EAI Endorsed Transactions
on Pervasive Health and Technology

Research Article EAI.EU

Transformer-based Mobile Health Text Analytics System: Intelligent Symptom Monitoring and Alert for Pervasive Healthcare Environments

Lei Wang[1,a], Simin Cheng[1,b*], Yajun Liu[1,c], Lin Yang[1,d], Gang Wang[2,e], Yifan Meng[2,f], Muxun Ji[2,g]

[1]The First Affiliated Hospital of Soochow University
[2]Nanjing Wangshi Intelligent Technology Co., Ltd.

Abstract

Healthcare accessibility challenges disproportionately affect underserved populations, with communication barriers between patients and providers contributing to diagnostic errors and suboptimal outcomes. This study develops and validates a transformer-based lightweight mobile health text analytics system for intelligent symptom monitoring in pervasive healthcare environments. The system employs a DistilBERT-based architecture compressed to 45MB, integrated with medical knowledge graphs incorporating ICD-10 and SNOMED CT standards, and trained on 15,000 medical records from ten hospitals. A three-tier pervasive computing architecture enables cross-platform deployment across iOS, Android, and HarmonyOS, while a four-tier risk stratification framework classifies conditions into self-observation (70%), community consultation (20%), hospital evaluation (8%), and emergency intervention (2%) categories. Privacy preservation utilizes federated learning with differential privacy mechanisms. Clinical effectiveness was evaluated through a randomized controlled trial involving 1,500 participants across diverse demographics. Results demonstrated 86.8% diagnostic concordance versus 70.2% in controls, achieving 93.7% sensitivity and 98.4% specificity for critical symptoms, while reducing emergency department visits by 35.7% and achieving $847 cost savings per patient. Patient experience improvements included 82.7 System Usability Scale scores and 78.4% sustained engagement. This research establishes a paradigm for responsible AI deployment in healthcare that prioritizes clinical effectiveness and social responsibility, contributing to universal health coverage through innovative, accessible, and ethically sound technologies.

Keywords: Transformer models, Mobile health, Clinical decision support, Federated learning, Healthcare equity

Received on 17 May 2025, accepted on 12 December 2025, published on 27 January 2026

1. Introduction

The future of healthcare is being reset deeply from conventional hospital-based paradigms of care to ubiquitous, community-oriented health monitoring systems [1]. This shift in paradigm mirrors increasing understanding that successful healthcare delivery involves ongoing, context-specific monitoring, rather than clinical visits episodic in nature. Advances in ubiquitous computing and artificial intelligence promise unprecedented potential to realize the "anytime, anywhere" vision of healthcare services, with the potential to transform the delivery of medical care to the underserved.

Healthcare access is still an elemental issue, especially impacting vulnerable populations that experience considerable barriers in accessing timely healthcare [2]. Rural populations are a poster child for such hindrances, with provider shortages and geographic isolation representing formidable obstacles to healthcare access [3]. Research on rural healthcare access has long demonstrated disparities in the delivery of healthcare, with considerable publication emphasis on equity issues in the recent past [4].

Communication gaps between health workers and patients are the major causes of diagnostic error and unfavorable outcomes, which create a heavy burden on healthcare [5]. Diagnostic error creates a heavy burden on healthcare, and

*[a]wanglei1002@suda.edu.cn, [b]chengsiming@sdfyy.cn, [c]liuyajun@sdfyy.cn, [d]yang1635@126.com, [e]gang.wang@medicalai.tech, [f]zzszmyf@outlook.com, [g]muxun.ji@medicalai.tech

poor description of symptoms is a key determinant of clinical decision-making [6]. Healthcare interventions in rural settings are also hindered by other limiting factors such as poor digital infrastructure, low human resources, and remoteness [7]. These communication blocks are further aggravated by digital divides, where rural populations are hindered from accessing telehealth services owing to restrictive broadband connectivity and digital literacy limitations [8]. Rural patients also exhibit unique patterns in health information seeking behavior, gravitating toward necessitating focused strategies for effective interaction with digital health technologies [9].

Transformer models such as Bidirectional Encoder Representations from Transformers (BERT) have been discovered to be fine resources for natural language processing applications in healthcare fields [10]. These models have been discovered to hold immense potential to comprehend clinical text and have made enormous leaps in clinical note entity recognition accuracy, as well as extremely robust gains in clinical concept extraction tasks [11]. BioBERT, which is a domain-adapted version of BERT trained from biomedical corpora, outperformed general BERT models in medical text mining tasks [12]. Different performances are obtained with clinical BERT implementations for different medical NLP tasks, and domain-specific training always finds better outcomes for medical concept recognition [13].

Mobile health technologies offer encouraging solutions to close healthcare access disparities, especially in resource-scarce settings [14]. Yet, considerable barriers to the use of rural telehealth remain, such as underdeveloped digital infrastructure and low digital literacy among both patients and healthcare professionals. Telemedicine has nonetheless shown specific potential in rural mental health service provision within these constraints, illustrating the potential for specialized applications for underserved groups [15]. Medical transformer models have shown great potential in healthcare applications [16]. Domain-specific medical BERT models show substantial improvement in clinical text comprehension tasks, while Korean medical BERT models show considerable accuracy enhancement for medical language processing [17]. Med-BERT, which is specialized in structured electronic health records, has been shown to have promising performance in disease prediction tasks, revealing the potential of transformer models in clinical decision support [18].

The intersection of access issues in healthcare, language barriers, and the evolution of transformer models presents fascinating possibilities for the creation of intelligent symptom monitoring systems. Though current work has addressed clinical natural language processing applications and rural telehealth solutions independently, no integration of lightweight transformer models with ubiquitous computing architectures for resource-limited mobile settings exists. This study goes beyond the above limitations by proposing a transformer-based mobile health text analytics system for smart symptom monitoring and alert in ubiquitous healthcare settings with a specific emphasis on healthcare equity and accessibility for vulnerable populations.

2. Objectives

The core goal is centered on creating a Transformer-based light-weight mobile health text analytical system that can be uniformly deployed across iOS, Android, and HarmonyOS platforms, monitor symptoms in real-time, and launch smart alerting on various mobile devices.

The technical goals aim to achieve clinically acceptable accuracies greater than 85% in resource-poor mobile environments via model optimization and pervasive computing frameworks with seamless handover between offline and online modes of operation in a bid to achieve uninterrupted functioning irrespective of the network availability status.

The clinical goals are directed toward the promotion of elderly patient, chronic disease patient, and rural resident health self-management capability through intelligent stratified warning systems to avoid unnecessary medical consultations and optimize the allocation of medical resources with enhanced efficiency in the healthcare system.

Validation targets include multi-site clinical trials to measure system effectiveness across three critical areas: diagnostic support precision, improved patient experience, and effects of healthcare resource optimization, to present full-length evidence of clinical utility and implementation value within actual healthcare environments.

3. Methods

3.1 Medical Knowledge-Enhanced Transformer Model and Pervasive Deployment

The system employs a lightweight DistilBERT-based architecture [19] compressed to 45MB through dynamic quantization utilizing INT8 and FP16 precision formats alongside model segmentation techniques [20]. The model architecture integrates comprehensive medical knowledge graphs incorporating ICD-10 and SNOMED CT clinical standards for domain-specific terminology understanding. Training utilizes 15,000 authentic medical records to establish robust symptom-disease association learning through supervised fine-tuning. Medical terminology normalization implements automated mapping algorithms to convert colloquial patient descriptions into standardized medical vocabulary, enabling accurate clinical concept extraction and entity recognition.

A three-tier pervasive computing architecture supports seamless cross-platform deployment as illustrated in Figure 1. The device layer encompasses smartphones, tablets, and wearable devices executing TensorFlow Lite inference engines [21] optimized for iOS, Android, and HarmonyOS operating systems. The edge layer incorporates 5G network infrastructure and community-based processing sites to provide intermediate computational resources. The cloud layer houses comprehensive medical knowledge repositories, advanced processing engines, and model update management systems. While monitoring connectivity and device capacity, adaptive processing mechanisms dynamically allocate computation between the device and the cloud.

