HEAL-TX: A Generative AI-Powered Context-Aware Preventive Healthcare

Integrating Multimodal AI for Proactive Health Equity

Nusrat Khan[1], Anjali Shah[2]

Abstract: Preventive healthcare faces a fundamental paradox: while 80% of chronic diseases are preventable, existing systems remain reactive, fragmented, and disconnected from individuals lived realities. This paper presents HEAL-TX (Holistic Ecosystem for Adaptive, Longitudinal, and Transparent Health), a generative AI-powered context-aware preventive healthcare system designed for statewide deployment in Texas. Unlike existing approaches that apply narrow AI models to isolated data streams, HEAL-TX employs a unified multimodal AI architecture that integrates physiological, behavioral, environmental, social, and healthcare system data into a continuous, privacy-preserving health intelligence platform. The system introduces four transformative AI innovations: (1) a Multimodal Fusion Transformer that learns cross-modal representations from heterogeneous temporal data without manual feature engineering; (2) a Generative Digital Twin Framework that creates personalized health simulations enabling counterfactual reasoning and "what-if" intervention planning; (3) a Federated Learning with Differential Privacy architecture that enables population-level model improvement without centralizing sensitive data; and (4) a Conversational AI Health Coach powered by fine-tuned large language models that delivers empathetic, culturally competent guidance in multiple languages. Evaluation using retrospective data from 50,000 patients and a prospective six-month pilot with 1,200 participants across three Texas counties demonstrates that HEAL-TX reduces preventable hospitalizations by 31.7% (surpassing our previous 28.3%), improves medication adherence by 44.2%, and achieves 91.3% user trust ratings. The paper details how AI enables: early warning systems that detect deterioration 5.2 days before clinical presentation, personalized intervention generation that adapts to individual preferences and constraints, and population health analytics that identify emerging outbreaks while preserving individual privacy. We address implementation challenges, partnership models, and policy recommendations for scaling AI-driven preventive healthcare across diverse communities.

Keywords: Generative AI, multimodal transformers, federated learning, preventive healthcare, context-aware systems, digital twins, health equity, large language models, state-level implementation.

Introduction

A. The Preventive Healthcare Imperative

The United States healthcare system spends $4.5 trillion annually—17.3% of GDP—yet ranks last among developed nations in health outcomes [1]. This paradox stems from a fundamental misalignment: 95% of healthcare expenditures target treatment rather than prevention, despite compelling evidence that 80% of heart disease, stroke, and type 2 diabetes, and 40% of cancers are preventable through lifestyle modification and early intervention [2].

Texas exemplifies both the urgency and complexity of this challenge. With 30 million residents spanning 268,000 square miles—from the sprawling Houston metroplex to remote West Texas frontier communities—the state confronts staggering health disparities. Life expectancy varies by 20 years between affluent Dallas suburbs and impoverished Rio Grande Valley colonias [3]. Rural counties have 40% fewer primary care physicians per capita than urban centers [4]. One in five Texans remains uninsured, the highest rate in the nation [5].

In this landscape, traditional preventive approaches fail. Mailed reminders about mammograms cannot reach women without stable addresses. Generic "exercise more" recommendations ignore that 23% of Texans live in food deserts with no safe places to walk [6]. Blood pressure alerts mean little to patients who cannot afford medications or take time off work for follow-up appointments.

B. Why AI Changes Everything

Artificial intelligence is not merely incremental improvement over traditional analytics—it represents a fundamental paradigm shift in what preventive healthcare can achieve. Traditional approaches rely on population averages, static rules, and retrospective analysis. AI enables:

1. **From Population Averages to Individual Trajectories**
   Traditional risk scores (e.g., Framingham, ASCVD) assign the same risk to everyone with identical clinical measurements. AI learns that two patients with identical blood pressure readings may have dramatically different trajectories based on subtle patterns in heart rate variability, sleep architecture, activity rhythms, and environmental exposures [7].

2. **From Single Time points to Continuous Dynamics**
   A clinic visit captures a 15-minute snapshot. AI analyzes months of continuous data, detecting that a patient's blood pressure has been rising gradually for weeks, that their sleep efficiency is declining, and that these changes correlate with worsening air quality near their workplace [8].

3. **From Single Diseases to Multimorbidity Patterns**
   Traditional models predict one disease at a time. AI discovers that depression, diabetes, and cardiovascular disease share common precursors—sleep disruption, inflammation, social isolation—and that interventions addressing these root causes prevent multiple conditions simultaneously [9].

4. **From Retrospective Analysis to Counterfactual Reasoning**
   Traditional analytics ask "what happened?" AI asks "what would happen if...?" enabling personalized simulation of intervention outcomes before implementation [10].

5. **From One-Way Communication to Empathetic Dialogue**
   Traditional reminders broadcast generic messages. Conversational AI engages patients in natural dialogue, understanding their unique barriers, motivations, and circumstances, and adapting guidance accordingly [11].

C. The Generative AI Revolution in Healthcare

Recent advances in generative AI—models that create rather than merely classify—have transformed what's possible. Large language models (LLMs) like GPT-4 and Gemini demonstrate human-level language understanding and generation [12]. Multimodal transformers can integrate text, images, time-series data, and structured records into unified representations [13]. Diffusion models generate realistic synthetic data that preserves statistical properties while protecting privacy [14].

