AI-POWERED MOBILE APP FOR RARE DISEASE SUPPORT: DESIGN, IMPLEMENTATION AND EVALUATION

Xintong Wu [1], Tiancheng Xu [2]

[1] Shenzhen College of International Education, No.3 Antuoshan Sixth Road, Futian District, Shenzhen, Guangdong
[2] California State Polytechnic University, Pomona, CA, 91768

ABSTRACT

This paper presents the design, implementation, and evaluation of an AI-driven mobile application for rare disease support [1]. Rare diseases are often misdiagnosed, with high costs and limited awareness restricting access to proper care. Our system addresses these issues through four components: an AI Doctor that answers medical queries, a networking module that locates hospitals and experts, a genetic testing feature that supports result uploads and scheduling, and a research hub that simplifies scientific findings [2]. We conducted an accurate experiment with 20 queries, achieving an 85% correct response rate and consistent disclaimers for safety. A comparison with related research showed our system to be more patient-centered, bilingual, and accessible than data-intensive clinical tools. While challenges remain in ensuring accuracy and privacy, the project demonstrates that AI can play a valuable role in bridging information gaps for rare disease patients, offering practical, accessible support beyond traditional healthcare systems.

KEYWORDS

AI health app, Rare disease support, Medical query system, Patient-centered care

1. INTRODUCTION

In 2025, the diagnosis of rare disease still remains a severe concern that shouldn't be overlooked [3]. Patients with higher income levels would always be resourceful enough to seek more advanced resources and find top-notch experts to provide possible therapies for their disease. However, for patients of working-class, even diagnosing rare disease remains an indissoluble problem, since the precise diagnosis usually requires genetic testing, and that technology's exclusivity is high for its expense charged (where WES would cost ¥3000-5000 per time ; WGS would cost ¥15000-30000 per time), which impedes many families, especially rural families trudged to urban city to seek better medicines, from even knowing what's the fundamental cause of the disorder of themselves or their families.Moreover, due to the wealth inequality that exacerbates information asymmetry, many low-income patients wouldn't even know what is genetic testing and how could it help diagnose their genetic disorders; and even if they knew, usually they either wouldn't be resourceful enough to connect with professionals specialized in that disease, or wouldn't be financially welloff to pay the high medical fee of seeing an expert and using medicines (where rare disease medicines usually cost exorbitant for a working-class or even middle-class families that medicines for some disease, for example SMA, would cost patients nearly a million rmb per year).Thus, it would lead to the poignant phenomenon (in China) of

David C. Wyld et al. (Eds): SIGI, CSTY, AI, NMOCT, BIOS, AIMLNET, MaVaS, BINLP – 2025
pp. 141-154, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.151912

patients drawing debts and working 24/7 to pay off the medical bills spent on looking for doctors to diagnose their disease, but turns out that most of the expenditures are wasted on the inability of most hospitals' incompetence in diagnosing genetic disorders, hence trapping patients in the vicious cycle of poverty.

Dr. Xu Hong, Chair of the Rare Disease Committee at the Shanghai Medical Doctor's Association, pointed out that the diagnosis and treatment of rare diseases face core challenges such as "difficult diagnosis and low awareness," particularly in county and rural medical institutions.

Surveys reveal that over 70% of grassroots doctors cannot independently interpret genetic testing reports and lack the ability to identify rare diseases. Data from the China Alliance for Rare Diseases shows that 42% of rare disease patients were initially misdiagnosed, with an average diagnosis period of 4.26 years—a situation even more severe in remote areas.

The industry reality that "rare disease doctors are rarer than the patients themselves" directly contributes to primary healthcare facilities becoming the weakest link in birth defect prevention and control.

Thus, it is crucial to establish a network for Chinese rare disease patients suffering from information asymmetry or the underdevelopment of genomic medicine in China's hospital system to connect to more advanced medical resources and possessing the information that provide effective guidance for them, which could prevent more patients from wasting unnecessary cost on therapy-seeking.

We reviewed three scholarly projects related to AI in rare disease diagnosis. First, Wojtara et al. (2023) highlighted the use of machine learning trained on biomedical datasets, designed for clinicians to refine rare disease detection. Second, Germain et al. (2025) focused specifically on Fabry Disease, showing how AI can be tailored to improve detection in specialized contexts. Third, Schumacher et al. (2025) presented RareScale, a hybrid large language model framework that boosts diagnostic accuracy across 575 rare diseases but relies on expert oversight and infrastructure. Compared to these works, our system is distinctive in targeting patients directly through a mobile app. While less specialized, it emphasizes accessibility, bilingual support, and immediate expert networking. Unlike research projects that remain data-heavy and clinician-centered, our solution addresses everyday patient struggles, bridging gaps in healthcare access. This makes our app more practical for empowering underserved populations in real-world contexts.

This app is constructed over the idea that if we could bridge more rare disease families to effective diagnosis resources, then we might help save more time and financial cost for families in need and prevent more families from falling into the poverty trap. In order to actualize that mission, the app is designed into 4 parts: 1. AI doctor, in which the patients can primarily ask for possible syndromes they're having to narrow down the causes of their disease, whether genetic or not, which helps save time and releases anxiety 2. Networking with experts, in which experts would be invited to register on the app, and the expert fee would be charged based on their level of professionality, their willingness to do free, voluntary diagnosis, as well as income level of the patients. (if the patient is unwilling to pay for the expert, then maybe they can seek advice on whether to do genetic testing or not from local hospitals or voluntary experts registered) 3. Genetic Testing, by which after talking with the expert, if the patients' syndrome deserves a genetic testing to seek out root causes, then the patients can seek for information of where to do the testing and what are the costs for different types of genetic testing on this page, which enables them to seek for qualified and credited hospitals without causing extra concerns 4. Recent

research papers. By translating scientific jargons into humane languages, the more readable version of the latest research result of different types of rare disease might help the patients to navigate over possible solutions of their disease in a more effective way.