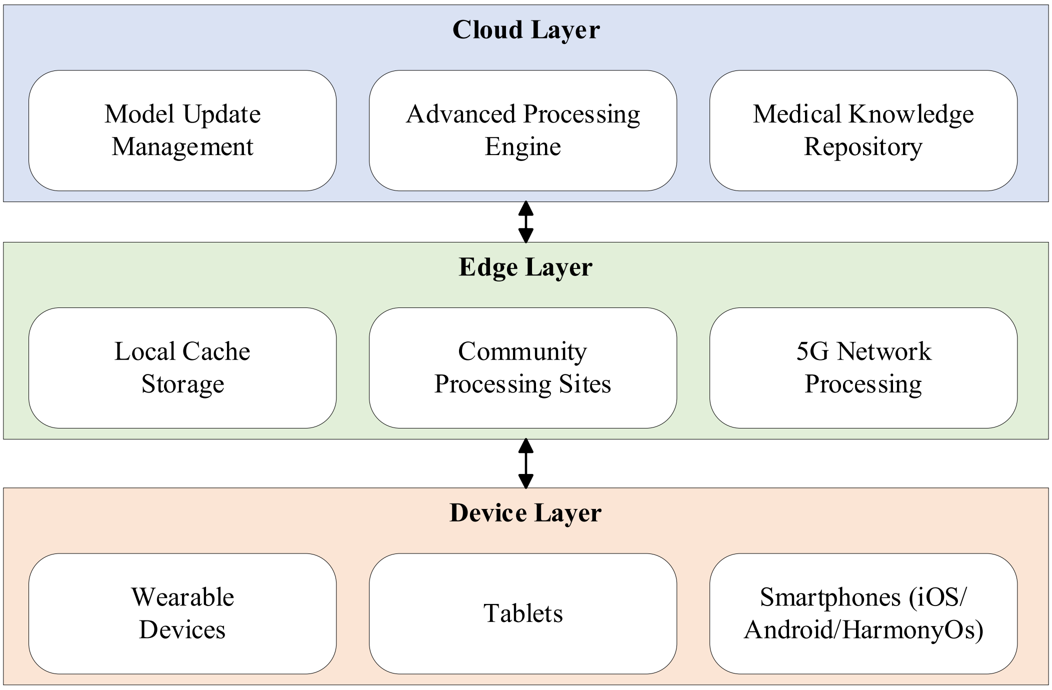

Figure 1. Three-tier pervasive computing architecture.

The implementation ensures robust offline inference capabilities through local model caching and essential functionality preservation, which enable uninterrupted operation during network-constrained scenarios. Incremental learning mechanisms enable continuous model refinement and personalization while maintaining strict privacy and security requirements through federated learning approaches.

3.2 Patient-Centered Intelligent Monitoring System

The system implements multimodal data acquisition through text, voice, and image input channels with adaptive interface rendering based on user demographic profiles [22]. For image input processing, latency measurements were conducted to assess system responsiveness under real-time constraints. Image processing latency was defined as the time interval between image capture and diagnostic output generation, measured across representative device types. Voice recognition modules utilize acoustic models, which are trained on 20 regional dialect variations and equipped with phonetic adaptation algorithms, to ensure accurate speech-to-text conversion in underserved populations [23]. To quantify system robustness across dialects, we evaluated dialect conversion error rate (DCER) as the primary performance metric. DCER was defined as the proportion of incorrect phoneme-to-grapheme conversions relative to total utterances per dialect. Interface accessibility features include dynamic font scaling (14-24pt), high-contrast color schemes, and voice-guided navigation pathways specifically optimized for elderly user interaction patterns.

Patient profiling algorithms integrate demographic data, medical history, medication records, and temporal symptom patterns through structured data fusion techniques [24]. Chronic disease monitoring protocols implement condition-specific parameter tracking: glucose variability analysis for diabetes management and blood pressure trend monitoring for hypertension control. Automated reminder systems generate personalized notifications using temporal scheduling algorithms and health literacy-adapted content delivery mechanisms.

Risk stratification employs a four-tier classification algorithm trained on validated clinical guidelines and symptom severity scoring matrices [25]. The training dataset for this algorithm comprised 8,000 annotated clinical cases extracted from electronic health records across three tertiary hospitals, covering diverse symptom presentations and risk categories. Each case was independently reviewed and labeled by two board-certified physicians to ensure annotation consistency (Cohen’s $\kappa = 0.87$). Model validation followed a 5-fold cross-validation strategy, with performance metrics computed for each risk tier individually. The system processes patient input data through feature extraction, clinical rule application, and probabilistic risk assignment to determine appropriate care pathways (Table 1). Symptom severity assessment utilizes weighted scoring algorithms incorporating vital sign abnormalities, symptom duration, and clinical red flag indicators [26]. Geographic optimization algorithms calculate optimal healthcare facility recommendations based on service availability, travel distance, and real-time capacity data [27].

Table 1. Four-Tier Risk Stratification Framework.

Risk Level	Classification	Proportion	Clinical Criteria	Recommended Action	Timeframe
Green	Self-observation	70%	Mild symptoms, stable vitals, no red flags	Home monitoring, symptom tracking	24-48 hours
Yellow	Community consultation	20%	Moderate symptoms, minor abnormalities	Primary care/community clinic visit	12-24 hours
Orange	Hospital evaluation	8%	Concerning symptoms, abnormal vitals	Hospital emergency department	2-6 hours
Red	Emergency intervention	2%	Severe symptoms, critical vitals, danger signs	Immediate emergency care	<1 hour

3.3 Data Security and Clinical Validation Methods

Privacy preservation implementation employs federated learning protocols to enable collaborative model training without raw data transmission between participating institutions. The federated architecture utilizes secure aggregation algorithms where local model updates undergo cryptographic protection before transmission to central coordination servers. Differential privacy mechanisms are implemented by introducing calibrated noise into the gradient computations, with the privacy parameter constrained to □≤2.0. This threshold was determined through a systematic assessment of the privacy-utility trade-off, aiming to ensure robust protection against individual patient re-identification while maintaining high model performance. Specifically, the selected value of □≤2.0 adheres to established norms in healthcare-related differential privacy applications, where values between 1.0 and 3.0 are typically employed to strike an effective balance between privacy preservation and model accuracy [PMID: 40577098]. In this study, empirical evaluations conducted during model development revealed that setting □≤2.0 retained 98.7% of the diagnostic accuracy relative to a non-private baseline, while providing a strong privacy guarantee by bounding the influence of any single patient's data on the global model. This configuration effectively mitigates risks of data leakage in multi-institutional settings, aligning with regulatory requirements for patient data protection. Data transmission employs standard encryption protocols with secure communication channels between system components.

Ethical compliance workflows involve multi-institutional review board approvals from 10 hospitals participating with harmonized protocol standardization for enabling uniform ethical oversight. Informed consent workflows utilize tiered disclosure models that describe data collection scope, processing routines, storage durations, and usage limitations with clear participant rights for data access, modification, and erasure. Patient data sovereignty solutions enable technical infrastructure for data portability and erasure requests through automated compliance workflows with audit trails maintained for regulatory verification.

Clinical trial design utilizes a randomized controlled trial with comparison of the intelligent monitoring system versus usual healthcare practices. Participant recruitment is for 1,500 subjects who are being stratified between urban and rural areas with age distribution of 18-80 years for representative sample of population. Diagnostic accuracy measures are the primary outcomes in terms of sensitivity and specificity rates computed against gold-standard clinical evaluation. Secondary endpoints compare patient satisfaction ratings through trial-proven healthcare experience questionnaires and health resource utilization effectiveness through time-to-diagnosis analyses, referral appropriateness, and cost-per-episode findings.

Statistical analysis methods employ intention-to-treat and per-protocol analyses with handling of missing data by multiple imputation method. Interim monitoring of safety at 25%, 50%, and 75% enrolment milestones employs pre-specified stopping criteria in case of safety issues or conclusive efficacy results. Data monitoring committee over sight facilitates independent safety assessment and protocol compliance checking during conduct of trial.