HEAL-TX harnesses these advances within a unified architecture designed for preventive healthcare. Unlike single-purpose AI applications that predict readmissions or detect arrhythmias, HEAL-TX employs generative AI as an integrated intelligence layer that:

- Continuously learns individual health patterns
- Simulates intervention outcomes before recommendation
- Generates personalized explanations and guidance
- Adapts to user preferences and health literacy
- Enables population-level insights without compromising privacy

D. Research Contributions

This paper presents the first comprehensive generative AI-powered context-aware preventive healthcare system designed for state-level deployment. Our contributions include:

1. **A Multimodal Fusion Transformer Architecture** that learns integrated representations from heterogeneous temporal data without manual feature engineering, achieving state-of-the-art prediction accuracy.
2. **A Generative Digital Twin Framework** that creates personalized health simulations enabling counterfactual reasoning, intervention optimization, and early warning generation.
3. **A Federated Learning with Differential Privacy Infrastructure** that enables continuous model improvement across institutions without centralizing sensitive data.
4. **A Conversational AI Health Coach** powered by fine-tuned LLMs that delivers empathetic, culturally competent guidance in English, Spanish, Vietnamese, and Mandarin.
5. **Empirical Evidence** from a six-month pilot demonstrating that AI-powered preventive healthcare reduces hospitalizations by 31.7%, improves medication adherence by 44.2%, and achieves 91.3% user trust.
6. **A Scalable Implementation Framework** for deploying AI-driven preventive healthcare across diverse populations and geographic regions.

E. Paper Organization

Section II reviews the evolution of AI in healthcare and identifies gaps addressed by HEAL-TX. Section III presents the system architecture and AI methodology in detail. Section IV describes experimental setup and results. Section V discusses implications, limitations, and future directions. Section VI concludes with policy recommendations for statewide AI deployment.

Literature Review: The Evolution of AI in Preventive Healthcare

A. First Generation: Rule-Based Expert Systems (1970s-1990s)

Early AI in healthcare relied on manually encoded expert knowledge. MYCIN, developed at Stanford in the 1970s, used 600 rules to diagnose bacterial infections and recommend antibiotics [15]. INTERNIST-I attempted to capture all internal medicine knowledge through disease-disease and disease-symptom relationships [16].

Limitations: Rule-based systems require exhaustive knowledge engineering, cannot learn from data, fail when encountering edge cases, and do not adapt to individual patients. They represent human knowledge encoded in software rather than genuine machine intelligence.

B. Second Generation: Machine Learning for Specific Predictions (2000s-2015)

The availability of electronic health records enabled machine learning models trained on historical data. Researchers developed predictors for readmission [17], sepsis onset [18], and medication non-adherence [19]. Random forests and support vector machines became standard tools.

Key achievements included:

- Predicting 30-day readmissions with AUC 0.72-0.76 [20]
- Detecting sepsis 4-6 hours earlier than clinical recognition [21]
- Identifying patients at risk for diabetes complications [22]

Limitations: These models were narrow (one prediction per model), required extensive feature engineering, could not integrate multiple data modalities, and provided no explanations for their predictions. They augmented specific clinical decisions but did not transform preventive care delivery.

C. Third Generation: Deep Learning for Multimodal Integration (2015-2022)

Deep learning enabled automatic feature learning from raw data. Convolutional neural networks analyzed medical images [23]; recurrent neural networks processed time-series from wearables [24]; natural language processing extracted insights from clinical notes [25].

The MIMIC-III benchmark demonstrated that multimodal deep learning outperforms single-modality approaches [26]. Huang et al. showed that integrating vital signs, labs, and notes improved mortality prediction by 12-18% [27].

Limitations: Deep learning models remained discriminative rather than generative—they classified existing data but could not create new insights. They required centralized training data, raising privacy concerns. They provided limited interpretability, reducing clinician trust.

D. Fourth Generation: Generative AI and Foundation Models (2022-Present)

The transformer architecture [28] enabled a revolution in AI capabilities. Large language models demonstrated that scaling data and compute produces emergent abilities—reasoning, instruction following, in-context learning—that were not explicitly programmed [29].

In healthcare, foundation models are now emerging:

- **Med-PaLM 2** achieves 86.5% accuracy on medical exam questions [30]
- **BioGPT** generates coherent biomedical text [31]
- **ClinicalBERT** learns representations from clinical notes [32]
- **HEAL** (our prior work) demonstrated multimodal fusion for health prediction [33]

The critical insight: foundation models learn general representations that can be adapted to multiple downstream tasks, eliminating the need for task-specific models.

E. Key AI Capabilities for Preventive Healthcare

Recent advances enable five capabilities essential for context-aware prevention:

1. **Multimodal Fusion**
   Transformers can integrate heterogeneous data types—numerical time-series, categorical variables, text, images—into unified representations. This is critical because health emerges from interactions among physiological, behavioural, and environmental factors [34].

2. **Temporal Modelling**
   Attention mechanisms capture long-range dependencies in time-series data, enabling detection of subtle patterns that precede health events [35]. The model learns that gradual changes over weeks matter more than acute fluctuations.

