dominate (given healthcare's inherent ties to well-being), however, our research revealed at least **five of the six dimensions** were consistently engaged. We use μ to show the mean of the responses.

- **Liberty/Oppression** ($\mu=2.16$) had the strongest impact, underscoring how much autonomy and consent matter to users.
- **Loyalty/Betrayal** ($\mu=1.94$), **Care/Harm** ($\mu=1.88$), **Authority/Subversion** ($\mu=1.8$) and **Fairness/Cheating** ($\mu=1.77$), also strongly shaped judgments, especially around data use, expert oversight, and equitable treatment.
- **Sanctity/Degradation** ($\mu=1.44$) was the least dominant but still present, often tied to fears of overreach or dehumanization.

Patients are not just worried about physical harm — they are evaluating AI through emotional, ethical, and societal lenses.

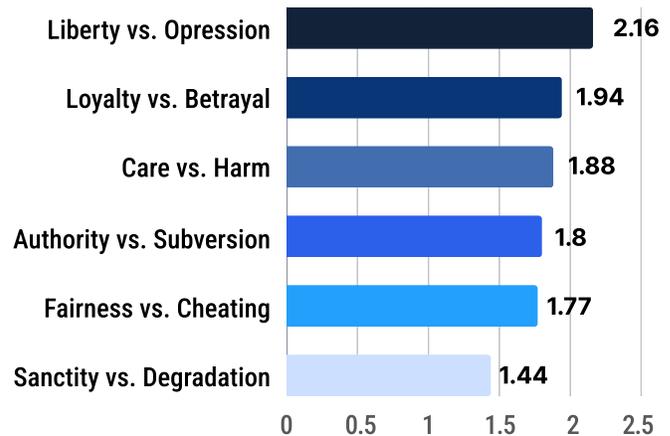


Figure 5. Average means of Haidt's moral dimensions at play in user judgments (US patient sample $n=275$, June 2025).

This broad engagement across moral foundations explains why purely technical messaging often fails. Effective communication must address not just safety and efficacy, but also autonomy, fairness, and trust.



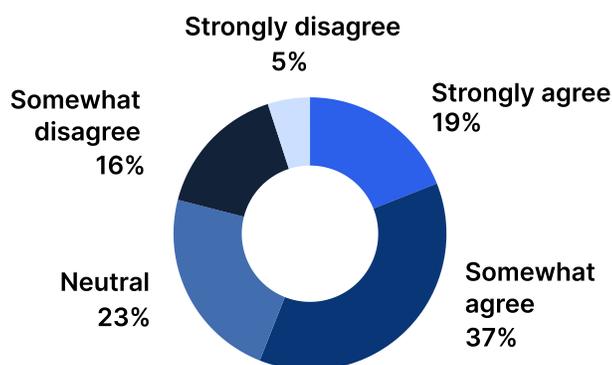
6. What Patients Want to Hear

Across scenarios, a clear communication pattern emerged. Patients respond positively to messages that affirm four key principles:

"You are in control. Your private relationship with your provider is protected. Medical experts are always involved. And you'll benefit in meaningful ways from our AI innovations."

This framework addresses multiple moral sensitivities simultaneously, moving beyond compliance-focused messaging to acknowledge the full spectrum of patient concerns.

"I am optimistic that AI will improve our lives and the world we live in."

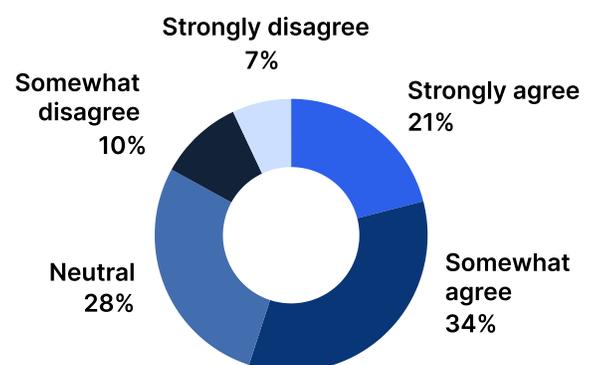


7. Optimism, Familiarity, and Demographic Insights

The final set of questions explored broader attitudes among non-rejectors:

