



Explainable Generative AI for Personalized Healthcare Decision Support: A Comprehensive Survey

Amit Kumar¹, Dr. Pawan Meena², Yogesh Rai³

¹M.Tech Scholar Technocrats Institute of Technology Excellence

^{2,3}Asst. Professor, CSE-AIML, Technocrats Institute of Technology Excellence

Abstract: Generative Artificial Intelligence (GenAI), particularly large language models (LLMs) and multimodal foundation models, is rapidly transforming the landscape of healthcare information systems. Unlike conventional machine learning systems that primarily focus on prediction and classification, generative models can interact with users through natural language, synthesize information, summarize clinical records, and produce personalized explanations. These capabilities create opportunities for developing personalized healthcare decision-support systems that consider patient characteristics, medical history, preferences, symptoms, and contextual information. However, the increasing personalization of generative AI introduces important challenges involving explainability, hallucination, privacy, fairness, bias, uncertainty, accountability, and clinical safety. A personalized recommendation without a transparent justification may be difficult for clinicians and patients to trust or evaluate.

This paper presents a comprehensive survey of explainable generative AI for personalized healthcare decision support. It examines the evolution from conventional clinical decision-support systems and machine learning toward explainable AI and generative AI. The paper reviews personalization mechanisms based on patient profiles, longitudinal electronic health records, clinical context, preferences, and multimodal information. It analyzes explanation techniques including feature attribution, natural-language explanations, counterfactual explanations, example-based explanations, evidence-grounded explanations, and uncertainty-aware explanations. The paper also investigates applications in disease-risk assessment, personalized treatment support, medication management, patient education, preventive healthcare, mental-health support, chronic disease management, and clinical workflow assistance. A conceptual Explainable Personalized Generative Healthcare Decision Support Framework is proposed, integrating patient profiling, predictive analytics, evidence retrieval, personalization, explanation generation, explanation verification, uncertainty

estimation, and human oversight. Finally, major research gaps and future directions are identified, including longitudinal personalization, causal explainability, multimodal personalization, multilingual healthcare AI, privacy-preserving personalization, fairness-aware recommendations, prospective clinical validation, and standardized evaluation. The survey concludes that explainable GenAI has significant potential to support personalized healthcare, but safe deployment requires evidence grounding, transparent explanations, continuous monitoring, privacy protection, and clinician involvement.

Index Terms— Generative AI, Explainable AI, Personalized Healthcare, Large Language Models, Clinical Decision Support, Patient-Centered AI, Healthcare Recommendation Systems, Medical AI, Personalization, XAI.

I. INTRODUCTION

Healthcare is increasingly moving toward personalized and patient-centered care. Traditional healthcare approaches frequently apply population-level clinical guidelines, while modern digital health systems increasingly seek to incorporate patient-specific characteristics into decision-making. Personalized healthcare attempts to answer a more specific question: **What information or decision-support recommendation is most appropriate for this particular patient?**[1]

Patient characteristics may include:

- Age
- Sex
- Medical history
- Family history
- Laboratory measurements
- Medication history
- Lifestyle
- Symptoms
- Previous diagnoses
- Preferences

- Treatment response
- Longitudinal health trends

The availability of electronic health records, wearable devices, mobile health applications, telemedicine platforms, and patient-generated health data has created unprecedented opportunities for personalization.[2][3] Machine learning has traditionally been used to identify disease patterns and estimate risks from these data. However, conventional predictive systems generally produce numerical or categorical outputs. They may answer:

Risk=0.78

but do not necessarily explain the prediction in a manner that is understandable to a patient or clinician.

Explainable Artificial Intelligence (XAI) was developed partly to address this problem.

An explainable system attempts to answer:

Why did the system produce this result?

Generative AI introduces another dimension.

An LLM can potentially answer:

What does this result mean for this particular patient, and how can it be communicated clearly?

This creates a promising combination:

Prediction+Personalization+Explanation+Natural Language

However, personalization also increases risk.

A generic incorrect answer may be inconvenient.

A personalized incorrect healthcare recommendation may be dangerous.