3. **Counterfactual Reasoning**
   Generative models can simulate alternative scenarios—"what if this patient had better sleep?"—enabling personalized intervention planning [36]. This moves beyond prediction to actionable insight.

4. **Natural Language Interaction**
   LLMs enable bidirectional communication between patients and AI systems. Patients can ask questions, report symptoms, express concerns; the AI can explain recommendations, provide education, offer encouragement [37].

5. **Privacy-Preserving Learning**
   Federated learning enables model training across distributed data sources without centralizing sensitive information [38]. Differential privacy ensures that individual records cannot be reconstructed from model updates [39].

F. Gaps Addressed by HEAL-TX

Despite these advances, no existing system integrates all capabilities into a unified preventive healthcare platform. Specific gaps include:

1. **Fragmentation**: Existing AI applications address single tasks—readmission prediction, arrhythmia detection, medication reminders—without coordination.
2. **Context Blindness**: Most models ignore environmental and social determinants, which account for 50% of health outcomes [40].
3. **Intervention Optimization Gap**: Prediction without actionable intervention leaves patients and clinicians uncertain what to do.
4. **Trust Deficit**: Black-box models reduce clinician and patient trust, limiting adoption [41].
5. **Scalability Gap**: Research prototypes rarely address the infrastructure, partnership, and policy requirements for statewide deployment.

HEAL-TX addresses each gap through integrated generative AI architecture designed from the outset for real-world impact.

Methodology: Generative AI Architecture

A. System Overview

HEAL-TX employs a six-layer AI architecture (Figure 1) designed for continuous learning, privacy preservation, and human-cantered interaction:

1. Data Acquisition Layer: Secure ingestion from wearables, environmental sensors, EHRs, and user input
2. Multimodal Fusion Layer: Transformer-based representation learning from heterogeneous data
3. Generative Digital Twin Layer: Personalized health simulation and counterfactual generation
4. Intervention Optimization Layer: Multi-objective reinforcement learning for recommendation generation
5. Conversational AI Layer: LLM-powered natural language interaction
6. Federated Learning Layer: Distributed model improvement across institutions

Figure 1: HEAL-TX Generative AI Architecture

B. Multimodal Fusion Transformer

Traditional approaches process each data modality separately, then combine features. This fails to capture cross-modal interactions—for example, how heart rate response to air pollution depends on activity level and time of day.

HEAL-TX employs a **Multimodal Fusion Transformer (MFT)** that learns integrated representations through cross-modal attention. The architecture builds on the Perceiver IO framework [42] but extends it with healthcare-specific adaptations.

- **Input Encoding**

Each data modality is encoded into a sequence of tokens:

- **Physiological time-series** (heart rate, HRV, SpO2, etc.) are encoded using 1D convolutional projections, creating tokens for each timepoint
- **Environmental data** (air quality, temperature, pollen) are similarly encoded and aligned temporally
- **Social determinants** (census tract data, social support indicators) are embedded as static tokens
- **Clinical text** (diagnoses, medications, notes) is encoded using a lightweight clinical BERT [32]
- **Cross-Modal Attention**

The MFT applies alternating layers of self-attention within modalities and cross-attention between modalities. This enables the model to learn that:

- Elevated heart rate during high-pollution periods signals different risk than during exercise
- Social isolation amplifies the health impact of poor sleep
- Medication non-adherence patterns differ by season and day of week

- **Temporal Context Window**

The model maintains a rolling context window of 90 days, with attention mechanisms that weight recent data more heavily while preserving long-term patterns. This enables detection of gradual deteriorations that unfold over weeks.

- **Mathematical Formulation**

Let $X = \{x_1, x_2, ..., x_m\}$ be the set of modality-specific input sequences. The MFT computes:

$H^0 = [E_1(x_1), E_2(x_2), ..., E_m(x_m)]$ # Initial encodings
For layer l = 1 to L:
H_l_self = MultiHeadSelfAttention(H^{l-1})
H_l_cross = MultiHeadCrossAttention(H_l_self, H^{l-1})
H^l = LayerNorm(H_l_cross + FFN(H_l_cross))

The final representation h_final aggregates information across all modalities and timepoints, serving as input to downstream tasks.

C. Generative Digital Twin Framework

The critical innovation in HEAL-TX is the **Generative Digital Twin**—a personalized generative model that simulates an individual's health dynamics and enables counterfactual reasoning.

- **Model Architecture**
  We employ a **Conditional Variational Autoencoder with Transformer backbone (CVAE-Trans)**. The model learns a latent representation of health trajectories conditioned on:
  - Static individual factors (age, sex, genetics, comorbidities)
  - Dynamic contextual factors (environment, behavior, social support)
  - Intervention history (medications, lifestyle changes, care received)

- **Training**
  The CVAE-Trans is trained on the retrospective cohort of 50,000 patients to reconstruct observed health trajectories and predict future states. The loss function combines:

L = Reconstruction + β * L_KL + Prediction

Where:
  - Reconstruction ensures accurate reconstruction of observed data
  - L_KL regularizes the latent space
  - Prediction optimizes forecasting accuracy

- **Counterfactual Generation**
  Once trained, the digital twin enables "what-if" simulation:
  For a given individual with latent representation z and context c, we can generate counterfactual trajectories by:
  1. Modifying the context (e.g., improving air quality, increasing social support)
  2. Sampling from the conditional distribution p(future | z, commodified)
  3. Comparing simulated outcomes with actual trajectory

This enables personalized answers to questions like:
  - "What would my 5-year diabetes risk be if I improved sleep consistency?"
  - "How much would increasing weekly exercise by 60 minutes reduce my cardiovascular risk?"
  - "Would moving to a neighborhood with better air quality improve my asthma control more than increasing medications?"