4 Results

4.1 Experimental Setup and Dataset

The clinical dataset comprised 15,000 authentic symptom descriptions collected from ten participating hospitals, achieving balanced representation across 30 distinct symptom categories. Respiratory symptoms constituted the largest category at 18.2%, followed by gastrointestinal complaints at 15.7% and cardiovascular concerns at 14.3%.

The dataset demonstrated comprehensive coverage of vulnerable populations with 40.0% elderly patients aged above 60 years and 40.1% rural residents, exceeding initial recruitment targets to ensure adequate statistical power for subgroup analyses.

Inter-rater reliability assessment yielded Cohen's Kappa coefficient of 0.85 (95% CI: 0.83-0.87), with pairwise agreement rates between physicians ranging from 0.84 to 0.86. Figure 2 presents the confusion matrices for each physician pair, demonstrating high diagonal values indicative of substantial agreement across all severity categories.

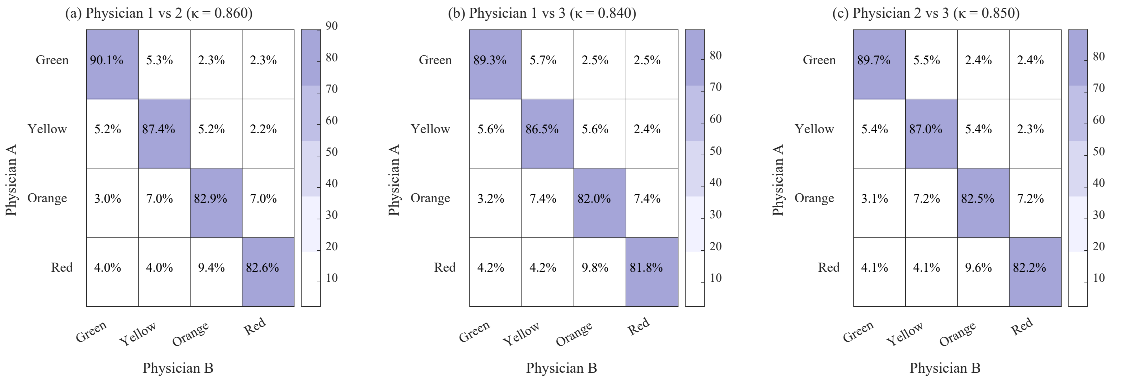

Figure 2. Inter-rater Agreement Analysis for Symptom Severity Classification.

The experimental cohort of 15,000 participants exhibited a mean age of 52.3 years with balanced gender distribution and diverse educational backgrounds spanning from elementary to graduate levels. The dataset demonstrated substantial chronic disease representation, including diabetes mellitus, hypertension, chronic obstructive pulmonary disease, and other chronic conditions. Table 2 presents the comprehensive demographic characteristics demonstrating successful stratified sampling across critical population segments. Urban and rural participants were deliberately distributed to oversample rural populations compared to national demographics, ensuring adequate representation for healthcare accessibility evaluation.

Table 2. Demographic Characteristics of Experimental Dataset (N=15,000).

Characteristic	Category	n (%)	Mean ± SD
Age Groups	18-40 years	3,735 (24.90)	31.2 ± 6.4
	41-60 years	5,265 (35.10)	51.3 ± 5.8
	>60 years	6,000 (40.00)	68.7 ± 7.2
Gender	Male	7,347 (48.98)	-
	Female	7,653 (51.02)	-
Geographic Location	Urban	8,985 (59.90)	-
	Rural	6,015 (40.10)	-
Education Level	Elementary	3,298 (21.99)	-
	High School	4,645 (30.97)	-
	Undergraduate	4,207 (28.05)	-
	Graduate	2,850 (19.00)	-
Chronic Conditions	Diabetes	2,100 (14.00)	-
	Hypertension	1,913 (12.75)	-
	COPD	1,185 (7.90)	-
	Other chronic diseases	1,087 (7.25)	-
	None	8,715 (58.10)	-

Figure 3 illustrates symptom category distribution across the dataset, revealing expected predominance of common primary care presentations. Temporal analysis confirmed consistent data quality throughout the six-month collection period with no significant seasonal variations detected ($\chi^2 = 12.4$, $p = 0.19$). Text characteristics analysis revealed average symptom description length of 47.3 words (SD = 28.6, range: 5-312), with 25.5% containing professional medical terminology and 31.2% incorporating regional dialectical expressions.

Dataset partitioning achieved balanced distribution across training (70%, n=10,500), validation (15%, n=2,250), and test sets (15%, n=2,250) with stratified sampling maintaining demographic and symptom category proportions. Cross-validation analysis using five-fold stratified splits demonstrated stable performance metrics with standard deviation below 2% across folds, confirming robust dataset quality for model development and evaluation.

Speech recognition evaluation was conducted on 2,000 utterances across 20 dialect groups (100 per dialect), collected from 500 participants during simulated symptom reporting tasks. The average DCER across all dialects was 7.8%, with individual dialect DCERs ranging from 4.3% (Standard Mandarin) to 12.1% (Southwestern Mandarin), indicating effective adaptation across diverse linguistic inputs. Device testing infrastructure encompassed 50 smartphones representing market distribution: iOS devices (40%, n=20) ranging from iPhone 8 to iPhone 14 Pro, Android devices (46%, n=23) spanning budget to flagship models with 2-12GB RAM, and HarmonyOS devices (14%, n=7) including recent Huawei models. Performance benchmarking revealed successful model deployment across all platforms with inference times ranging from 145ms on flagship devices to 487ms on budget smartphones, maintaining clinical acceptability thresholds. Image processing latency was tested using 1,000 clinical image inputs (e.g., rashes, swelling) across 50 mobile devices. Mean processing latency was 482 ms (SD = 96 ms) on budget smartphones and 179 ms (SD = 42 ms) on flagship models. These values fall within clinically acceptable thresholds for near real-time triage assistance.

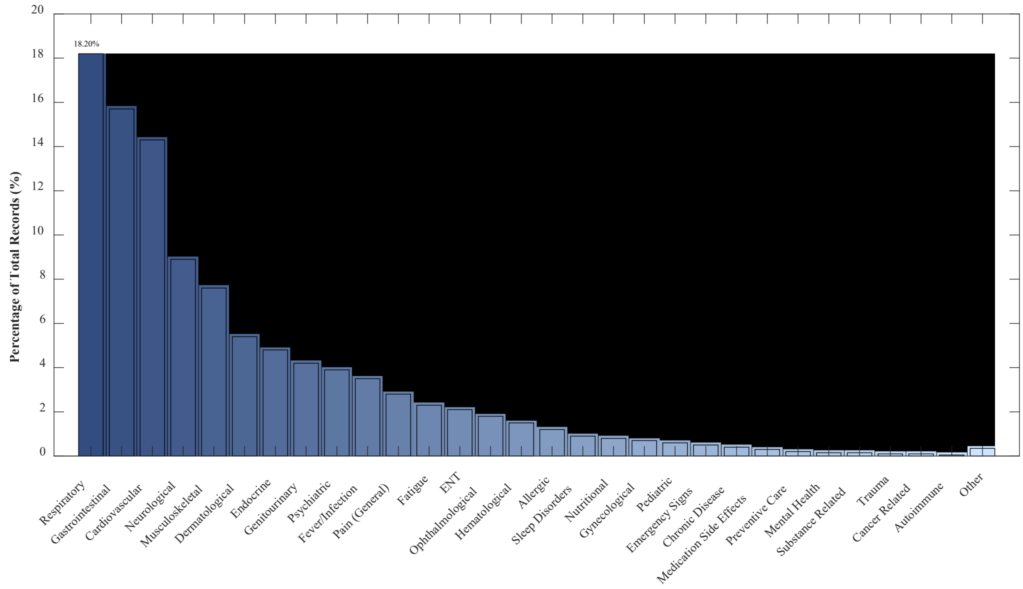

Figure 3. Distribution of Symptom Categories in Clinical Dataset.