Consequently, explainability becomes particularly important in personalized GenAI.

The World Health Organization has emphasized that AI in healthcare must address safety, transparency, accountability, equity, and human autonomy. Its guidance on large multimodal models specifically recognizes both the opportunities and risks of generative AI in health.[4][11]

Recent reviews also demonstrate rapid growth in healthcare LLM research, including clinical decision-making, patient information, education, summarization, and documentation, while identifying continuing challenges involving incorrectness, bias, safety, and reproducibility.

This paper therefore investigates the convergence of:

Generative AI+Explainable AI+Personalization

for healthcare decision support.

II. OBJECTIVES OF THE SURVEY

The major objectives are:

1. To examine the evolution of personalized healthcare AI.
2. To review conventional personalized clinical decision-support systems.
3. To analyze XAI techniques applicable to personalized healthcare.
4. To examine GenAI and LLM applications in healthcare.
5. To investigate personalization mechanisms for GenAI.
6. To analyze explainable generative healthcare systems.
7. To identify risks associated with personalized GenAI.
8. To propose a conceptual explainable personalized GenAI framework.
9. To identify research gaps.
10. To recommend future research directions.[5]

III. PERSONALIZED HEALTHCARE

A. Definition

Personalized healthcare refers to healthcare decision-making that incorporates individual patient characteristics rather than relying exclusively on population-level averages.[12]

Conceptually:

Decision=f(Patient Characteristics,Clinical Evidence,Context)

Patient characteristics can include:

$P=\{D,H,L,M,S,W\}$

Where:

D = demographic information

H = health history

L = laboratory information

M = medication history

S = symptoms

W = wearable/behavioral information

B. Personalized Versus Generalized Healthcare AI

Characteristic	General AI	Personalized AI
Input	General question	Patient-specific information
Context	Limited	Extensive
Recommendation	Population-oriented	Individual-oriented
Data requirement	Moderate	High
Privacy requirement	Moderate	Very high
Risk	General	Potentially higher
Explainability	Important	Critical

IV. EVOLUTION OF PERSONALIZED HEALTHCARE AI

Healthcare decision support has evolved through several stages.

Stage 1: Rule-Based Systems

Early clinical decision-support systems relied on explicit rules.

Example:

IF glucose > threshold → High Risk

Stage 2: Statistical Models

Regression-based methods introduced probabilistic predictions.

Examples include:

- Logistic regression
- Cox regression
- Bayesian models

Stage 3: Machine Learning

ML introduced more complex pattern recognition.

Examples:

- Random Forest
- SVM
- XGBoost
- Neural Networks

Stage 4: Explainable AI

XAI introduced mechanisms for interpreting complex models.

Examples:

- SHAP
- LIME

- Integrated Gradients
- Counterfactual explanations

Stage 5: Generative AI

LLMs introduced natural-language interaction.

Stage 6: Explainable Personalized GenAI

The emerging paradigm combines:

Patient Context+Prediction+Evidence+Explanation+Generation

This represents the primary focus of this survey.

V. GENERATIVE AI IN HEALTHCARE

Generative artificial intelligence (GenAI), particularly large language models (LLMs), has emerged as an important technology for healthcare information processing, communication, and decision-support applications. Unlike conventional predictive models that primarily generate numerical or categorical outputs, GenAI can transform heterogeneous healthcare information into structured and natural-language responses. Potential applications include clinical summarization, patient education, clinical question answering, documentation assistance, and evidence synthesis.[6][9]

A. Clinical Summarization: LLM-based systems can summarize heterogeneous clinical information, including patient histories, clinical notes, discharge summaries, laboratory reports, and medication records. Such summarization can reduce information overload and assist clinicians in rapidly identifying relevant patient information. However, generated summaries require verification because omission or distortion of clinically significant information can adversely affect subsequent decisions.[7]

B. Patient Education: GenAI can transform complex medical terminology into patient-oriented explanations. This capability can improve health-information accessibility by adapting the complexity, vocabulary, and structure of explanations to individual users. Nevertheless, generated educational content must remain consistent with validated medical evidence.[8]