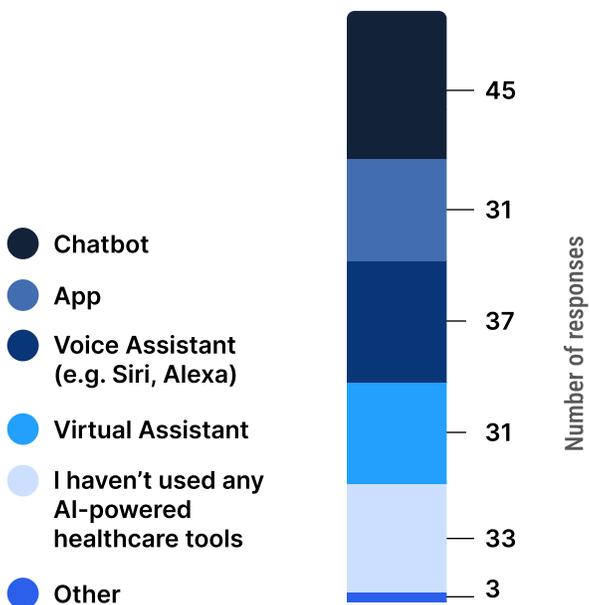
- A majority reported optimism about AI in general and in healthcare (Figures 6-7).
- Two-thirds to three-quarters felt they had already used AI tools to manage or answer questions about their health (Figures 8-9).
- This shows AI is not entirely foreign: there is a foundation of familiarity, even if confidence remains low.

"I am optimistic that AI will improve my experience with healthcare."

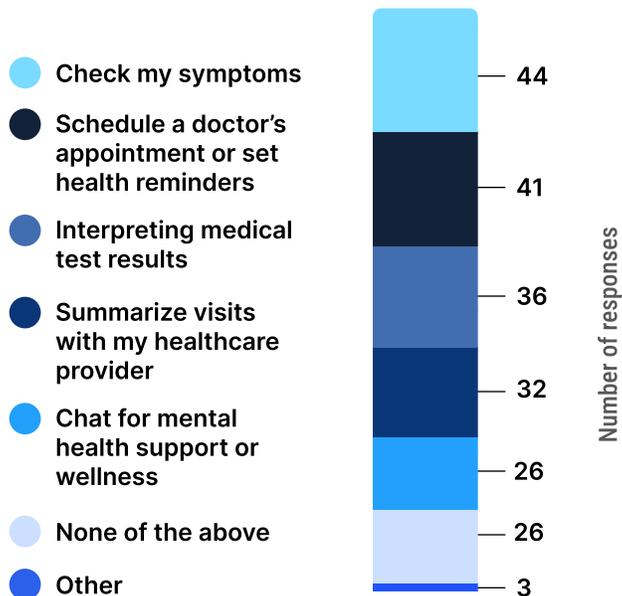


Figures 6-7. AI optimism among non-rejectors (US patient sample n=275, June 2025).

What type of AI-powered healthcare tools have you used to seek answers to healthcare questions? (Select all that apply)



Besides healthcare questions, what else have you used an AI-powered tool for? (Select all the apply)



Figures 8–9. AI usage among non-rejectors (US patient sample $n=275$, June 2025).

Demographically, the sample was gender-balanced. Educational attainment was higher than the U.S. average: **63% held a bachelor's degree or higher**, compared to the U.S. Census figure of 40%. This educational skew offers both insight and caution.

Current messaging strategies may not reach less educated populations who are equally deserving of the protections and benefits that ethical AI can offer.

This underscores a final communication challenge: **messaging to the AI-hesitant should match their literacy levels**, with clear, respectful language that avoids technical jargon and prioritizes patient agency.

Together, these findings reveal that patient hesitancy toward healthcare AI is neither irrational nor insurmountable. Instead, it reflects predictable human responses to questions of autonomy, trust, and benefit. The path forward lies in designing AI systems that honor these concerns while delivering genuine value.

What kind of health coverage do you have, if any? Select all that apply.