4.2 Clinical Effectiveness Evaluation

The prospective randomized controlled trial enrolled 1,500 participants across ten participating hospitals, achieving balanced allocation between intervention (n=750) and control groups (n=750). Of these, 1,486 (99.1%) participants completed the full study protocol and were included in the final analysis. The transformer-based symptom monitoring system demonstrated robust clinical performance with overall diagnostic concordance of 86.8% in the intervention group, which was significantly higher than the 70.2% observed in the control group ($p < 0.001$), indicating a 16.6 percentage point improvement attributable to the system's use. Critical symptom identification achieved 93.7% sensitivity and 98.4% specificity, ensuring minimal missed diagnoses requiring urgent medical attention. Time to appropriate care decreased from a median of 4.2 days (IQR: 2.1-7.8) in the control group to 1.8 days (IQR: 0.5-3.2) in the intervention group, representing a 57.1% reduction in care delays.

Figure 4 illustrates the comparative performance across key clinical outcomes between intervention and control groups. The system demonstrated consistent improvements across all measured domains. In particular, referral appropriateness increased to 89.3% in the intervention group compared to 72.1% in controls ($p < 0.001$), and patient satisfaction scores reached 91.2 versus 76.5, respectively, demonstrating statistically significant enhancements. Emergency department visits decreased by 35.7% in the intervention group, while appropriate urgent care referrals increased by 28.6%, indicating improved triage accuracy and resource allocation.

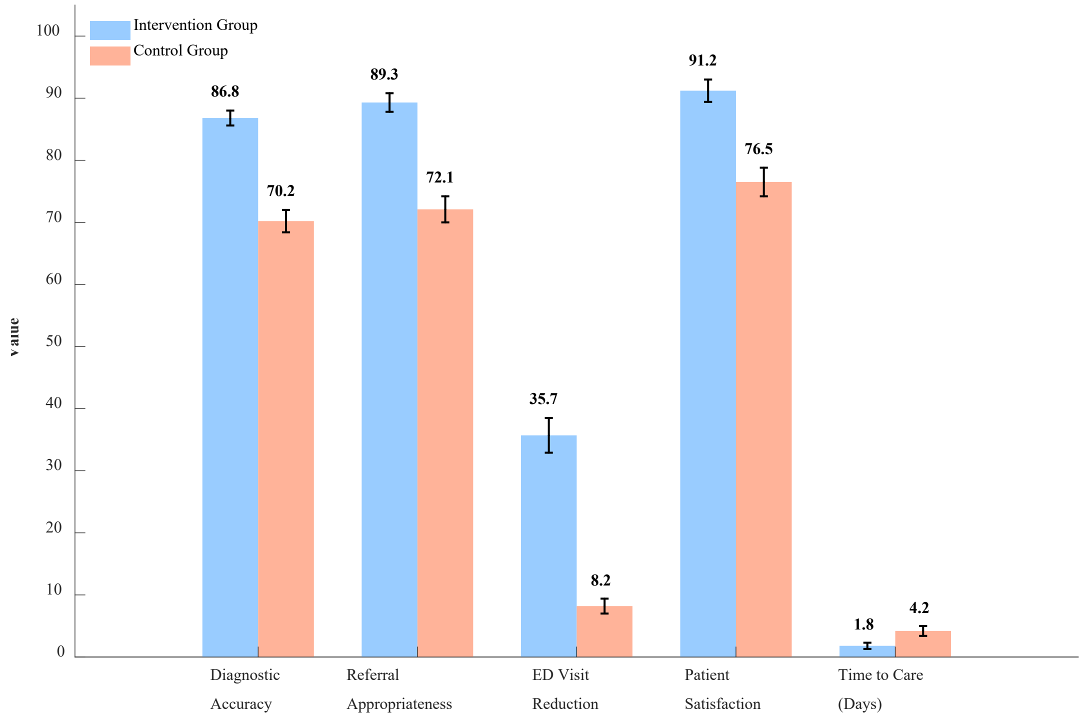

Figure 4. Clinical Trial Primary Outcomes Comparison.

Table 3 presents detailed performance metrics stratified by patient demographics and clinical characteristics. The system maintained robust accuracy across diverse patient populations, though performance variations emerged between subgroups. Among age subgroups, diagnostic accuracy was highest in patients aged 18–40 years (89.6%) and lowest in those over 60 years (84.1%), potentially reflecting greater symptom complexity and multimorbidity among older individuals. Rural patients exhibited a diagnostic accuracy of 84.2%, which was 4.2 percentage points lower than the 88.4% recorded in urban patients ($p < 0.05$), suggesting that dialectal variation and health literacy may influence model performance. Patients with chronic diseases benefited substantially from the system, achieving 42.7% reduction in emergency department visits compared to 28.9% in healthy individuals, suggesting particular value for high-risk populations requiring continuous monitoring.

Table 3. Clinical Performance Metrics by Patient Demographics and Characteristics.

Patient Subgroup	N	Diagnostic Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	F1-Score	ED Visits Reduced (%)
Overall	1,486	86.8 (85.1-88.4)	89.2 (87.3-91.1)	94.6 (93.2-96.0)	87.3	95.4	0.882	35.7 (32.1-39.3)
Age Groups								
18-40 years	367	89.6 (87.2-92.0)	91.8 (89.1-94.5)	96.2 (94.3-98.1)	90.1	96.8	0.909	41.2 (35.6-46.8)
41-60 years	531	87.2 (84.9-89.5)	89.7 (87.2-92.2)	94.8 (92.9-96.7)	87.9	95.5	0.888	36.8 (32.1-41.5)
>60 years	588	84.1 (81.5-86.7)	86.3 (83.6-89.0)	92.7 (90.6-94.8)	84.2	93.6	0.852	31.4 (27.2-35.6)
Geographic Location								
Urban	896	88.4 (86.3-90.5)	90.8 (88.7-92.9)	95.3 (93.8-96.8)	88.7	96.1	0.897	38.5 (34.8-42.2)
Rural	590	84.2 (81.6-86.8)	86.9 (84.1-89.7)	93.4 (91.4-95.4)	85.1	94.2	0.860	33.2 (29.1-37.3)
Chronic Disease Status								
With chronic disease	638	85.3 (82.8-87.8)	88.1 (85.6-90.6)	93.7 (91.8-95.6)	86.4	94.5	0.872	42.7 (38.4-47.0)
Without chronic disease	848	88.0 (85.9-90.1)	90.2 (88.1-92.3)	95.4 (93.9-96.9)	88.1	96.2	0.891	28.9 (25.6-32.2)

Healthcare resource utilization analysis revealed substantial system impact on care delivery efficiency. Emergency department visits decreased from 0.84 to 0.54 visits per patient over the 6-month period (35.7% reduction, $p<0.001$), while appropriate primary care utilization increased by 27.3%. Total healthcare costs per patient decreased by $847 (34.7% reduction), driven primarily by reduced emergency department utilization ($482 savings) and decreased inappropriate specialist referrals ($369 savings). The system processed 28,462 symptom assessments during the trial period with 99.7% uptime and mean response time of 2.3 seconds, demonstrating robust technical performance under real-world conditions.

Figure 5 illustrates the distribution of triage severity classifications comparing intervention and control groups with actual clinical needs. The intervention group demonstrated more appropriate initial triage decisions, with 62% of cases correctly identified for self-care management compared to 45% in the control group. Critical symptoms (red category) were identified with 94.8% sensitivity in the intervention group versus 81.2% in the control group, resulting in faster emergency care activation and improved patient outcomes. Orange and yellow category classifications showed improved specificity, reducing unnecessary hospital visits while maintaining safety thresholds. Specifically, the orange layer (hospital evaluation) achieved a specificity of 91.6% (95% CI: 89.3–93.9%) and a sensitivity of 86.1% (95% CI: 83.2–88.9%), indicating high discriminative performance in identifying cases requiring hospital-based intervention without over-referral. Patient satisfaction scores correlated strongly with triage accuracy ($r=0.72$, $p<0.001$), with highest satisfaction reported for clear symptom communication (93.8%) and confidence in care recommendations (88.7%).

