Translational Medicine and Digital Health
REVIEW | COMMISSIONED SYNTHESIS

Patient-Facing Generative AI in Telemedicine

Patrick Sullivan - Corresponding author

Abstract

This review examines patient-facing generative artificial intelligence in telemedicine. The organizing question is how generated explanations and triage suggestions can be evaluated for comprehension, equity, safety, and escalation. Ten related scholarly sources are synthesized through a decision-centered framework spanning problem definition, mechanism, measurement, evaluation, implementation, and governance. The review does not invent experiments, pooled estimates, or unreported quantitative results. It instead evaluates the strength and transferability of the available evidence, with particular attention to allowing conversational fluency to mask clinical uncertainty and delayed care. The resulting framework links technical or empirical performance to explicit use conditions and identifies tests that should precede wider adoption in teleconsultation and patient communication.

Keywords: generative AI, telemedicine, patient communication, safety, health equity

Synthesis lens	Principal question
Mechanism	the chain connecting evidence inputs, patient-facing generative artificial intelligence in telemedicine, intermediate outputs, and downstream decisions
Evaluation	construct validity, comparative performance, uncertainty, transfer across settings, and decision utility
Primary risk	allowing conversational fluency to mask clinical uncertainty and delayed care

1. Introduction

Research on patient-facing generative artificial intelligence in telemedicine sits at the boundary between a promising component and a consequential decision. The core question is how generated explanations and triage suggestions can be evaluated for comprehension, equity, safety, and escalation. Answering it requires more than assembling favorable results. It requires a chain of evidence that makes the problem boundary, mechanism, measurement, comparator, uncertainty, and accountable decision visible to the reader.

This article presents a structured narrative review of ten related publications. Sources were selected for topical relevance, traceable publication metadata, and complementary coverage of technical, empirical, and governance questions. No meta-analytic pooling is attempted because the included studies differ in designs, populations, system boundaries, and outcomes. The purpose is decision-oriented synthesis: to identify what each source can support, where transfer remains uncertain, and which evidence would justify the next stage of use.

The central risk is allowing conversational fluency to mask clinical uncertainty and delayed care. A top-line metric or broad policy label can appear decisive while concealing the conditions that produced it. Accordingly, the synthesis separates observation from interpretation and implementation from validation. This separation is especially important when the same system creates benefits for one user or institution while transferring cost, risk, or administrative burden to another.

2. Evidence Base and Problem Boundary

A defensible review begins by defining the unit of analysis. For the present topic, the relevant boundary includes inputs, institutional or technical context, the mechanism producing an output, the person or system acting on that output, and the downstream outcome. Evidence falls outside the boundary when it measures only an intermediate proxy yet is used to claim an end-to-end benefit.

Sezgin and McKay (2024) addresses "Behavioral health and generative AI: a perspective on future of therapies and patient care." In this synthesis, that work is used as a source on problem definition within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Mandal et al. (2025) addresses "Utilization of Generative AI-drafted Responses for Managing Patient-Provider Communication." In this synthesis, that work is used as a source on conceptual boundary within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Baig et al. (2024) addresses "Generative AI in Improving Personalized Patient Care Plans: Opportunities and Barriers Towards Its Wider Adoption." In this synthesis, that work is used as a source on measurement within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Together, these works provide complementary entry points, but their conclusions should not be flattened into a single ranking. Differences in data generation, setting, temporal horizon, and outcome definition may be substantively meaningful. A review should therefore preserve those differences and state where analogy, rather than direct replication, connects one domain to another.

3. Mechanism and System Design

The operative mechanism is the chain connecting evidence inputs, patient-facing generative artificial intelligence in telemedicine, intermediate outputs, and downstream decisions. This formulation discourages component-only reasoning. A model, platform, policy, assay, or infrastructure service acquires practical meaning only when its interfaces and use conditions are specified. The evidence chain should describe what information is available, what transformation occurs, which assumptions make that transformation credible, and who is authorized to act.