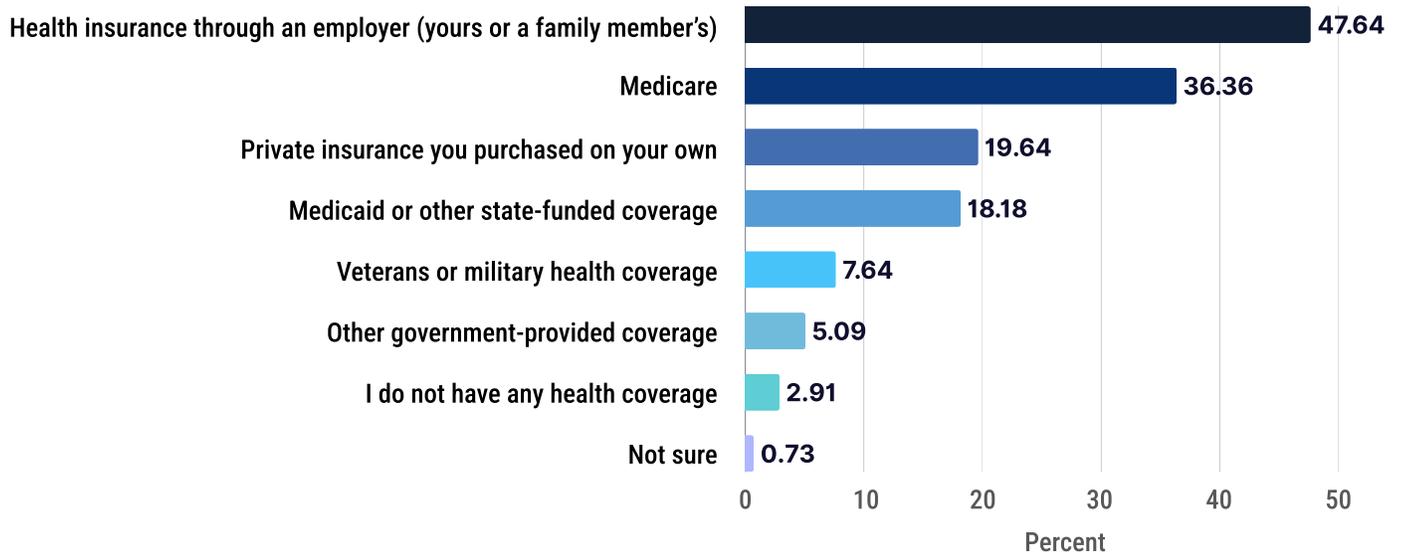


Figure 10. Health coverage amongst respondents (US patient sample n=275, June 2025).

Which of the follow types of medical care have you personally received?