C. Clinical Question Answering: LLMs can assist healthcare professionals and patients in retrieving, organizing, and communicating healthcare information. Retrieval-augmented generation (RAG) can further improve this capability by grounding generated responses in authoritative clinical resources rather than relying exclusively on information encoded in model parameters.[13]

D. Decision Support: GenAI can synthesize information from multiple sources, including electronic health records (EHRs), clinical guidelines, medical literature, and predictive-model outputs. This synthesis can support clinical reasoning; however, the generated response should function as decision support rather than an autonomous clinical decision mechanism.[15]

E. Clinical Documentation: LLMs can assist in generating structured clinical documentation from notes, conversations, and other healthcare information. Potential benefits include reduced documentation burden and improved information organization. Such systems require safeguards against hallucination, inaccurate transcription, privacy violations, and inappropriate insertion of unsupported clinical information.

Consequently, although GenAI presents considerable opportunities for healthcare transformation, its integration into clinical workflows must be accompanied by evidence grounding, uncertainty communication, privacy protection, and human oversight.

VI. EXPLAINABLE AI IN PERSONALIZED HEALTHCARE

Explainability is particularly important when artificial intelligence systems generate individualized predictions or recommendations. Consider a predictive model that estimates an individual's disease risk as: **Risk = 0.85**

A numerical prediction alone does not explain why the estimated risk is high. An explainable AI (XAI) component can identify influential patient characteristics such as family history, blood glucose, body mass index (BMI), blood pressure, and age.

The explanation can conceptually be represented as:

$$E(x) = \{(F_i, C_i)\}_{i=1}^m$$

Such explanations can improve transparency and support more meaningful interaction between patients, clinicians, and AI systems. However, feature attribution should not automatically be interpreted as causal evidence. Therefore, explanation methods must be evaluated for both faithfulness and clinical relevance.

VII. TYPES OF EXPLANATIONS

A. Feature-Based Explanations: Feature-based explanations identify the variables that contribute most strongly to a prediction. For example, an explanation may indicate that elevated glucose and BMI were among the dominant contributors to an estimated disease risk.[14][16]

B. Counterfactual Explanations: Counterfactual explanations describe hypothetical changes that could produce a different model output. Conceptually,

$$x' = \arg \min_z D(x, z)$$

Where (x) represents the observed patient state and (x') represents a hypothetical alternative state associated with a different prediction.

Such explanations can potentially make AI outputs more actionable. However, clinically meaningful counterfactuals must distinguish between variables that are realistically modifiable and variables that are immutable.

Example-based approaches explain predictions by identifying historically similar cases. A simplified representation is:

$$Patient\ x \rightarrow TopK_{z \in D} Sim(x, z) \rightarrow Similar\ Cases \rightarrow Explanation$$

This approach can provide contextual information but requires careful similarity definitions to avoid misleading comparisons.

D. Natural-Language Explanations: GenAI can transform technical XAI outputs into understandable natural-language explanations. This capability enables the same underlying evidence to be communicated differently to clinicians and patients.[15][18]

E. Evidence-Based Explanations: Evidence-based explanations associate model outputs with supporting information obtained from clinical guidelines, peer-reviewed literature, and authorized patient records. Evidence grounding can improve the traceability of generated responses.

F. Uncertainty Explanations: Healthcare AI systems should explicitly communicate uncertainty. Rather than presenting predictions as definitive diagnoses, systems should distinguish

between model estimates, evidence-supported conclusions, and clinically established diagnoses.

VIII. PERSONALIZATION MECHANISMS

Personalization can operate at multiple levels within healthcare AI systems.

A. Patient Profile Personalization: The system can incorporate demographic and clinical attributes such as age, sex, medical history, diagnoses, and current health conditions.

B. Context Personalization: The relevance of information depends on the current clinical context. Preventive care, emergency assessment, chronic disease monitoring, and post-treatment follow-up may require different information and communication strategies.