Uddin et al. (2025) addresses "Health Communication on the Internet: Promoting Public Health and Exploring Disparities in the Generative AI Era." In this synthesis, that work is used as a source on system mechanism within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Vaccari et al. (2021) addresses "A Generative Adversarial Network (GAN) Technique for Internet of Medical Things Data." In this synthesis, that work is used as a source on representation choice within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Sun and Ortiz (2024) addresses "Rapid Review of Generative AI in Smart Medical Applications." In this synthesis, that work is used as a source on failure analysis within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

These studies also show why modular claims require interface tests. A component may perform well while the complete pathway fails because inputs drift, users interpret outputs differently, recovery semantics are ambiguous, or institutional incentives change. Design documentation should therefore include not only the intended pathway but also foreseeable deviations, degraded modes, and conditions under which the system should stop or defer.

4. Evaluation and Uncertainty

Evaluation should center on construct validity, comparative performance, uncertainty, transfer across settings, and decision utility. Average performance remains useful, but it should be accompanied by distributions, error categories, relevant subgroups or operating regimes, and uncertainty at the point where a decision changes. Comparators must isolate the value of the proposed addition rather than compare a mature intervention with an intentionally weak baseline.

Meskó (2023) addresses "The Impact of Multimodal Large Language Models on Health Care’s Future." In this synthesis, that work is used

as a source on comparative evaluation within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Xu et al. (2021) addresses "Chatbot for Health Care and Oncology Applications Using Artificial Intelligence and Machine Learning: Systematic Review." In this synthesis, that work is used as a source on uncertainty within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

External validation should introduce meaningful shifts in setting, time, population, workload, market structure, or environmental conditions. A nominally independent sample can still be too similar to test transfer. Conversely, a failed transfer is scientifically informative when the changed conditions are measured and the mechanism of failure is examined rather than hidden in an aggregate score.

5. Translation and Governance

Translation to teleconsultation and patient communication depends on more than technical readiness. Decision rights, documentation, maintenance, monitoring, appeal, and recovery are part of the intervention. The governance requirement is traceable evidence, named responsibility, version control, monitoring, appeal, and explicit revalidation. Each control should be connected to a named failure mode and should identify an owner, evidence source, review interval, and escalation path.

Yang et al. (2025) addresses "Peer perceptions of clinicians using generative AI in medical decision-making." In this synthesis, that work is used as a source on implementation within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Babu and Boddu (2024) addresses "BERT-Based Medical Chatbot: Enhancing Healthcare Communication through Natural Language Understanding." In this synthesis, that work is used as a source on governance within patient-facing generative artificial intelligence in telemedicine. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for teleconsultation and patient communication. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

A credible implementation plan also defines sunset and revalidation conditions. New data, software updates, institutional restructuring, population change, or shifts in incentives can invalidate previously reasonable assumptions. Monitoring should therefore test the validity of the evidence chain, not merely whether a service remains online or a headline metric remains stable.

6. Research Agenda

Future studies should preregister or otherwise predefine the intended decision, principal outcomes, exclusions, and analysis plan when feasible. They should report missingness, sensitivity analyses, negative or null findings, and the cost of errors to differently situated stakeholders. Reproducible artifacts should include sufficient metadata to reconstruct data provenance, model or analytical versions, parameter choices, and the sequence from evidence to conclusion.

Cross-disciplinary work needs a shared object of explanation rather than a collection of adjacent vocabularies. For patient-facing generative artificial intelligence in telemedicine, that shared object is the complete decision pathway. Technical, clinical, social, environmental, or economic expertise should each change the design or interpretation of the study. When evidence remains incomplete, the useful output is a bounded hypothesis, a transparent uncertainty statement, or a request for additional measurement.