- **Early Warning Detection**
  The digital twin continuously compares predicted health trajectory with actual observations. Deviations beyond confidence intervals trigger early warnings. For example, if the model predicts stable blood pressure based on historical patterns but observes a rising trend unexplained by context changes, it flags potential deterioration 5-7 days before clinical thresholds are crossed.

D. Intervention Optimization via Multi-Objective Reinforcement Learning

Generating recommendations requires optimizing across potentially conflicting objectives: reducing cardiovascular risk, improving mental health, minimizing medication burden, respecting patient preferences, and ensuring feasibility.

HEAL-TX formulates this as a **multi-objective reinforcement learning (MORL) problem** [43].

- **State Space**

The state s_t at time t includes:

- Current digital twin latent representation
- Recent health metrics and trends
- Recent intervention adherence
- User-reported preferences and barriers
- Available resources (medications, facilities, support)

- **Action Space**

Actions a_t include:

- Medication adjustments (with clinician oversight)
- Lifestyle recommendations (exercise, diet, sleep)
- Environmental modifications (air purifiers, relocation guidance)
- Social support interventions (support groups, community connections)
- Healthcare utilization (appointment reminders, screening prompts)

- **Reward Functions**

Multiple reward functions capture different objectives:

- r_1: Reduction in predicted disease risk
- r_2: Improvement in quality of life (from patient-reported outcomes)
- r_3: Adherence feasibility (estimated probability patient will follow)
- r_4: Cost-effectiveness
- r_5: Patient preference alignment

- **Pareto Frontier Optimization**

Rather than combining rewards into a single scalar (which requires arbitrary weighting), we compute the **Pareto frontier** of non-dominated policies—those where no objective can be improved without degrading another [44].

The MORL algorithm explores the policy space using a variant of Deep Q-Networks extended to multi-objective settings [45]. For each state, it returns the set of Pareto-optimal actions with their projected outcome vectors.

- **User Preference Integration**

When multiple Pareto-optimal options exist, the system engages the user through the conversational AI to understand preferences:

"Based on your health data, I've found three ways to reduce your heart attack risk. Option A focuses on medication adjustment and would give you the largest risk reduction but requires a new prescription. Option B emphasizes lifestyle changes—increasing your daily steps and improving sleep—which takes more time but has no medication side effects. Option C combines smaller changes in both areas. Which approach feels most manageable for you?"

E. Conversational AI Health Coach

The conversational interface is powered by a **fine-tuned large language model** based on Gemini Pro architecture [12]. Fine-tuning uses:

1. **Domain adaptation data**: De-identified clinical conversations, patient education materials, and health behavior change dialogues
2. **Instruction tuning**: Examples of explaining medical concepts, providing emotional support, and guiding behavior change
3. **Safety filtering**: Reinforcement learning from human feedback (RLHF) to ensure responses are accurate, empathetic, and never harmful

- **Key Capabilities**

1. Multi-lingual Support

The model operates in English, Spanish, Vietnamese, and Mandarin—the four most common languages in Texas—with real-time translation for less common languages via API integration.

2. Empathetic Response Generation

Beyond information provision, the model recognizes and responds to emotional states. If a user expresses frustration about exercise, the AI might respond:

"It sounds like you're feeling discouraged about your activity goals. That's completely understandable—making lifestyle changes is genuinely hard. Many people I work with find it helpful to start with very small goals they're confident they can achieve. What's one tiny change you feel you could make this week?"

3. Health Literacy Adaptation

The model assesses user health literacy through conversation and adjusts explanation complexity accordingly. For users with limited health literacy, it uses plain language, concrete examples, and frequent comprehension checks.

4. Behavior Change Support

Drawing on motivational interviewing principles [46], the AI:

- Elicits user's own motivations for change
- Explores ambivalence without judgment
- Supports self-efficacy
- Helps develop concrete action plans

5. Clinician Collaboration

When users report symptoms or concerns requiring clinical attention, the AI:

- Gathers structured information
- Generates a summary for the care team
- Schedules follow-up (if integrated with clinic systems)
- Provides guidance on when to seek emergency care

F. Federated Learning with Differential Privacy

Training powerful AI models typically requires centralizing large datasets, raising privacy concerns. HEAL-TX solves this through **federated learning** [38] with **differential privacy** [39].

- **Federated Learning Workflow**
  1. **Local Training**: Each participating institution (hospital, clinic, health system) maintains local patient data. The global model is distributed to each site.
  2. **Local Updates**: Each site computes model updates (gradients) using its local data, improving the model based on its patient population.
  3. **Secure Aggregation**: Updates are encrypted and sent to a central server that aggregates them without ever seeing individual updates.
  4. **Global Model Update**: The aggregated update improves the global model, which is then redistributed.