C. Preference Personalization: Patient preferences can influence communication and decision-support presentation. Examples include preferred language, explanation complexity, communication style, and treatment preferences.[18]

D. Temporal Personalization: Personalized healthcare systems can incorporate longitudinal information rather than relying exclusively on current observations. Thus, Temporal personalization can reveal trends in blood pressure, glucose levels, medication response, body weight, and recurring symptoms.

IX. LONGITUDINAL PATIENT MODELING

Longitudinal modeling represents a major opportunity for personalized GenAI. A patient's evolving health state can be represented as

$$P_{1:t} = \{X_1, X_2, \dots, X_t\}$$

Where (X_t) represents the patient's available health information at time (t).

Longitudinal modeling enables AI systems to analyze temporal patterns rather than isolated measurements. Such patterns may include progressive changes in blood pressure, fluctuations in glucose levels, medication response, weight trajectories, and recurring symptoms.

This temporal representation can provide more contextualized decision support than static patient profiles. Nevertheless, longitudinal systems require appropriate handling of missing observations, temporal inconsistencies, changes in clinical practice, and evolving patient conditions.

X. PROPOSED EXPLAINABLE PERSONALIZED GENERATIVE HEALTHCARE FRAMEWORK

This paper proposes an Explainable Personalized Generative Healthcare Decision Support (EPG-HDS) framework for integrating predictive analytics, XAI, knowledge retrieval, personalization, and GenAI within a human-supervised healthcare architecture.[20]

The proposed architecture consists of the following major stages:

Patient/Clinician → Patient Profile → Longitudinal Health Data → Context Identification → Predictive Model + Knowledge Retrieval → XAI Explanation → Personalization Engine → Generative AI → Explanation/Recommendation → Safety Verification → Clinician → Final Decision

The architecture follows a modular design in which predictive modeling, evidence retrieval, explanation generation,

personalization, and safety verification are treated as distinct but interconnected components.[19]

The primary objective is not to replace clinicians but to provide evidence-grounded, interpretable, personalized, and appropriately qualified decision-support information.

XI. PATIENT PROFILE MODULE

The patient profile module constructs a structured representation of an individual's relevant healthcare information. The profile can include four major categories.

A. Demographic Information

- Age
- Sex
- Geographic or regional information

B. Clinical Information

- Diagnoses
- Symptoms
- Laboratory measurements
- Medications
- Clinical observations

C. Behavioral Information

- Physical activity
- Dietary patterns
- Sleep characteristics
- Tobacco or other relevant behavioral factors

D. Historical Information

- Previous treatments
- Previous diagnostic results
- Hospitalizations
- Longitudinal clinical observations

The module must implement appropriate authorization and access-control mechanisms. Only information that the system and user are authorized to access should contribute to personalization.

XII. CONTEXT ENGINE

The same patient may require different decision-support information depending on the current clinical situation. Therefore, the proposed framework introduces a context engine that determines the relevance of patient information according to the current scenario.

The decision-support function can be represented as:

Recommendation = f(Patient, Context).

Relevant contexts may include preventive care, chronic disease management, acute symptoms, post-treatment monitoring, and medication review.[16]

Context identification can reduce irrelevant information and improve the appropriateness of generated responses.

XIII. PREDICTIVE ANALYTICS

A conventional machine learning model can generate patient-specific risk estimates represented by $P(\text{Disease} \mid \text{Patient Features})$.

Potential predictive algorithms include logistic regression, random forests, gradient-boosting methods such as XGBoost, neural networks, and temporal models.

The predictive component should be independently trained and validated before its outputs are incorporated into a GenAI

pipeline. Appropriate evaluation should include discrimination, calibration, sensitivity, specificity, and robustness across relevant patient populations.[9]

Separating predictive modeling from language generation also provides a clearer basis for evaluating whether the generated explanation faithfully represents the underlying prediction.

XIV. KNOWLEDGE RETRIEVAL

The knowledge-retrieval component provides evidence required for grounded generation. Potential information sources include clinical guidelines, peer-reviewed medical literature, drug databases, institutional protocols, and authorized patient records. A retrieval-augmented generation architecture can be represented as

Response = LLM(Query, Retrieved\ Evidence, Patient\ Context). Grounding the generation process in authoritative evidence can reduce reliance on unsupported model-generated information. However, retrieval itself requires validation because irrelevant, outdated, or inappropriate sources can introduce errors into the final response.