7. Conclusion

The literature supports a disciplined approach to patient-facing generative artificial intelligence in telemedicine: define the decision, preserve system boundaries, test consequential failure modes, connect uncertainty to action, and assign accountable control. The strongest conclusion is not that one method is universally superior. It is that credible use in teleconsultation and patient communication requires an auditable link from source evidence to design choice, evaluation result, and governed decision.

Declarations

Ethics approval: Not applicable. This article synthesizes previously published literature and reports no new research involving humans, animals, human tissue, or identifiable sensitive data.

Consent to participate and consent for publication: Not applicable.

Clinical trial registration: Not applicable; no intervention or participant enrollment was conducted.

Reporting guidance: This narrative synthesis is not a clinical trial, primary observational study, prediction-model development study, or case report. CONSORT, STROBE, TRIPOD, and CARE are therefore not directly applicable. Any later conversion to a systematic review or primary clinical study requires the corresponding reporting checklist and protocol or registration.

Funding: No external funding is reported for this commissioned synthesis.

Competing interests: The authoring group declares no competing interests for this commissioned synthesis.

AI-assisted writing: Generative AI supported drafting, source organization, and language editing. Accountable authors and editors must verify every claim, citation, and disclosure before publication.

Data, code, and materials availability: No new dataset or code was generated. All evidence is identified in the reference list.

References

Babu, A., & Boddu, S. B. (2024). BERT-Based Medical Chatbot: Enhancing Healthcare Communication through Natural Language Understanding. Exploratory Research in Clinical and Social Pharmacy, 13, 100419. https://doi.org/10.1016/j.rcsop.2024.100419

Baig, M. M., Hobson, C., GholamHosseini, H., Ullah, E., & Afifi, S. (2024). Generative AI in Improving Personalized Patient Care Plans: Opportunities and Barriers Towards Its Wider Adoption. Applied Sciences, 14(23), 10899.

https://doi.org/10.3390/app142310899

Mandal, S., Wiesenfeld, B. M., Szerencsy, A. C., Small, W. R., Major, V., Richardson, S., Schoenthaler, A., Mann, D., & Nov, O. (2025). Utilization of Generative AI-drafted Responses for Managing Patient-Provider Communication. npj Digital Medicine, 8(1), 591. https://doi.org/10.1038/s41746-025-01972-w

Meskó, B. (2023). The Impact of Multimodal Large Language Models on Health Care's Future. Journal of Medical Internet Research, 25, e52865. https://doi.org/10.2196/52865

Sezgin, E., & McKay, I. (2024). Behavioral health and generative AI: a perspective on future of therapies and patient care. npj Mental Health Research, 3(1), 25. https://doi.org/10.1038/s44184-024-00067-w

Sun, Y., & Ortiz, J. (2024). Rapid Review of Generative AI in Smart Medical Applications. International Journal of Computer Science and Information Technology, 3(2), 80-93. https://doi.org/10.62051/ijcsit.v3n2.10

Uddin, J., Feng, C., & Xu, J. (2025). Health Communication on the Internet: Promoting Public Health and Exploring Disparities in the Generative AI Era. Journal of Medical Internet Research, 27, e66032. https://doi.org/10.2196/66032

Vaccari, I., Orani, V., Paglialonga, A., Cambiaso, E., & Mongelli, M. (2021). A Generative Adversarial Network (GAN) Technique for Internet of Medical Things Data. Sensors, 21(11), 3726. https://doi.org/10.3390/s21113726

Xu, L., Sanders, L., Li, K., & Chow, J. C. L. (2021). Chatbot for Health Care and Oncology Applications Using Artificial Intelligence and Machine Learning: Systematic Review. JMIR Cancer, 7(4), e27850. https://doi.org/10.2196/27850

Yang, H., Dai, T., Mathioudakis, N., Knight, A. M., Nakayasu, Y., & Wolf, R. M. (2025). Peer perceptions of clinicians using generative AI in medical decision-making. npj Digital Medicine, 8(1), 530. https://doi.org/10.1038/s41746-025-01901-x