- **Differential Privacy Guarantees**

To prevent inference attacks that might reconstruct patient information from model updates, we add calibrated noise to each update:

$\theta\_shared = \theta\_local + Lap(0, \lambda)$

Where $Lap(0, \lambda)$ is Laplace noise with scale λ chosen to satisfy (ε, δ)-differential privacy. This ensures that even if an adversary accesses the aggregated model, they cannot determine whether any specific individual's data contributed to training.

- **Empirical Privacy-Accuracy Trade-off**

We evaluated privacy-accuracy trade-offs across ε values from 0.1 (high privacy) to 10 (low privacy). At $\varepsilon = 2$ (the common industry standard), we observed only a 3.2% reduction in prediction accuracy compared to non-private training, demonstrating that strong privacy guarantees are compatible with high performance.

G. Human-Centered AI Design Principles

Technical sophistication must be balanced with human needs. HEAL-TX incorporates five design principles derived from user research:

1. **Explainability by Default**

All AI recommendations include natural language explanations of why they were made and what evidence supports them. Users can drill down for more detail or request simpler explanations.

2. **User Control**

Users can view, correct, or delete their data; can adjust AI behavior (e.g., "remind me less often," "focus on mental health this month"); and can override AI recommendations.

3. **Appropriate Trust**

The AI communicates uncertainty: "I'm 85% confident this recommendation will help, based on patterns from people with similar health profiles." It never presents predictions as certainties.

5. **Cultural Competence**

The AI is trained on diverse data and explicitly calibrated for Texas populations. It understands that health beliefs, dietary patterns, and family structures vary across communities and adapts accordingly.

6. **Continuous Improvement**

User feedback (explicit ratings and implicit signals like recommendation acceptance) continuously refines the model. The system learns which approaches work for which individuals.

Experiments and Results

A. Study Design

HEAL-TX was evaluated through three complementary approaches:

1. **Retrospective Validation**: Simulation using de-identified EHR data from 50,000 patients across three Texas health systems (2019-2023), comparing HEAL-TX predictions against actual outcomes.
2. **Prospective Pilot**: Six-month deployment with 1,200 participants across three diverse Texas counties: Harris (urban, high diversity, air quality challenges), Travis (urban, technology-rich, high health literacy), and Reeves (rural, underserved, limited broadband).
3. **Ablation Studies**: Systematic removal of AI components to isolate contributions of each innovation.

Participant demographics reflected Texas diversity: 42% Hispanic/Latino, 24% non-Hispanic White, 18% African American, 8% Asian American, 8% other; age range 18-89; 52% female, 48% male; 23% rural, 77% urban/suburban.

B. Primary Clinical Outcomes

C. Reduction in Preventable Hospitalizations

HEAL-TX participants experienced 31.7% fewer preventable hospitalizations compared to matched controls ($p < 0.001$). This exceeds our previous 28.3% [33], demonstrating that generative AI capabilities add significant value.

The effect varied by population segment:

- **Rural participants**: 34.2% reduction (95% CI: 28.7-39.8%)
- **Urban participants**: 29.8% reduction (95% CI: 25.3-34.2%)
- **Multiple chronic conditions (≥3)**: 37.1% reduction (95% CI: 31.5-42.6%)
- **Single chronic condition**: 26.3% reduction (95% CI: 21.8-30.9%)

D. Medication Adherence

Adherence improved by 44.2% (from baseline 58.3% to 84.1% at 6 months, $p < 0.001$). The conversational AI coach was cited by 82% of participants as the primary reason for improvement. Qualitative analysis revealed that the AI's ability to understand individual barriers—forgetfulness, cost concerns, side effect fears, complex schedules—and address them empathetically was critical.

E. Early Detection Performance

HEAL-TX identified 237 potential health deteriorations before clinical presentation, with 178 confirmed (positive predictive value: 75.1%). Detected conditions included:

*Depression confirmation challenging due to subjective nature; many unconfirmed cases may still represent genuine deterioration

Mean lead time before clinical presentation was 5.2 days—sufficient for early intervention in most cases.

F. AI Component Ablation Studies

To understand each AI component's contribution, we systematically removed capabilities and measured performance impact

Key insights:

- Multimodal fusion and digital twin contribute most to prediction accuracy
- Conversational AI is critical for behavior change (adherence)
- Federated learning preserves performance while adding privacy
- Each component adds measurable value

G. Comparison with State-of-the-Art

HEAL-TX achieves state-of-the-art performance across all metrics, with particularly large gains in adherence prediction and user trust.

H. Federated Learning Performance

We simulated federated learning across 10 distributed sites with heterogeneous patient populations. Key findings:

- Convergence: The federated model reached 98.2% of centralized performance after 50 communication rounds
- Communication efficiency: Model compression reduced communication costs by 73% with minimal accuracy loss
- Heterogeneity robustness: Performance varied by only 2.1% across sites with different population characteristics
- Privacy-utility trade-off: At $\varepsilon=2$, accuracy decreased by 3.2% compared to non-private training

I. User Experience and Trust

Participant surveys revealed exceptional engagement and trust:

- 91.3% agreed "HEAL-TX respects my privacy"
- 89.7% agreed "recommendations feel genuinely personalized"
- 88.2% agreed "I understand how HEAL-TX uses my data"
- 93.1% agreed "the conversational AI is easy to talk to"
- 87.4% agreed "I trust the AI's recommendations"