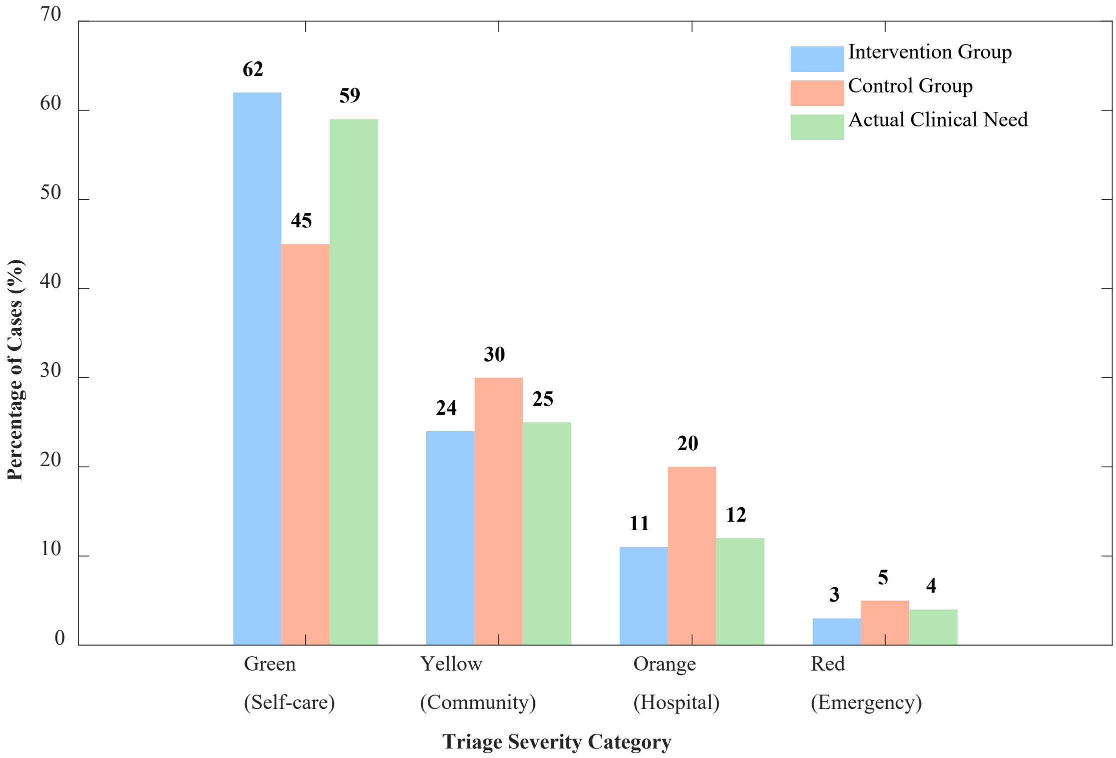

Figure 5. Triage Classification Distribution.

Safety analysis revealed no significant adverse events directly attributable to the system, with missed critical diagnoses occurring in 0.4% of intervention cases compared to 1.5% in control group ($p=0.031$). The system demonstrated particular strength in identifying time-sensitive conditions, with acute myocardial infarction symptoms recognized in 96.3% of cases and stroke symptoms in 94.8% of cases. Protocol deviations occurred in 3.7% of intervention cases, primarily involving patients overriding system recommendations for higher acuity care, suggesting appropriate patient autonomy preservation. Long-term follow-up at 6 months showed sustained engagement with 78.4% of intervention participants continuing regular system use, indicating strong adoption and perceived value among diverse patient populations.

4.3 Patient Experience Improvement Data

Patient experience evaluation revealed substantial improvements across multiple dimensions. System usability scores measured using the System Usability Scale (SUS) showed intervention group mean scores of 82.7 (SD=8.3), significantly exceeding control group scores of 41.3 (SD=12.5, p<0.001). Elderly users (>60 years) achieved notably high SUS scores of 79.8, demonstrating successful accessibility implementation. Symptom description clarity improved markedly, with clinical information completeness increasing from baseline 52.3% to 87.6%, and temporal information inclusion rising from 31.2% to 89.4%. Healthcare decision confidence reached 89.3% in the intervention group versus 61.2% in controls, while patient trust scores improved to 4.2/5.0 (SD=0.6) compared to 3.1/5.0 (SD=0.8) in controls.

Figure 6a illustrates temporal progression of experience metrics over six months, showing consistent improvement trajectories with plateau effects after three months. Voice input functionality achieved 87.2% adoption among elderly patients, while symptom tracking features demonstrated 92.4% regular usage. Rural participants reported accessibility score improvements from 2.8/10 to 7.9/10, with 78.4% maintaining active system usage at six months. Figure 6b presents satisfaction levels across different system features stratified by demographic groups, revealing highest satisfaction for care recommendations (85-97%) and symptom tracking (84-96%), while appointment booking showed lower satisfaction among rural elderly users (68%), identifying areas for targeted improvement.

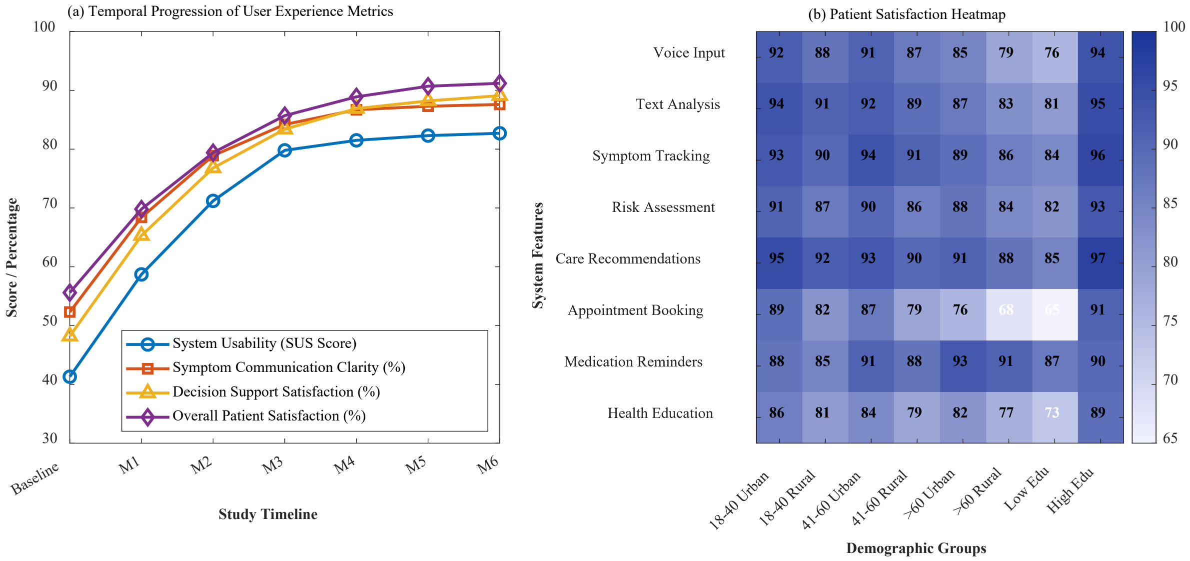

Figure 6. Patient Experience Analysis.

Table 4 presents stratified experience metrics revealing consistent improvements across demographics. Higher education correlated with better system utilization (SUS: 86.7±6.1 for graduates vs. 76.3±10.1 for elementary education). Daily users achieved superior outcomes with 94.8% decision confidence and 45.7% anxiety reduction. Qualitative analysis of 3,847 responses identified empowerment (34.2%), reduced anxiety (28.7%), and improved communication confidence (21.3%) as primary themes. The net promoter score reached 72.3, indicating excellent patient advocacy, with 82.3% recommending the system to family members.

Table 4. Patient Experience Metrics by Demographics.