XV. PERSONALIZATION ENGINE

The personalization engine integrates patient characteristics, clinical context, and retrieved evidence:

$$Y_p = f(P, C, E) = f(P_{profile}, C_{clinical}, E_{retrieved})$$

Patients with the same diagnosis may differ considerably in age, comorbidities, medications, risk factors, lifestyle characteristics, and preferences. Consequently, a single generic response may not provide equally useful information for every patient.

The personalization engine therefore determines which information should be emphasized and how it should be communicated to the intended user.

XVI. GENERATIVE AI ENGINE

The GenAI engine converts structured model outputs, explanations, evidence, and patient context into natural-language communication.

Its input can be represented as:

$$I = \{P, E, X, C\}$$

where (P) represents the predictive output, (E) represents retrieved evidence, (X) represents XAI information, and (C) represents patient context.

The generated response can then be represented as:

$$R = \text{LLM}(I).$$

The generation component should be constrained so that unsupported medical claims are not introduced. Prompt constraints, retrieval grounding, structured output formats, and post-generation verification can contribute to this objective.

XVII. EXPLANATION VERIFICATION

Explanation verification is a critical safety layer in the proposed framework. Before an AI-generated response is presented to a clinician or patient, the system should evaluate the following dimensions:

1. **Evidence Support:** Whether generated claims are supported by retrieved evidence.
2. **Patient Consistency:** Whether the response accurately represents the available patient information.

3. **Clinical Consistency:** Whether generated statements conflict with validated medical knowledge.
4. **Uncertainty:** Whether uncertainty is appropriately communicated.
5. **Safety:** Whether the output contains potentially harmful or unsupported recommendations.

The verification layer can therefore be conceptualized as a gate between generation and clinical presentation.

XVIII. EXAMPLE OF PERSONALIZED EXPLANATION

Consider a hypothetical patient for whom a predictive model produces an estimated risk of Risk=0.78.

Suppose the major contributing variables are glucose, BMI, age, and physical activity. The resulting explanation could state:

“The predictive model estimates an elevated risk based primarily on the patient's glucose measurement and BMI, while physical activity contributes in the opposite direction. This is a model-based risk estimate rather than a diagnosis and should be considered alongside clinical assessment.”[11][15]

This example demonstrates the integration of four components: Prediction + Evidence + Explanation + Safety.

The output does not convert a model prediction into a definitive medical diagnosis.

XIX. PERSONALIZED PATIENT COMMUNICATION

GenAI enables adaptive communication by presenting identical underlying information at different levels of technical complexity. For example, a technical representation may emphasize model features, a clinician-oriented representation may describe relevant clinical variables, and a patient-oriented representation may use accessible terminology.

Thus, personalization can be applied not only to what information is communicated, but also to how that information is communicated.

This capability is particularly relevant to health literacy and patient engagement.

XX. MULTILINGUAL PERSONALIZATION

Language is an important consideration in equitable healthcare communication. Personalized GenAI systems may generate healthcare explanations in English, Hindi, and other regional languages.

For multilingual healthcare applications, evidence retrieval and clinical knowledge may be maintained in authoritative source languages while the final explanation is generated in the user's preferred language.

However, multilingual medical generation requires rigorous validation because translation errors, terminology ambiguity, and culturally inappropriate expressions may alter clinical meaning. Consequently, multilingual systems should undergo language-specific clinical evaluation rather than assuming that performance in one language transfers directly to another.

XXI. MULTIMODAL PERSONALIZATION

Future healthcare GenAI systems are likely to integrate multiple modalities, including Text + EHR + Imaging + Laboratory + Wearable\ Data.

Potential inputs include clinical histories, radiological images, blood-test results, electrocardiograms, and wearable-sensor measurements. Multimodal GenAI can potentially synthesize these heterogeneous data sources into a unified patient representation. However, multimodal systems introduce additional challenges related to data alignment, missing modalities, modality-specific uncertainty, privacy, and validation.