Qualitative themes from interviews:

Empathy: "It doesn't feel like a robot. When I told it I was stressed about work, it didn't just tell me to meditate—it asked what was causing the stress and helped me think through solutions." (Participant 382, Travis County)

Understanding: "It knows I work nights, so it doesn't tell me to get sunlight in the morning like every other app. It actually gets my schedule." (Participant 156, Harris County)

Empowerment: "For the first time, I feel like I'm in control of my health instead of just going to the doctor when something's wrong." (Participant 891, Reeves County)

J. Scalability Simulation

We simulated statewide deployment for 10 million users (approximately one-third of Texas population) with the following results:

- Infrastructure costs: $38-52 million (edge devices, servers, networking)
- Annual operating costs: $16-22 million (cloud compute, model retraining, support)
- Projected annual savings: $410-520 million (reduced hospitalizations, ER visits)
- Break-even timeline: 11-15 months
- Bandwidth requirements: 2.3 Mbps average per user (peak 5.1 Mbps) — within existing infrastructure with modest rural upgrades ($8-12 million)

Discussion

A. Interpretation of Findings

HEAL-TX demonstrates that generative AI, thoughtfully integrated across the preventive healthcare continuum, delivers clinically meaningful improvements that exceed previous approaches. The 31.7% reduction in preventable hospitalizations represents not just statistical significance but genuine transformation in patients' lives—fewer crises, less suffering, more time at home with family.

The 44.2% improvement in medication adherence addresses one of healthcare's most intractable problems. Non-adherence causes approximately 125,000 deaths annually and costs $100-300 billion [48]. That a conversational AI, deployed at scale, can achieve adherence rates approaching those of intensive human coaching (which costs 20-50x more) suggests a fundamental rethinking of chronic disease management.

The early warning performance—5.2 days lead time with 75% positive predictive value—demonstrates that continuous context monitoring enables genuinely preventive action. Clinicians receiving these alerts can intervene before patients deteriorate to the point of requiring hospitalization.

B. How AI Enables These Outcomes

The ablation studies reveal why AI matters. Removing multimodal fusion dropped effectiveness by 24%, confirming that health outcomes emerge from interactions among physiological, behavioral, environmental, and social factors. Models trained on single modalities miss these interactions.

Removing the generative digital twin reduced early detection by 22%. Traditional discriminative models classify risk but cannot simulate counterfactuals or detect subtle deviations from expected trajectories. The digital twin's ability to ask "is this patient deviating from their personalized expected path?" enables earlier, more sensitive detection.

Removing conversational AI reduced adherence improvement by 34%—the largest single-component impact on behavior change. Information alone rarely changes behavior; people need understanding, support, and motivation tailored to their circumstances. The AI coach provides this at scale.

C. The Privacy Paradox Resolved

A common concern with AI in healthcare is that better performance requires sacrificing privacy. HEAL-TX demonstrates that federated learning with differential privacy resolves this paradox. By training models across distributed data without centralizing sensitive information, and by adding mathematical guarantees against inference attacks, we achieve 98% of centralized performance while preserving privacy.

The high user trust ratings (91.3%) suggest that privacy-preserving architecture, when transparently communicated, actually increases adoption. Users who understand that their data never leaves their device or hospital are more willing to engage.

D. Health Equity Implications

The larger effects in rural (34.2%) and multi-morbid (37.1%) populations suggest that AI-powered preventive healthcare may help reduce disparities. Populations with limited access to traditional care benefit disproportionately from continuous monitoring and support.

However, realizing these equity benefits requires intentional design. Our loaner device program for low-income participants, multi-lingual support, and community health worker partnerships were essential. Technology alone cannot overcome structural inequities, but thoughtfully deployed technology can avoid exacerbating them.

E. Limitations

This study has important limitations:

1. **Pilot Duration**: Six months is insufficient to assess long-term outcomes for slowly progressing conditions. Three-year follow-up is ongoing.
2. **Selection Bias**: Participants were volunteers who may differ from the general population in motivation and digital readiness.
3. **Texas Specificity**: Findings may not generalize to states with different demographics, healthcare infrastructure, or policy environments.
4. **Algorithm Limitations**: Performance degrades as number of conditions increases beyond 10 due to combinatorial complexity.
5. **Clinician Integration**: Pilot focused on patient-facing AI; integration with clinician workflows requires further development.

6. **Cost Estimates**: Projected costs are based on simulations; actual deployment may reveal unforeseen expenses.

F. Future Work

Ongoing and planned research includes:

1. **Reinforcement Learning from Human Feedback**
   We are implementing RLHF to continuously improve the conversational AI based on user ratings and outcomes [49]. This will enable the system to learn which communication styles work best for different populations.

2. **Multimodal Foundation Model**
   We are developing a healthcare-specific foundation model trained on >100 million patient-years of data, enabling few-shot adaptation to new conditions and populations [50].

3. **Causal Inference Integration**
   Current counterfactuals are based on observational correlations. We are integrating causal inference methods to better estimate true intervention effects [51].

4. **Clinician Decision Support**
   We are developing clinician-facing interfaces that summarize AI insights, explain reasoning, and support shared decision-making.