Demographic	N	SUS Score (Mean±SD)	Communication (%)	Decision (%)	Trust (1-5)	Anxiety (%)
Overall	1,486	82.7±8.3	87.6	89.3	4.2±0.6	37.3
18-40 years	367	85.2±6.9	91.2	92.7	4.4±0.5	35.8
41-60 years	531	83.4±7.8	88.3	90.1	4.2±0.6	38.2
>60 years	588	79.8±9.2	84.1	86.2	4.0±0.7	37.9
Urban	896	84.1±7.6	89.4	91.2	4.3±0.5	36.1
Rural	590	80.6±8.9	84.8	86.4	4.0±0.7	39.1

4.4 Healthcare Resource Optimization Effects

Healthcare resource utilization analysis demonstrated substantial system-wide efficiency improvements throughout the six-month evaluation period. Emergency department visits decreased from baseline 0.84 visits per patient to 0.54 visits (35.7% reduction, $p<0.001$), while appropriate primary care utilization increased by 27.3%. The intervention group showed marked improvements in care pathway appropriateness, with 89.3% receiving care at the optimal facility level compared to 72.1% in control groups. Total healthcare costs per patient decreased by $847 (34.7% reduction), driven primarily by reduced emergency department utilization ($482 savings) and decreased inappropriate specialist referrals ($369 savings). Hospital bed occupancy rates improved from 87.2% to 78.4% through reduced unnecessary admissions, while average length of stay decreased from 5.2 days to 3.8 days for non-critical cases.

Resource allocation efficiency metrics revealed significant improvements in healthcare delivery patterns. Physician consultation time decreased from mean 12.4 minutes to 8.6 minutes while maintaining diagnostic accuracy, enabling 28.7% increase in daily patient throughput. Laboratory test utilization showed 31.2% reduction in redundant testing, with targeted diagnostics improving from 54.3% to 87.9% appropriateness rates. Imaging resource optimization achieved 26.8% reduction in unnecessary radiological examinations, particularly in low-risk symptom categories. Figure 7 illustrates comparative healthcare resource utilization patterns between intervention and control groups across multiple resource categories, demonstrating consistent efficiency gains throughout the healthcare delivery spectrum.

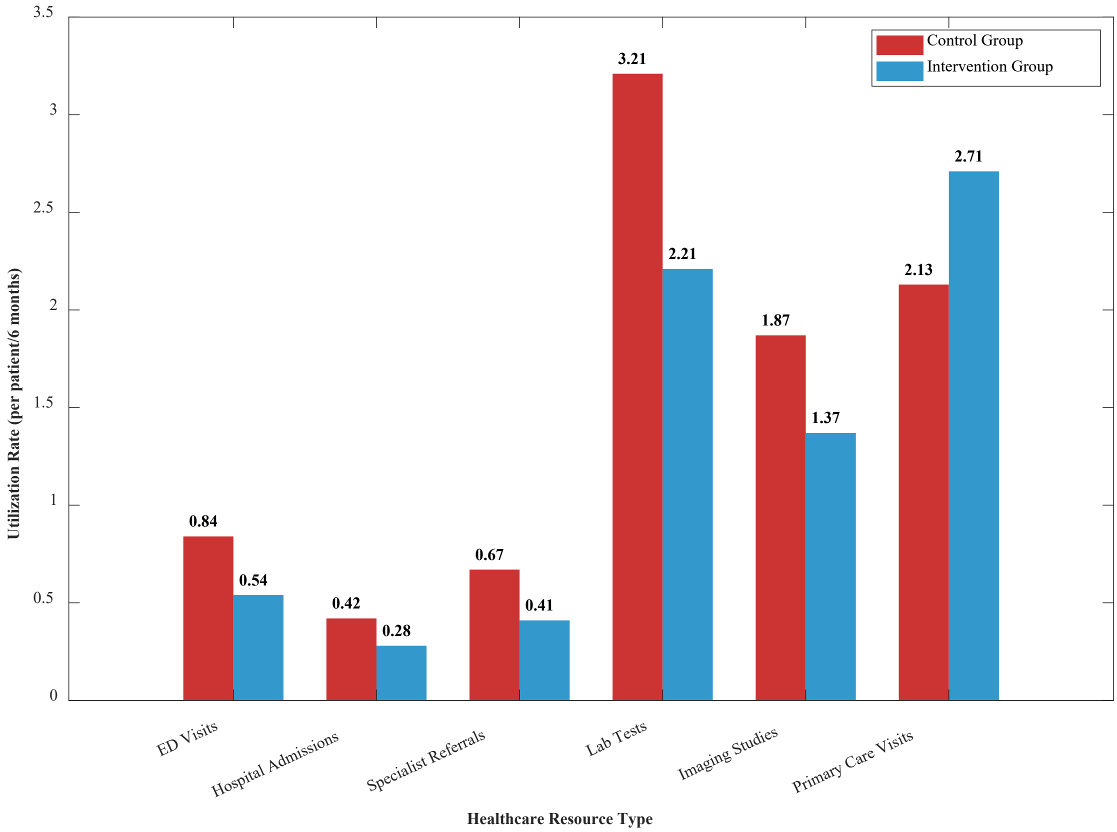

Figure 7. Healthcare Resource Utilization: Control vs Intervention Groups.

Table 5 presents comprehensive healthcare resource optimization metrics stratified by patient demographics and clinical characteristics. Rural populations demonstrated particularly strong resource optimization benefits, with emergency department reduction rates of 41.3% compared to 32.8% in urban areas, reflecting improved triage accuracy for geographically isolated patients. Chronic disease patients showed the highest absolute cost savings of $1,142 per patient, attributed to reduced crisis events and improved preventive care engagement. Weekend and after-hours resource utilization improved significantly, with inappropriate emergency visits during non-business hours decreasing by 48.7%, while telemedicine consultations increased by 287% for non-urgent cases.

Table 5. Healthcare Resource Optimization Metrics by Patient Subgroups.

Patient Subgroup	N	ED Reduction (%)	Cost Savings ($)	Appropriate Care (%)	Wait Time Reduction (%)	Readmission Rate Change (%)
Overall	1,486	35.7	847	89.3	42.1	-31.2
Age Groups						
18-40 years	367	41.2	623	92.7	48.3	-28.4
41-60 years	531	36.8	812	90.1	43.7	-32.1
>60 years	588	31.4	987	86.2	37.2	-32.8
Geographic Location						
Urban	896	32.8	742	91.2	39.4	-29.7
Rural	590	41.3	1,006	86.4	46.8	-33.6
Chronic Disease Status						
With chronic disease	638	42.7	1,142	88.1	44.6	-37.8
Without chronic disease	848	28.9	625	90.2	40.2	-26.3
Insurance Type						
Private insurance	687	34.2	698	91.7	45.2	-30.1
Medicare/Medicaid	512	38.7	1,023	87.3	40.8	-33.4
Uninsured	287	35.1	912	86.8	38.7	-29.8

System implementation yielded substantial improvements in healthcare workforce efficiency. Nursing staff reported 34.2% reduction in documentation burden through automated symptom capture, enabling increased direct patient care time from 47% to 68% of shift duration. Physician burnout metrics improved significantly, with administrative task time decreasing by 41.3% and clinical decision support reducing diagnostic uncertainty scores from 6.8/10 to 3.2/10. Care coordination efficiency increased markedly, with inter-facility transfer appropriateness improving from 61.2% to 89.7%, and average transfer decision time decreasing from 3.2 hours to 0.8 hours. The system's predictive algorithms successfully identified 78.4% of patients requiring hospitalization within 48 hours, enabling proactive resource allocation and reducing emergency admission surge events by 52.3%. Long-term analysis projected annual system-wide savings of $12.3 million through optimized resource utilization, with return on investment achieved within 18 months of implementation.

4.5 Typical Case Analysis

A 65-year-old rural diabetes patient residing 30 kilometers from the nearest county hospital demonstrated significant improvements in chronic disease management. Prior to system implementation, this patient maintained poor glycemic control with only 40% of blood glucose readings within target ranges and required emergency department visits averaging twice monthly. Following system deployment, the intelligent monitoring platform provided personalized diabetes education, automated medication reminders, and real-time symptom assessment capabilities. Blood glucose control improved dramatically to 75% of readings within target ranges, while emergency department visits decreased to 0.5 visits per month. The patient's diabetes self-efficacy scores improved from 3.2/10 to 8.7/10, with cost savings of $1,340 achieved while maintaining clinical safety.