XXII. APPLICATION AREAS

A. Diabetes Management

A personalized system can integrate glucose measurements, BMI, medication information, diet, and physical activity to provide individualized educational and decision-support information.

B. Cardiovascular Risk Assessment

Relevant information may include blood pressure, cholesterol, ECG measurements, smoking history, age, and family history. GenAI can assist in communicating risk factors and supporting evidence.

C. Cancer Support

Potential applications include patient education, report summarization, risk communication, and treatment-information support. Autonomous treatment determination should remain outside the scope of a general-purpose GenAI system unless specifically validated and regulated for that purpose.

D. Chronic Disease Management

Longitudinal GenAI systems can support monitoring and communication for chronic conditions such as diabetes, hypertension, kidney disease, and cardiovascular disease.

E. Medication Management

AI can assist in summarizing medication histories, identifying potential interaction information, and organizing adherence-related information. Such outputs require clinical verification before medication changes are made.

F. Preventive Healthcare

Personalized systems can provide evidence-grounded information concerning screening, lifestyle factors, risk factors, and preventive-health measures.

XXIII. COMPARISON WITH CONVENTIONAL SYSTEMS

System	Predictio n	Personalizati on	Explanatio n	Natural- Language Interactio n
Rule-Based System	Yes	Low	High	Low
Statistical ML	Yes	Moderate	Moderate	Low
Deep Learning	Yes	High	Low	Low
ML + XAI	Yes	High	High	Low
LLM	Limited	Moderate	Limited	High
GenAI + RAG	Moderate	High	Moderate	High

System	Predictio n	Personalizati on	Explanatio n	Natural- Language Interactio n
Explainabl e Personalize d GenAI	Yes	High	High	High

The comparison demonstrates that the proposed architecture attempts to combine predictive modeling, explainability, personalization, evidence grounding, and natural-language interaction within a unified framework.

XXIV. MAJOR CHALLENGES

Despite its potential, personalized GenAI introduces substantial technical, ethical, and clinical challenges.

A. Hallucination: GenAI models can generate fluent but unsupported medical statements.

B. Personalization Error: Incorrect, outdated, or incomplete patient information can result in inappropriate personalization.

C. Privacy: Patient-level personalization requires access to sensitive health information, increasing privacy and security requirements.

D. Bias: Training data and model behavior may contain demographic or socioeconomic biases that can affect personalized outputs.

E. Explanation Faithfulness: A fluent explanation may not accurately reflect the actual computational process of the underlying predictive model.

F. Overtrust: Patients or clinicians may place excessive confidence in AI-generated outputs because of their fluent and authoritative presentation.

G. Automation Bias: Clinicians may accept AI recommendations without sufficient independent evaluation.

H. Model Drift: Changes in patient populations, clinical practices, data distributions, and disease patterns can reduce system performance over time.

XXV. PROPOSED EVALUATION FRAMEWORK

The proposed EPG-HDS system should be evaluated at five complementary levels.

Evaluation Level	Representative Metrics
Prediction	Accuracy, Precision, Recall, F1-score, ROC-AUC
Calibration	Brier Score, Expected Calibration Error
Explanation	Fidelity, Completeness, Clarity
Generation	Factuality, Relevance, Hallucination Rate
Clinical	Safety, Utility, Expert Agreement

A conceptual aggregate score can be represented as

$$S = w_pP + w_eE + w_gG + w_cC$$

Where,

S = overall system evaluation score
P = predictive performance

E = explanation quality
 G = generative response quality
 C = clinical utility
 wp = weight assigned to predictive performance
 we = weight assigned to explanation quality
 wg = weight assigned to generation quality
 wc = weight assigned to clinical utility

XXVI. RESEARCH GAP ANALYSIS

The literature reveals several important research gaps.

Limited Integration: GenAI, XAI, and personalization are frequently studied as separate components. Integrated architectures combining all three remain comparatively underexplored.