5. **Multi-State Expansion**
   Partnerships with California, Florida, and New York health systems will test generalizability across different populations and policy contexts.

6. **Policy Engagement**
   Working with Texas HHSC and CMS to develop reimbursement models for AI-powered preventive services.

Conclusion: Toward AI-Powered Preventive Healthcare at Scale

HEAL-TX demonstrates that generative AI, integrated across the full spectrum of contextual factors and deployed with privacy-preserving architecture, can transform preventive healthcare. The system achieves what previously seemed impossible: continuous, personalized, empathetic health support for millions of people at sustainable cost.

For Texas—a state of profound diversity and disparity—such a system offers particular promise. Rural communities gain access to continuous monitoring that partially offsets provider shortages. Urban populations benefit from environmental alerts tailored to their specific locations and health conditions. Underserved groups receive support that adapts to their circumstances rather than imposing one-size-fits-all recommendations.

The path to statewide implementation requires:

1. Infrastructure Investment

Expanding rural broadband, subsidizing device access for low-income populations, and building secure data aggregation infrastructure.

2. Partnership Development

Collaboration among universities (research and validation), health systems (clinical integration), community organizations (trust building and support), and state agencies (policy and reimbursement).

3. Workforce Training

Preparing clinicians to work effectively with AI systems, and training community health workers to support AI adoption among digitally hesitant populations.

4. Policy Innovation

Creating reimbursement models that reward prevention rather than procedures, and establishing regulatory frameworks that ensure AI safety without stifling innovation.

5. Continuous Community Engagement

Ensuring that AI systems reflect community values, address community priorities, and remain accountable to the people they serve.

The ultimate measure of HEAL-TX is not technological sophistication but human impact. When a grandmother in rural Reeves County receives an alert that her health trajectory is diverging from her personalized baseline, along with a supportive conversation about what might be causing it and what she can do—and when that alert prevents a hospitalization that would have separated her from her family—technology fulfills its highest purpose.

Generative AI, thoughtfully designed and equitably deployed, can help us build a future where healthcare is not something that happens to people in moments of crisis, but something that supports them continuously in living healthier, fuller lives. That future is within reach.

Acknowledgments

This research was supported by the Texas Health and Human Services Commission (Grant No. THHSC-2024-1782), the National Institute of Biomedical Imaging and Bioengineering (Award No. R01EB032345), and a Google Research Collaboration Award. The authors thank the 1,200 participants in Harris, Travis, and Reeves counties for their time and trust; the community health workers who made recruitment and retention possible; and the clinicians who provided guidance and feedback throughout.

References

[1]. Centers for Medicare and Medicaid Services, "National Health Expenditure Projections 2023-2031," CMS Office of the Actuary, Baltimore, MD, 2023.

[2]. R. Raghupathi and W. Raghupathi, "An empirical study of chronic diseases in the United States: A visual analytics approach," *Int. J. Environ. Res. Public Health*, vol. 15, no. 3, p. 431, 2018.

[3]. University of Wisconsin Population Health Institute, "County Health Rankings & Roadmaps 2024," Madison, WI, 2024.

[4]. Texas Department of State Health Services, "Texas Health Workforce Strategic Plan 2024," Austin, TX, 2024.

[5]. U.S. Census Bureau, "Health Insurance Coverage in the United States: 2023," Washington, DC, 2024.

[6]. U.S. Department of Agriculture, "Food Access Research Atlas," Economic Research Service, Washington, DC, 2023.

[7]. A. Mohamed, M. Abdelrehim, and R. Al-Barazie, "Context matters in machine learning based disease prediction with insights from diverse clinical and symptom data," *Sci. Rep.*, vol. 15, no. 1, pp. 1-15, 2025.

[8]. M. Snyder, "The digital transformation of health care," *N. Engl. J. Med.*, vol. 385, no. 6, pp. 489-491, 2021.

[9]. C. Lu, T. Han, and Y. Ning, "Context-aware health event prediction via transition functions on dynamic disease graphs," in *Proc. 36th AAAI Conf. Artif. Intell.*, 2022, pp. 4567-4574.

[10]. IMPACT System Research Group, "IMPACT: An interactive multi-disease prevention and counterfactual treatment system using explainable AI and a multimodal LLM," *PeerJ Comput. Sci.*, vol. 11, p. e2839, 2025.

[11]. B. Meskó, "Prompt engineering as an important emerging skill for medical professionals," *JMIR Med. Educ.*, vol. 9, p. e50638, 2023.

[12]. Gemini Team Google et al., "Gemini: A family of highly capable multimodal models," arXiv preprint arXiv: 2312.11805, 2023.

[13]. A. Radford et al., "Learning transferable visual models from natural language supervision," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 8748-8763.

[14]. J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 6840-6851.

[15]. E. H. Shortliffe, *Computer-Based Medical Consultations: MYCIN*. New York: Elsevier, 1976.

[16]. R. A. Miller, H. E. Pople Jr., and J. D. Myers, "INTERNIST-I, an experimental computer-based diagnostic consultant for general internal medicine," *N. Engl. J. Med.*, vol. 307, no. 8, pp. 468-476, 1982.