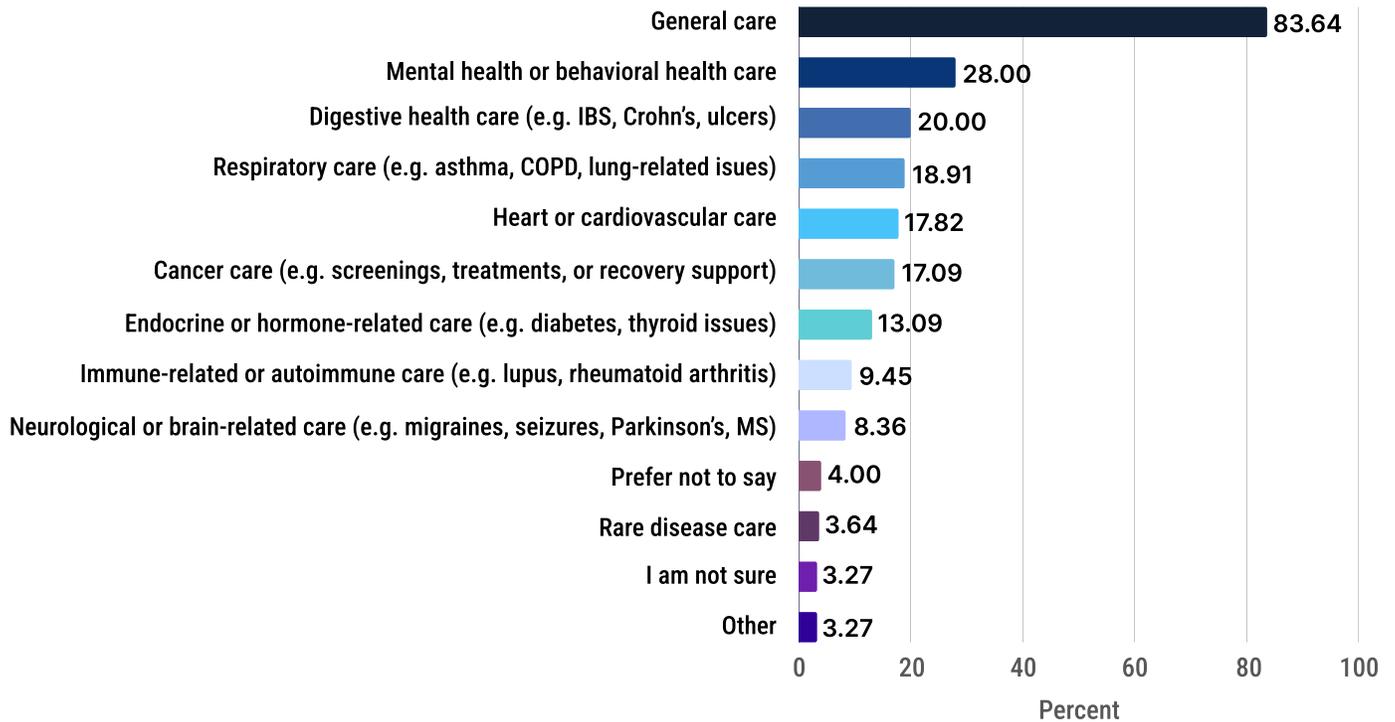


Figure 11. Types of medical care received by respondents (US patient sample n=275, June 2025).

Respondent gender

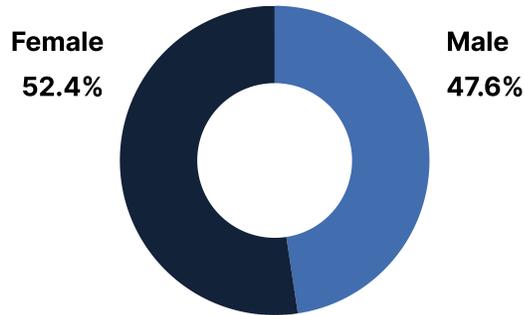


Figure 12. Respondent genders of non-rejectors (US patient sample n=275, June 2025).

What is your highest level of education?

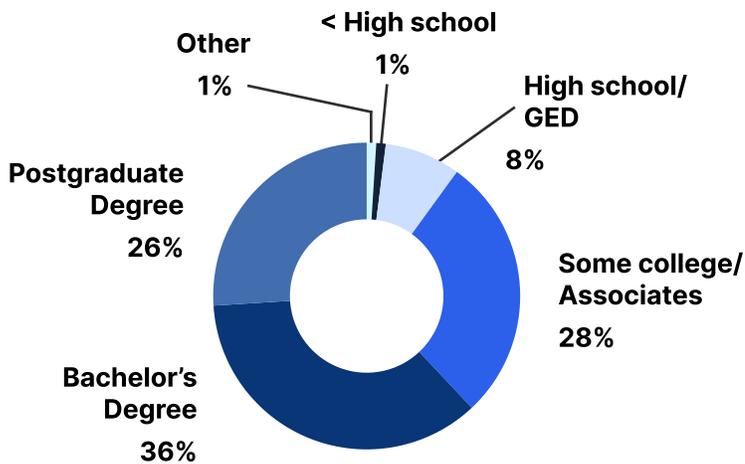


Figure 13. Respondent educational attainment of non-rejectors (US patient sample n=275, June 2025).



Recommendations: Designing for Trust, Building for Adoption

This research reveals a central truth: patient acceptance of AI hinges not just on what the technology does, but on how it feels, what it signals, and whether it aligns with deeply held moral intuitions. The 38.5% who reject AI entirely and the 61.5% who remain hesitant share substantial concerns — they want control, transparency, expert oversight, and tangible benefits. Successful implementation requires systems designed to address these human needs from the outset.

Patient acceptance of AI hinges not just on what the technology does, but on how it makes them feel.



Product and process recommendations:

AI-powered healthcare tools should be designed from the outset with two priorities: ethical alignment and patient agency. Specifically, systems should:

- Empower patient control by offering broad and granular options to decline, delete, or disable the use of personal health data, especially for secondary uses like marketing.
- Ensure meaningful consent by securing HIPAA-compliant, scenario-specific agreements before using patient data to train AI models.
- Deliver tangible benefits that patients can immediately recognize, such as clearer visit summaries, early risk alerts, or personalized treatment guidance.
- Make AI visible, not invisible by clearly labeling AI-generated content with consistent standards and clear indicators when AI assistance influenced a recommendation or summary.