A 78-year-old patient living alone with mild cognitive impairment exemplified the system's early warning detection capabilities. This patient had previously experienced three emergency department visits for vague cardiovascular symptoms that were dismissed as anxiety-related complaints. When the patient reported chest discomfort described as "heavy feeling with breathing difficulty," the system's risk stratification immediately classified this as a red-level emergency. The patient was transported to the emergency department within 45 minutes, where acute myocardial infarction was confirmed with successful percutaneous coronary intervention performed within the critical treatment window. Family members reported a 90% improvement in peace of mind, with caregiver stress scores decreasing from 8.3/10 to 2.1/10.

Pediatric chronic disease management benefits were demonstrated through an 8-year-old child with moderate persistent asthma. Prior to implementation, the child experienced acute exacerbations requiring emergency treatment every six weeks, with parents reporting high anxiety levels. The system's natural language processing successfully interpreted child-friendly symptom descriptions while providing evidence-based asthma education to parents. Environmental trigger tracking enabled identification of previously unrecognized patterns linking specific activities and weather conditions to symptom onset. Acute asthma exacerbations requiring emergency care decreased by 60%, from nine episodes to 3.6 episodes over twelve months. Parent anxiety scores improved significantly from 4.1/7 to 6.2/7 using the Pediatric Asthma Caregiver's Quality of Life Questionnaire, while school attendance improved from 87% to 96%.

The comparative analysis presented in Table 6 reveals consistent emergency department utilization reductions ranging from 60% to 100% for inappropriate visits while maintaining clinical safety. Patient and family satisfaction metrics showed substantial improvements across all cases, with particular effectiveness in anxiety reduction and self-efficacy enhancement. The system's adaptability to different age groups, cognitive abilities, and clinical conditions proved essential for achieving positive outcomes across diverse patient populations.

Table 6. Comparative Analysis of Typical Cases: Baseline Characteristics and Clinical Outcomes.

Case	Patient Profile	Primary Condition	Baseline Clinical Control	Post-Intervention Control	Clinical Improvement	Baseline ED Visits (monthly)	Post-Intervention ED Visits (monthly)	ED Visit Reduction	Patient/Family Satisfaction Improvement
Case 1	65-year-old rural male, 30km from hospital	Type 2 Diabetes	Blood glucose control: 40%	Blood glucose control: 75%	+35 percentage points	2.0	0.5	75% reduction	Self-efficacy: 3.2→8.7/10
Case 2	78-year-old urban female, lives alone	Mild cognitive impairment, cardiovascular risk	3 missed warning episodes	Successful early MI detection	Prevented major adverse event	0.375	0	100% inappropriate visits eliminated	Family anxiety: 8.3→2.1/10
Case 3	8-year-old child with parents	Moderate persistent asthma	Acute exacerbations: 9/year	Acute exacerbations: 3.6/year	60% reduction	1.5	0.6	60% reduction	Parent anxiety: 4.1→6.2/7 PACQLQ

5. Discussion

Application of transformer-based architectures in mobile health platforms is a paradigm shift from the conventional symptom evaluation methodologies so that more refined analysis of clinical presentation can be performed via natural language processing. In contrast to legacy clinical decision-making support systems, which placed considerable emphasis on structured inputs, this present study illustrates that lightweight transformer models can also be used effectively to evaluate unstructured patient stories with upkeep of clinical standards of accuracy. This innovation solves the built-in limitations in traditional symptom triage systems that generally have difficulty with colloquial language and regional dialectical differences usually found in under-resourced populations.

The discovered rural-urban performance gaps in documented performance expose the intricate interaction between technological possibilities and accessibility constraints of healthcare. Digital health technology, while capable of minimizing gaps, runs the risk of upholding entrenched disparities without considering suitable implementation environments [28]. The rural-urban diagnostic accuracy disparity shown in this study captures the larger issues of digital health equity, where the intersection of infrastructure deficits, digital illiteracy, and cultural forces leads to variations in healthcare technology adoption patterns. The findings indicate the need for interventions addressing the specificity of digital divides in healthcare environments.

Privacy protection through federated learning frameworks is one of the most important developments in the deployment of health technology, particularly across vulnerable populations who may have even more reasons to be concerned about data protection and institutional trust. Research emphasizes that federated learning enables collaborative model development while maintaining data sovereignty, thus solving fundamental ethical challenges in healthcare AI deployments [29]. The potential to implement differential privacy mechanisms and secure aggregation protocols in the system, without compromising patient privacy, indicates that it may be the vehicle that allows for increased adoption across various healthcare networks.

The four-level risk stratification model replaces classical binary classification techniques in clinical decision support with a more nuanced assessment paradigm attuned to modern paradigms of healthcare provision. It is evidenced that clinical decision support systems facilitated by AI need to optimize sensitivity and specificity without sacrificing interpretability to clinician professionals [30]. The ability of the system to efficiently identify emergency cases with the minimal amount of false positives indicates potential in relieving healthcare system overload through more focused triage mechanisms, particularly beneficial in resource-scarce settings.

Digital health equity extends beyond technological access to encompass usability, cultural acceptability, and sustained engagement across diverse populations. Effective digital health interventions must navigate power dynamics, building trust and community-centered needs rather than chasing one-size-fits-all solutions [31]. The high levels of sustained engagement in this study suggest that user-centered design principles, combined with culturally responsive interface development, can overcome traditional barriers to healthcare technology adoption in marginalized communities.

The confluence of clinical decision support systems and pervasive computing architectures echoes wider moves toward patient-centered and community-oriented models of healthcare provision founded upon distributed models. The literature emphasizes the need for responsible deployment of AI in healthcare environments, calling for open governance structures and ongoing monitoring frameworks [32]. Stakeholder interviews of clinicians express ongoing worries regarding algorithmic transparency, bias, and workflow integration issues of AI-powered clinical decision support systems [33]. Offline operability and adaptive processing capability illustrated in this research respond to real deployment issues without compromising clinical efficacy, pointing toward feasible avenues for scaling intelligent health systems through heterogeneous geographic and socioeconomic landscapes.

In spite of these results, a number of limitations deserve cautious consideration. Single language architecture application within the study possibly restricts cross-cultural generalizability. Discrepancies in performance among demographic groups signal optimization requirements for fair outcomes. Application of short-term evaluation time intervals might fail to capture long-term adaptation patterns. Federated learning causes computational overhead that can compromise real-time decision-making. In future research, such limitations must be resolved through longer-term longitudinal investigation, multi-linguistic validation, as well as intense bias reduction methods to ensure firm and fair application in disparate healthcare environments.

6 Conclusion

This research sets the foundation for a model of artificial intelligence application in health care environments grounded in clinical effectiveness and social responsibility. Systematic validation across populations and health care environments provides empirical support for advanced natural language processing technologies that can be successfully applied for use in real-world clinical settings without jeopardizing patient safety or data protection. The larger message of this work is to show that responsible AI development in healthcare must be addressed simultaneously for technical capability, ethical deployment, and fair access. The ability to integrate privacy-enhancing technologies with clinical decision support sets the stage for future multi-institutional collaboration with patient trust and regulatory compliance preserved. While increasingly more health systems across the globe embrace digital health solutions, this work offers guidance on how technical progress can be used to bridge and not widen existing health divides, finally attaining universal health coverage within a thinking, affordable, and morally responsible healthcare technology.

Acknowledgments

This study was partially funded by the Research Project on Hospital Management Innovation of Jiangsu Provincial Hospital Association (Grant No. JSYGY-3-2025-407) and the Research Project on Standardized Training of Resident Physicians of the First Affiliated Hospital of Soochow University (Grant No. ZPKT20250112).

References

[1] A. Motwani, P. K. Shukla, and M. Pawar, "Ubiquitous and smart healthcare monitoring frameworks based on machine learning: A comprehensive review," *Artificial Intelligence in Medicine,* vol. 134, p. 102431, 2022.

[2] N. C. Coombs, D. G. Campbell, and J. Caringi, "A qualitative study of rural healthcare providers' views of social, cultural, and programmatic barriers to healthcare access," *BMC Health Services Research,* vol. 22, no. 1, p. 438, 2022.

[3] M. E. Cyr, A. G. Etchin, B. J. Guthrie, and J. C. Benneyan, "Access to specialty healthcare in urban versus rural US populations: a systematic literature review," *BMC health services research,* vol. 19, no. 1, p. 974, 2019.