[17]. S. F. Jencks, M. V. Williams, and E. A. Coleman, "Rehospitalizations among patients in the Medicare fee-for-service program," *N. Engl. J. Med.*, vol. 360, no. 14, pp. 1418-1428, 2009.

[18]. K. E. Henry, D. N. Hager, P. J. Pronovost, and S. Saria, "A targeted real-time early warning score (TREWScore) for septic shock," *Sci. Transl. Med.*, vol. 7, no. 299, p. 299ra122, 2015.

[19]. L. Osterberg and T. Blaschke, "Adherence to medication," *N. Engl. J. Med.*, vol. 353, no. 5, pp. 487-497, 2005.

[20]. A. Hammoudeh, M. Al-Tawil, and R. Al-Sayyed, "Predicting 30-day hospital readmission using machine learning: A systematic review," *J. Med. Syst.*, vol. 45, no. 8, p. 78, 2021.

[21]. M. E. Klompas, "Early detection of sepsis: Are we there yet?" *Clin. Infect. Dis.*, vol. 61, no. 7, pp. 1047-1049, 2015.

[22]. R. W. Grant et al., "Development and validation of a model to predict diabetes complications," *Diabetes Care*, vol. 42, no. 8, pp. 1463-1470, 2019.

[23]. A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115-118, 2017.

[24]. N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," in *Proc. 25th Int. Joint Conf. Artif. Intell.*, 2016, pp. 1533-1540.

[25]. E. Alsentzer et al., "Publicly available clinical BERT embeddings," in *Proc. 2nd Clin. Nat. Lang. Process. Workshop*, 2019, pp. 72-78.

[26]. A. E. Johnson et al., "MIMIC-III, a freely accessible critical care database," *Sci. Data*, vol. 3, p. 160035, 2016.

[27]. S. Huang, G. Nalepa, and W. Z. Wang, "Multimodal learning for mortality prediction in ICU," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 6, pp. 2198-2208, 2021.

[28]. A. Vaswani et al., "Attention is all you need," in *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5998-6008.

[29]. J. Kaplan et al., "Scaling laws for neural language models," arXiv preprint arXiv:2001.08361, 2020.

[30]. K. Singhal et al., "Large language models encode clinical knowledge," *Nature*, vol. 620, pp. 172-180, 2023.

[31]. R. Luo et al., "BioGPT: Generative pre-trained transformer for biomedical text generation and mining," *Brief. Bioinform.*, vol. 23, no. 6, p. bbac409, 2022.

[32]. E. Alsentzer et al., "Publicly available clinical BERT embeddings," in *Proc. 2nd Clin. Nat. Lang. Process. Workshop*, 2019, pp. 72-78.

[33]. R. Sharma et al., "HEAL: A multimodal transformer for health event prediction," *IEEE J. Biomed. Health Inform.*, vol. 28, no. 3, pp. 1456-1467, 2024.

[34]. T. Vats et al., "Explainable context-aware IoT framework using human digital twin for healthcare," *Multimedia Tools Appl.*, pp. 1-25, 2023.

[35]. Z. C. Lipton, D. C. Kale, and R. Wetzel, "Modeling missing data in clinical time series with RNNs," in *Proc. Mach. Learn. Healthc. Conf.*, 2016, pp. 1-15.

[36]. R. Moraffah et al., "Causal interpretability for machine learning: Problems and solutions," in *Proc. 27th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2021, pp. 3997-3998.

[37]. S. Lee et al., "Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine," *N. Engl. J. Med.*, vol. 388, no. 13, pp. 1233-1239, 2023.

[38]. B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Stat.*, 2017, pp. 1273-1282.

[39]. C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3-4, pp. 211-407, 2014.

[40]. World Health Organization, "Social determinants of health," Geneva, Switzerland, 2023. [Online]. Available: https://www.who.int/health-topics/social-determinants-of-health

[41]. R. Caruana et al., "Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2015, pp. 1721-1730.

[42]. A. Jaegle et al., "Perceiver IO: A general architecture for structured inputs & outputs," in *Int. Conf. Learn. Represent.*, 2022.

[43]. C. F. Hayes et al., "A practical guide to multi-objective reinforcement learning and planning," *Auton. Agents Multi-Agent Syst.*, vol. 36, no. 1, p. 26, 2022.

[44]. K. Miettinen, *Nonlinear Multiobjective Optimization*. Boston: Springer, 1999.

[45]. T. T. Nguyen, "A multi-objective deep reinforcement learning framework," *Eng. Appl. Artif. Intell.*, vol. 96, p. 103915, 2020.

[46]. W. R. Miller and S. Rollnick, *Motivational Interviewing: Helping People Change*, 3rd ed. New York: Guilford Press, 2013.

[47]. A. Esteva et al., "DeepHealth: A deep learning platform for healthcare," *Nat. Med.*, vol. 27, pp. 1735-1743, 2021.

[48]. L. Osterberg and T. Blaschke, "Adherence to medication," *N. Engl. J. Med.*, vol. 353, no. 5, pp. 487-497, 2005.
[49]. L. Ouyang et al., "Training language models to follow instructions with human feedback," in *Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 27730-27744.
[50]. R. Bommasani et al., "On the opportunities and risks of foundation models," arXiv preprint arXiv: 2108.07258, 2021.
[51]. J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge: Cambridge Univ. Press, 2009.