Design must also support the emotional arc of adoption:

- Use gradual onboarding to help hesitant users build confidence and agency over time, especially within patient portals, chat interfaces, and digital care tools.
- Never use misleading signals such as labeling AI output as if it came from a human provider.
- Proactively surface settings and preferences, guiding patients to configuration options without burying opt-outs in legalese or dark patterns.

- Never use negative-option consent models where silence or inaction is treated as agreement.
- Design graceful error handling with user-friendly messaging and clear recovery paths that assume low baseline trust and high privacy sensitivity.

Implementation Priority: Start with the foundations that matter most to patients (e.g., control over data use, clear expert oversight, and visible benefits) before advancing to more complex AI applications. Begin with low-stakes applications (appointment scheduling, basic health information) before moving to clinical decision support.

Start with the foundations that matter most to patients before advancing to more complex AI applications.



Communication Recommendations

Effective communication about healthcare AI must do more than explain functionality. It must resonate with the moral values patients bring into the exam room or app experience.

- Lead with benefits, reinforce with values. Moral judgments begin with cost-benefit assessments, but deepen around commitments to liberty, loyalty, and legitimate expertise.
- Speak to multiple moral foundations, not just care and safety, but also autonomy, fairness, authority, and community. This is essential across onboarding flows, UX writing, call-center scripts, marketing messages, and face-to-face conversations.
- Invest in patient-specific education and training, not just provider enablement. Patients' general exposure to AI shapes their expectations and their anxiety when encountering it in healthcare. Consider personal follow-up for first-time delivery of AI-assisted content.
- Reframe the AI conversation away from replacement fears. Instead, position AI as a tool that supports clinicians, enhances quality, and solves real human problems.
- Test messages across different segments, ensuring various confidence levels, age groups, and health literacy levels feel respected and represented.
- Use human stories (i.e., real patient testimonials and relatable scenarios) to demystify AI and create emotional connection.
- Proactively address fears and misconceptions with plainspoken, culturally relevant content that explains how AI works and what it can (and cannot) do.

- Offer ongoing literacy resources like short tutorials, embedded tooltips, or patient-facing FAQs to support understanding as use evolves.
- Track trust metrics, not just adoption rates: measure patient-reported confidence levels and comfort with AI-assisted care.



Organizational Strategy Recommendations

Beyond product and communication design, successful AI implementation requires thoughtful organizational preparation:

- Treat the 40% rejection rate as market research, not a verdict. Develop targeted outreach and education programs specifically for AI-hesitant populations.
- Build cross-functional AI ethics teams that include patient advocates, community representatives, and diverse stakeholders, not just technologists and compliance officers.
- Establish clear governance frameworks for AI deployment with patient feedback loops built in and regular reassessment of trust metrics.
- Create phased rollout plans that allow for course correction based on real patient responses and emerging concerns.



Marketing Strategy Recommendations

Trust starts with the first message. Our research reveals that patients want to hear four key assurances, and marketing communications should reflect this framework:

Lead with the core message: “You are in control. Your relationship with your provider is protected. Medical experts are involved. And you’ll benefit meaningfully from our AI innovations.”

- Emphasize compliance with established medical standards while clearly explaining that AI operates under human oversight, not as an autonomous medical advisor.
- Position AI as enhancement, not replacement: a supportive tool that amplifies clinician expertise rather than substituting for human judgment.
- Address moral foundations beyond just safety and efficacy: speak to autonomy, fairness, and transparency to create deeper resonance with patient values.



Customer Service Recommendations

AI will raise questions, and patients need human answers that respect their autonomy. Organizations should prepare for the confidence gap our research identified by:

- Train support teams specifically on the concerns of low-confidence AI users, who represent the majority of patients open to healthcare AI.
- Equip staff with clear, respectful scripts and comprehensive FAQs addressing the top triggers our research identified: data privacy, marketing use, unauthorized sharing, and expert validation.
- Empower representatives to guide patients through choices about how AI fits into their care, recognizing that patient control is the strongest predictor of AI acceptance.

Customer service becomes a frontline for trust-building.

Every conversation either reinforces or undermines the ethical foundation your organization has built.



Public Relations Recommendations

Public trust is fragile, and transparency remains the best defense.

Our research shows patients are most concerned about data marketing and unauthorized sharing, making proactive communication essential:

- Communicate proactively about AI system roles, limits, and oversight before concerns arise in public forums.
- Label all AI-generated content clearly and consistently to avoid confusion or misinterpretation.
- Prepare crisis communication plans specifically focused on the scenarios our research identified as most triggering: ethical breaches, data mishandling, or perceived harm to patient autonomy.
- Respond quickly and sincerely to concerns. The difference between a temporary stumble and lasting reputational damage often lies in the speed and authenticity of organizational response.