Limited Patient-Level Validation: A significant proportion of research relies on retrospective, benchmark, or synthetic datasets. Prospective evaluation using realistic patient-level workflows remains necessary.

Weak Explanation Evaluation: Generated explanations are frequently evaluated qualitatively, creating a need for standardized measures of faithfulness, completeness, usefulness, and safety.

Personalization Without Adequate Safety: Personalization can improve relevance while simultaneously increasing the consequences of incorrect patient information or unsupported recommendations.

Multilingual Personalization: Healthcare GenAI research remains disproportionately concentrated on English, leaving substantial opportunities for multilingual and culturally appropriate systems.

Longitudinal Personalization: Many systems rely on static snapshots rather than complete longitudinal patient trajectories.

Causal Explainability: Feature importance indicates association with model output but does not establish causal relationships.

Uncertainty Communication: Personalized systems require mechanisms for communicating predictive and generative uncertainty in clinically understandable forms.

XXVII. PROPOSED RESEARCH METHODOLOGY

A practical experimental methodology for evaluating the proposed framework can be organized into ten phases.

Phase 1: Dataset Collection: A clinically relevant healthcare dataset containing patient-level information is selected according to the target prediction task.

Phase 2: Patient Profiling: Patient records are transformed into structured profiles containing demographic, clinical, behavioral, and longitudinal information.

Phase 3: Predictive Modeling: Multiple machine learning models are trained for the selected healthcare prediction problem.

Phase 4: Model Selection: Models are compared using F1-score, ROC-AUC, sensitivity, specificity, and calibration metrics.

Phase 5: Explainability: Methods such as SHAP, LIME, and counterfactual analysis are applied to obtain interpretable information about model predictions.

Phase 6: Knowledge Retrieval: Relevant clinical guidelines and evidence sources are retrieved using a RAG architecture.

Phase 7: GenAI Integration: The GenAI component generates personalized explanations using patient information, model predictions, XAI outputs, and retrieved evidence.

Phase 8: Verification: Generated responses are evaluated for evidence support, patient-data consistency, hallucination, uncertainty, and clinical safety.

Phase 9: Expert Evaluation: Healthcare professionals assess the generated explanations for correctness, relevance, usefulness, and safety.

Phase 10: User Evaluation: Representative patients or intended end users evaluate clarity, usefulness, trust, and understanding.

XXVIII. FUTURE RESEARCH DIRECTIONS

A. Causal AI: Future systems should move beyond correlational feature attribution toward clinically meaningful causal reasoning. The distinction between association and causation must remain explicit when communicating explanations.

B. Counterfactual Personalized AI: Counterfactual models can generate individualized what-if scenarios and identify potentially actionable changes while respecting clinical constraints.

C. Multimodal GenAI: Future systems can integrate text, EHR information, medical images, laboratory measurements, physiological signals, and wearable-sensor data.

D. Multilingual Healthcare AI: Development of robust healthcare GenAI for Indian and other underrepresented languages represents an important research direction.

E. Privacy-Preserving Personalization: The combination of personalization with federated learning, differential privacy, secure computation, and data minimization can provide a foundation for privacy-aware healthcare AI.

F. Continuous Learning: Future systems may incorporate updated patient information and evolving clinical evidence while maintaining strict validation and safety controls.

G. Digital Twin Integration: Patient-specific computational representations could support simulation and longitudinal decision-support scenarios. However, digital-twin outputs would require extensive clinical validation.

H. Agentic Healthcare AI: Future architectures may use specialized agents for patient profiling, evidence retrieval, prediction, and explanation, verification, and safety assessment. Such systems will require explicit coordination, access control, traceability, and human oversight.

I. Prospective Clinical Validation: Ultimately, the effectiveness of personalized healthcare GenAI should be established through prospective studies conducted within realistic healthcare workflows rather than relying exclusively on offline benchmarks.