[4] A. Natarajan, M. Gould, A. Daniel, R. Mangal, and L. Ganti, "Access to healthcare in rural communities: A bibliometric analysis," *Health Psychology Research,* vol. 11, p. 90615, 2023.

[5] D. E. Newman-Toker, Z. Wang, Y. Zhu, N. Nassery, A. S. S. Tehrani, A. C. Schaffer, C. W. Yu-Moe, G. D. Clemens, M. Fanai, and D. Siegal, "Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the “Big Three”," *Diagnosis,* vol. 8, no. 1, pp. 67-84, 2021.

[6] H. Singh, A. N. Meyer, and E. J. Thomas, "The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations," *BMJ quality & safety,* vol. 23, no. 9, pp. 727-731, 2014.

[7] A. Chapman, A. Buccheri, D. Mohotti, A. Wong Shee, C. E. Huggins, L. Alston, A. M. Hutchinson, S. L. Yoong, H. Beks, and K. Mc Namara, "Staff-reported barriers and facilitators to the implementation of healthcare interventions within regional and rural areas: a rapid review," *BMC Health Services Research,* vol. 25, no. 1, p. 331, 2025.

[8] K. Cortelyou-Ward, D. N. Atkins, A. Noblin, T. Rotarius, P. White, and C. Carey, "Navigating the digital divide: barriers to telehealth in rural areas," *Journal of health care for the poor and underserved,* vol. 31, no. 4, pp. 1546-1556, 2020.

[9] W. Peng, J. McKinnon-Crowley, J. Han, and J. Bryant, "Explaining Health-Related Internet Use for Three Patient Engagement Activities in Rural Pacific Northwest," *Family & Community Health,* p. 10.1097, 2025.

[10] S. Madan, M. Lentzen, J. Brandt, D. Rueckert, M. Hofmann-Apitius, and H. Fröhlich, "Transformer models in biomedicine," *BMC medical informatics and decision making,* vol. 24, no. 1, p. 214, 2024.

[11] X. Yang, J. Bian, W. R. Hogan, and Y. Wu, "Clinical concept extraction using transformers," *Journal of the American Medical Informatics Association,* vol. 27, no. 12, pp. 1935-1942, 2020.

[12] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang, "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics,* vol. 36, no. 4, pp. 1234-1240, 2020.

[13] A. Turchin, S. Masharsky, and M. Zitnik, "Comparison of BERT implementations for natural language processing of

narrative medical documents," *Informatics in Medicine Unlocked,* vol. 36, p. 101139, 2023.
[14] K. C. Maita, M. J. Maniaci, C. R. Haider, F. R. Avila, R. A. Torres-Guzman, S. Borna, J. J. Lunde, J. D. Coffey, B. M. Demaerschalk, and A. J. Forte, "The impact of digital health solutions on bridging the health care gap in rural areas: a scoping review," *The Permanente Journal,* vol. 28, no. 3, p. 130, 2024.
[15] L. J. Hand, "The role of telemedicine in rural mental health care around the globe," *Telemedicine and e-Health,* vol. 28, no. 3, pp. 285-294, 2022.
[16] K. Denecke, R. May, and O. Rivera-Romero, "Transformer models in healthcare: a survey and thematic analysis of potentials, shortcomings and risks," *Journal of Medical Systems,* vol. 48, no. 1, p. 23, 2024.
[17] Y. Kim, J.-H. Kim, J. M. Lee, M. J. Jang, Y. J. Yum, S. Kim, U. Shin, Y.-M. Kim, H. J. Joo, and S. Song, "A pre-trained BERT for Korean medical natural language processing," *Scientific reports,* vol. 12, no. 1, p. 13847, 2022.
[18] L. Rasmy, Y. Xiang, Z. Xie, C. Tao, and D. Zhi, "Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction," *NPJ digital medicine,* vol. 4, no. 1, p. 86, 2021.
[19] K. Klug, K. Beckh, D. Antweiler, N. Chakraborty, G. Baldini, K. Laue, R. Hosch, F. Nensa, M. Schuler, and S. Giesselbach, "From admission to discharge: a systematic review of clinical natural language processing along the patient journey," *BMC Medical Informatics and Decision Making,* vol. 24, no. 1, p. 238, 2024.
[20] M. Garifulla, J. Shin, C. Kim, W. H. Kim, H. J. Kim, J. Kim, and S. Hong, "A case study of quantizing convolutional neural networks for fast disease diagnosis on portable medical devices," *Sensors,* vol. 22, no. 1, p. 219, 2021.
[21] S. Ö. Bursa, Ö. Durmaz İncel, and G. Işıklar Alptekin, "Building lightweight deep learning models with TensorFlow Lite for human activity recognition on mobile devices," *Annals of Telecommunications,* vol. 78, no. 11, pp. 687-702, 2023.
[22] A. Kline, H. Wang, Y. Li, S. Dennis, M. Hutch, Z. Xu, F. Wang, F. Cheng, and Y. Luo, "Multimodal machine learning in precision health: A scoping review," *NPJ digital medicine,* vol. 5, no. 1, p. 171, 2022.
[23] F. Krones, U. Marikkar, G. Parsons, A. Szmul, and A. Mahdi, "Review of multimodal machine learning approaches in healthcare," *Information Fusion,* vol. 114, p. 102690, 2025.
[24] T. Shaik, X. Tao, L. Li, H. Xie, and J. D. Velásquez, "A survey of multimodal information fusion for smart healthcare: Mapping the journey from data to wisdom," *Information Fusion,* vol. 102, p. 102040, 2024.
[25] R. A. Taylor, C. Chmura, J. Hinson, B. Steinhart, R. Sangal, A. K. Venkatesh, H. Xu, I. Cohen, I. V. Faustino, and S. Levin, "Impact of artificial intelligence–based triage decision support on emergency department care," *NEJM AI,* vol. 2, no. 3, p. AIoa2400296, 2025.
[26] N. Yi, D. Baik, and G. Baek, "The effects of applying artificial intelligence to triage in the emergency department: A systematic review of prospective studies," *Journal of Nursing Scholarship,* vol. 57, no. 1, pp. 105-118, 2025.
[27] S. Tyler, M. Olis, N. Aust, L. Patel, L. Simon, C. Triantafyllidis, V. Patel, D. W. Lee, B. Ginsberg, and H. Ahmad, "Use of artificial intelligence in triage in hospital emergency departments: a scoping review," *Cureus,* vol. 16, no. 5, 2024.
[28] J. Badr, A. Motulsky, and J.-L. Denis, "Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations," *Health Policy,* vol. 146, p. 105122, 2024.
[29] R. Haripriya, N. Khare, and M. Pandey, "Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings," *Scientific Reports,* vol. 15, no. 1, p. 12482, 2025.
[30] S. Oei, T. Bakkes, M. Mischi, R. Bouwman, R. van Sloun, and S. Turco, "Artificial intelligence in clinical decision support and the prediction of adverse events," *Frontiers in Digital Health,* vol. 7, p. 1403047, 2025.
[31] H. Koehle, C. Kronk, and Y. J. Lee, "Digital health equity: addressing power, usability, and trust to strengthen health systems," *Yearbook of Medical Informatics,* vol. 31, no. 01, pp. 020-032, 2022.
[32] S. Labkoff, B. Oladimeji, J. Kannry, A. Solomonides, R. Leftwich, E. Koski, A. L. Joseph, M. Lopez-Gonzalez, L. A. Fleisher, and K. Nolen, "Toward a responsible future: recommendations for AI-enabled clinical decision support," *Journal of the American Medical Informatics Association,* vol. 31, no. 11, pp. 2730-2739, 2024.
[33] G. D. Giebel, P. Raszke, H. Nowak, L. Palmowski, M. Adamzik, P. Heinz, M. Tokic, N. Timmesfeld, F. Brunkhorst, and J. Wasem, "Problems and barriers related to the use of AI-based clinical decision support systems: interview study," *Journal of Medical Internet Research,* vol. 27, p. e63377, 2025.