Long-Term Strategic Value: Benefits of Ethical AI Implementation

The research findings point to a clear business case: ethical AI implementation is a moral imperative and a strategic advantage. Organizations that address the specific concerns identified in our study (e.g., patient control, expert oversight, transparency, and tangible benefits) will be better positioned to capture the 60% of patients who remain open to AI while potentially converting some of the 40% who currently reject it.

As generative AI becomes more deeply embedded in healthcare, organizations that center patient trust and moral design from the outset will thrive, adapt, and lead. The benefits extend well beyond compliance or reputation management and shape long-term performance, resilience, and market position.

Financial Performance

Building AI systems with patient trust and transparency at their core reduces long-term financial risk. Companies that invest in ethical design early avoid the steep costs of retroactive compliance, legal exposure, or reputational repair. Organizations with high patient trust scores typically see 20–25% higher treatment adherence rates and 15% lower customer acquisition costs, as satisfied patients become advocates rather than skeptics.

Ethical implementation also reduces operational friction, lowering call volume, minimizing patient confusion, and decreasing support costs. When patients understand and trust AI systems, they require less support and engage more meaningfully with their care, directly impacting the bottom line through improved efficiency and outcomes.

Trust compounds over time: human-centered AI systems foster deeper user engagement and higher customer lifetime value.

Regulatory Resilience

AI systems that respect patient autonomy, align with HIPAA, and anticipate future regulatory trends are both compliant and sustainable. Our research shows that patients specifically value HIPAA compliance and expert oversight, suggesting that systems built around these principles will face fewer regulatory challenges as the landscape evolves.

Proactive alignment with global standards like GDPR beyond the EU further strengthens a company's ability to operate across markets and adapt to changing legal requirements. Ethical AI reduces

business volatility by preventing the kind of missteps that lead to public backlash, regulatory penalties, or policy intervention that can derail entire product lines.

Market Positioning and Leadership

First-Mover Advantage: Organizations that implement patient-centered approaches now — while many competitors focus purely on technical capabilities — establish market leadership that becomes increasingly difficult to challenge as AI adoption accelerates.

Ethical, human-centered AI gives organizations the opportunity to set industry benchmarks rather than follow them. As the field matures, early adopters of transparent, patient-first practices will shape public expectations and regulatory standards. This positioning is especially valuable in healthcare, where trust, safety, and access remain under intense scrutiny.

Trusted AI expands access to personalized care, reaching underserved populations and growing the market while improving public health and community trust

Brand Value and Public Goodwill

Companies that address multiple moral foundations (e.g., autonomy, fairness, and transparency) create deeper emotional connections with patients. This multi-

dimensional trust translates into stronger brand equity, higher patient loyalty, and a more resilient market position.

Our research demonstrates that patients don't evaluate AI systems solely on technical merit; they judge them through emotional, ethical, and social lenses. Organizations that understand and respond to these deeper human needs are positioned as trustworthy healthcare partners, not just technology vendors.

Over time, companies that integrate ethics into AI design and delivery become the standard-setters for an entire industry moving toward more human-centered care.

The Competitive Moat Strategy

By investing in these organizational enablers, healthcare companies create an ecosystem that transforms the challenge of patient hesitancy into a competitive advantage. Organizations that earn patient trust through ethical AI implementation will define the healthcare transformation, setting standards that elevate an entire industry toward more human-centered care.

The result is not just better technology adoption, but a stronger, more trusted, and more sustainable healthcare future. By committing to openness, speaking to moral values, and embedding patient priorities into AI design, healthcare organizations can transform patient hesitancy into trust. This

approach delivers more than just compliance or acceptance. It creates deeper patient engagement, stronger therapeutic relationships, and ultimately better health outcomes.