XXIX. CONCLUSION

This paper examined the emerging role of explainable generative artificial intelligence in personalized healthcare decision support. Conventional machine learning provides valuable predictive capabilities but often lacks patient-centered communication and interpretable natural-language explanations. Explainable AI addresses aspects of interpretability, while GenAI enables flexible communication, summarization, and information synthesis. Personalization extends these capabilities by incorporating individual patient characteristics, clinical context, longitudinal health information, and communication preferences.

The proposed Explainable Personalized Generative Healthcare Decision Support (EPG-HDS) framework integrates these capabilities through a modular architecture comprising patient profiling, longitudinal modeling, context identification, predictive

analytics, knowledge retrieval, XAI, personalization, GenAI generation, and safety verification.

The framework emphasizes that personalization should not be implemented as a simple mechanism for adding more patient data to a language model. Instead, personalized healthcare AI requires evidence grounding, explanation verification, uncertainty communication, privacy protection, fairness assessment, and clinical oversight.

The major research opportunities include longitudinal patient modeling, causal and counterfactual explainability, multimodal healthcare GenAI, multilingual systems, privacy-preserving personalization, standardized evaluation methodologies, agentic healthcare architectures, and prospective clinical validation.

Ultimately, the goal should not be to develop an AI system that independently determines what is best for a patient. Rather, the objective should be to develop a system capable of understanding patient context, retrieving trustworthy evidence, explaining model outputs, communicating uncertainty, adapting communication to user needs, and assisting clinicians and patients in making better-informed decisions. Such an approach can contribute to the development of healthcare AI that is not only intelligent and personalized, but also explainable, evidence-grounded, safe, and clinically responsible.

XXX. REFERENCES

- [1] M. M. Raza, K. P. Venkatesh, and J. C. Kvedar, "Generative AI and large language models in health care: Pathways to implementation," *npj Digital Medicine*, vol. 7, 2024.
- [2] Y.-J. Park *et al.*, "Assessing the research landscape and clinical utility of large language models: A scoping review," *BMC Medical Informatics and Decision Making*, 2024.
- [3] F. Busch *et al.*, "Current applications and challenges in large language models for patient care: A systematic review," *Communications Medicine*, 2025.
- [4] L. Wang *et al.*, "Applications and Concerns of ChatGPT and Other Conversational Large Language Models in Health Care: Systematic Review," *Journal of Medical Internet Research*, 2024.
- [5] S. Bedi *et al.*, "Testing and Evaluation of Health Care Applications of Large Language Models: A Systematic Review," *JAMA*, 2025.
- [6] A. Vaswani *et al.*, "Attention Is All You Need," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [7] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," in *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016.
- [8] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [9] A. Rajkomar, J. Dean, and I. Kohane, "Machine Learning in Medicine," *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347–1358, 2019.
- [10] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- [11] World Health Organization, *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*, Geneva, Switzerland, 2021.
- [12] World Health Organization, *Ethics and Governance of Artificial Intelligence for Health: Large Multi-Modal Models—WHO Guidance*, Geneva, Switzerland, 2024.
- [13] T. Y. C. Tam *et al.*, "A framework for human evaluation of large language models in healthcare derived from literature review," *npj Digital Medicine*, 2024.
- [14] J. Haltaufderheide and R. Ranisch, "The ethics of ChatGPT in medicine and healthcare: A systematic review on Large Language Models," *npj Digital Medicine*, 2024.
- [15] P. Lewis *et al.*, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems*, vol. 33, 2020.
- [16] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proc. EMNLP-IJCNLP*, 2019, pp. 3982–3992.
- [17] Y. Gao *et al.*, "Retrieval-Augmented Generation for Large Language Models: A Survey," *arXiv preprint arXiv:2312.10997*, 2023.
- [18] L. M. Amugongo *et al.*, "Retrieval augmented generation for large language models in healthcare: A systematic review," *PLOS Digital Health*, 2025.
- [19] S. Liu *et al.*, "Improving large language model applications in biomedicine with retrieval-augmented generation: A systematic review, meta-analysis, and clinical development guidelines," *Journal of the American Medical Informatics Association*, 2025.
- [20] U.S. Food and Drug Administration, "Good Machine Learning Practice for Medical Device Development: Guiding Principle.