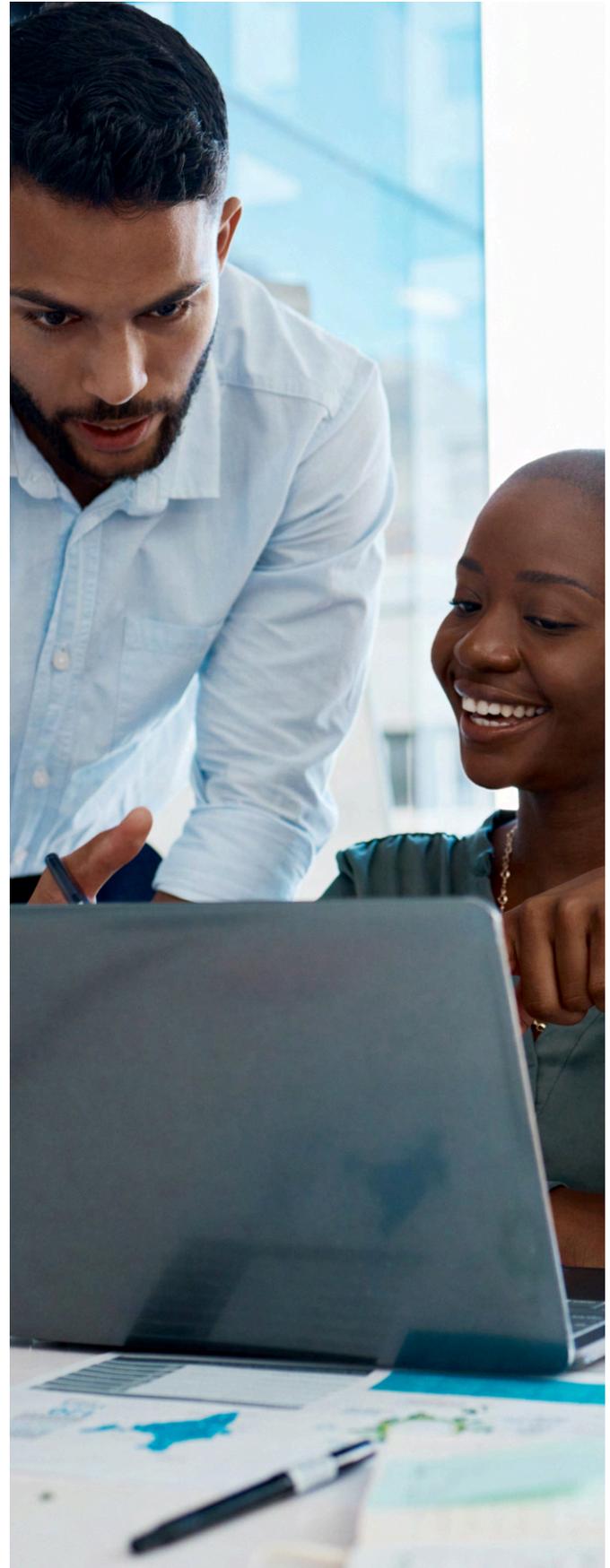
The organizations that get this right won't only survive the AI transformation, they'll lead it.

Ethical implementation is not a constraint on innovation, but a competitive advantage that builds lasting patient loyalty and clinical excellence.

Implementation Sequencing for Maximum Impact

Begin with foundation-building: Start with low-risk, high-visibility applications that demonstrate the four key patient values our research identified. Use early wins in areas like appointment scheduling or health education to build organizational confidence and patient trust before advancing to complex clinical applications.

Scale Strategically: Once foundational trust is established, expand to higher-stakes applications with the credibility and user confidence needed for successful adoption.



Conclusion: Human-Centered AI Is the Way Forward

AI is already embedded throughout healthcare, powering insurance algorithms, triaging emergency calls, and suggesting treatments in EHR systems. Yet most patients don't realize how extensively AI shapes their care, and when they do become aware, 40% reject it entirely.

The future of healthcare AI won't be decided by technical capabilities alone, but by whether patients trust, accept, and engage with these systems.

This research reveals that patient resistance to healthcare AI follows predictable moral patterns. Through our analysis of 275 patients across 33 scenarios, we've identified the specific triggers that drive rejection (e.g., data marketing, lack of control, absence of expert oversight) and the mitigators that build support. Most importantly, we've shown that even AI-hesitant patients share common ground about what makes AI acceptable: transparency, control, expert validation, and tangible benefits.

Healthcare organizations face a critical choice. They can continue deploying AI systems designed primarily for operational efficiency, risking widespread patient backlash and regulatory intervention. Or,

they can learn from these insights to build AI that patients actually want: systems that respect autonomy, ensure expert oversight, and deliver clear value that improves care experiences.

Healthcare organizations that embrace these insights will avoid the pitfalls of patient rejection and unlock AI's true potential. They'll create systems that patients trust enough to engage with fully, providers feel confident recommending, and regulators view as gold standards for responsible innovation. These organizations will discover that ethical AI implementation is the foundation for sustainable competitive advantage in an industry built on trust.

By embedding moral intelligence into AI strategy from the start, healthcare leaders can transform inevitable patient skepticism into lasting market differentiation. The organizations that act on these insights now while the field is still taking shape will define the standards that others must follow.

The question isn't whether AI will reshape healthcare, it is whether patients will allow it into their healthcare journey. Leaders who implement a human-centered thinking in their AI systems are poised to define the future.

Authors' Note:

1. Richa and Lanyi contributed equally to this paper. The foundational research paper written by them won the 2025 UW Communication Faculty Award for Outstanding Research.
2. The opinions expressed in this paper are those of the contributors and do not necessarily reflect any official position or viewpoint of their employers. The authors recommend that users always make sure AI-generated content is accurate and appropriate before using.
3. For their contributions to pilot tests of this study, the authors wish to thank all of the students in the Persuading Ethical UX Design class taught by David Evans Ph.D. in the Master of Communication Leadership program at the University of Washington. The authors would like to especially thank classmates Dany Rios, Valentina Wang, Vivian Bian for their invaluable support while developing the initial study.
4. The authors extend their gratitude to Caroline Ma, UX Researcher, for her meaningful contributions to the development of the survey questions used in this study. Her thoughtful input helped shape a tool that captured diverse perspectives clearly and effectively.
5. The authors also thank Alli Bradshaw for the design of this white paper and for making the research more accessible through clear, user-friendly visuals.
6. The UW Human Subjects Division approved this study on May 7, 2025 (IRB ID: STUDY00023024).

About the Authors:

Richa Nevatia (UW) is a marketing and product strategist specializing in the intersection of technology, business and marketing with over 4 years of professional experience. She earned a Master's in Communication Leadership from the University of Washington, where she was awarded the 2025 Master's in Communication Research Award (with Lanyi Zhu) for their foundational paper on AI in healthcare. Richa also holds an MBA from MICA, and a Bachelors degree in finance from the University of Mumbai, bringing a global, multidisciplinary lens to her work at the nexus of innovation and impact. Connect with Richa on LinkedIn: [linkedin.com/in/nevatiaricha](https://www.linkedin.com/in/nevatiaricha)

Lanyi Zhu (UW) is a mixed methods UX Researcher with 2+ years of industry experience, specializing in agentic AI, responsible AI, and AI ethics. She has PhD-level training in social science and statistics, and holds master's degrees of Communication in Digital Media and Political Science, along with graduate certificate in User-Centered Design (UCD) — all from the University of Washington. Connect with Lanyi on LinkedIn for future research collaborations: <https://www.linkedin.com/in/lanyi-zhu/>

Emma Wigdahl (Intechnic) is a user experience leader passionate about the intersection of digital design and health outcomes. As the UX Team Lead at Intechnic, she directs a research team specializing in creating better digital experiences for patients and healthcare providers. Leveraging her background in Psychology from the University of New Mexico, Emma utilizes user interviews and usability testing to design interfaces to build empathetic digital tools that improve the entire healthcare experience for patients and providers alike. Connect with Emma on LinkedIn: www.linkedin.com/in/emmawigdahl

Alla Zarifyan (Intechnic) is a healthcare communications strategist with a background in molecular biology and pharmacology. With 15+ years of experience translating complex scientific topics into clear, accessible narratives, she has led communications across biotech, healthtech, and AI-driven healthcare organizations, bridging scientific insight and strategic storytelling to engage both professionals and the general public. Connect with Alla on LinkedIn: [linkedin.com/in/alla-zarifyan](https://www.linkedin.com/in/alla-zarifyan)

Andrew Kucheriavy (Intechnic) is a technology strategist and founder with over two decades of experience designing digital systems for global organizations. His current work focuses on integrating AI, behavioral science, and clinical insight to improve how people engage with healthcare, decision-making, and complex information. At Intechnic, he leads initiatives at the intersection of intelligent systems and human-centered design, developing frameworks that support more adaptive, trustworthy, and impactful experiences. Connect with Andrew on LinkedIn: [linkedin.com/in/uxmaster](https://www.linkedin.com/in/uxmaster)