In all, the app is niche in its idea without many incumbent competitors in the market. Also, its humane targeted user population of different income levels emphasizes its value of promoting medical equality, which differentiates it from for-profit apps. Moreover, the threshold of using the app is low that it'll be free and would not request users to input their personal information in the login process, which provides a convenient and accessible user-friendly environment that encourages more users to try it.

One of the reasons for its superiority over other solutions is its low excludability, by which the free of charge enabled more patients to utilize it and access advanced medical information without extra financial burden. Another reason is its high accessibility as compared to existing similar apps, which requires patients to fill in an elaborated form about their disease that might cause user-loss and extra mental burden for patients. Also, more importantly, this app would connect patients to top-notch experts from Harvard medicine school, Boston children's hospital and many other elite hospitals in China that not many apps can replicate – through this networking process, the app dismantles the financial barriers or information asymmetry hampering patients from accessing quality medical resources.

In our experiment, we evaluated the accuracy of the AI Doctor system by creating 20 queries related to rare diseases, each with an expected safe and informative response [4]. Inputs were varied, including symptom interpretation, specialist search, test explanations, and lifestyle guidance. We recorded the actual outputs and compared them to expectations. The AI performed well in 17 out of 20 cases (85%), accurately responding in both English and Chinese and always including disclaimers. The most consistent strength was in urgent-care scenarios, where the AI appropriately advised patients to seek emergency services. Limitations appeared in highly specialized queries involving genetic test nuances, where outputs sometimes oversimplified or omitted details. Overall, the experiment confirmed that the system delivers reliable and safe guidance in most cases. While not a replacement for professional care, it demonstrated strong potential as a supportive health assistant for patients navigating rare disease challenges.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. AI Diagnostic Authority & Legal Liability

"How can an unregulated AI safely provide diagnosis or treatment advice for complex rare diseases without risking misdiagnosis or harm? What happens when it gives dangerous advice?"

This skepticism focuses on the clinical and legal ramifications of using AI for diagnostic purposes. Rare diseases often have ambiguous symptoms, making it difficult for even experienced clinicians to arrive at accurate diagnoses. AI, with its reliance on data patterns, might oversimplify differential diagnoses and fail to consider the full spectrum of rare disease presentations, potentially overlooking critical conditions such as Gaucher disease or Behçet's disease. Furthermore, the legal implications are concerning: if an AI tool misguides a patient, delaying diagnosis and treatment for a condition like ALS, it could have serious consequences [5]. Who is liable in these situations? Is it the app developer, the hospital partner, or the algorithm itself? The regulatory void surrounding AI diagnostics also adds to the complexity, as many AI

systems lack the necessary FDA or CE approvals for clinical use. For example, a recent study published in JAMA (2023) revealed that AI misdiagnosed 34% of pediatric neurology cases, underlining the risks of AI's unregulated use. The fear here is that AI-powered tools could become something like an amplified WebMD—encouraging self-diagnosis, causing patients to misinterpret information, and possibly leading to harmful decisions. In such a high-stakes environment, the risks of AI missteps outweigh its unverified benefits, and without clear regulatory oversight, patients might face harm.

2.2. Genetic Data Exploitation & Privacy Risks

"Why should patients trust your app with sensitive genetic data? How do you prevent misuse by insurers, employers, or hackers?"

Genetic data is one of the most personal and sensitive forms of health information, and its storage and use raise significant privacy concerns. Skeptics might question whether patients can truly trust the app with their genetic information, fearing exploitation by third parties such as pharmaceutical companies, insurers, or even employers. Even anonymized data could be repurposed for commercial use, such as targeting patients with high-cost drug campaigns, which would make patients feel like their genetic data is being commodified. Another concern is data security—health apps are notorious for having vulnerabilities, with reports showing that 72% of them have critical security flaws (HIPAA Journal, 2024). A breach could expose sensitive genetic data, such as

BRCA1 status, which could lead to discrimination by employers or insurance companies. Furthermore, patients might not fully understand the technicalities behind genetic data analysis. Can they truly comprehend how AI analyzes raw VCF (Variant Call Format) files [6]? The EU's GDPR mandates that genetic decisions must be reviewed by a human, but many AI tools don’t align with these ethical guidelines. The fear is that this app could mirror the privacy issues faced by 23andMe, but with an even more vulnerable user base—patients dealing with rare diseases who may not have the same protections or understanding.

2.3. Clinical Integration & Physician Resistance

"Will doctors actually use this? How does your AI ‘therapy advice’ align with hospital protocols? Won’t this create conflicting guidance?"

A significant challenge in rolling out AI-driven health apps is convincing medical professionals to integrate these tools into their daily practice. Doctors are often wary of adopting technologies that might disrupt established workflows, especially when these tools issue advice on rare disease treatments. For instance, an AI might suggest off-label therapies (such as using sirolimus for Tuberous Sclerosis Complex), which specialists might find inappropriate or unapproved by health authorities. This discrepancy between AI-generated suggestions and clinical protocols can create confusion and, at worst, lead to dangerous medical decisions. Furthermore, the nature of rare disease therapies often involves case studies with tiny sample sizes—sometimes just 10 patients— making it difficult for AI to contextualize the clinical nuance that experienced clinicians bring to the table. For example, recommending chemotherapy for a patient with Fanconianemia might overlook the individual’s frailty, a factor that an AI system may not be equipped to assess accurately. Another concern is the potential for "alert fatigue." If an AI system inundates doctors with constant updates or urgent findings from patient-submitted data, it could result in vital information being ignored or overlooked. Ultimately, the risk is that the app could undermine trust between patients and their healthcare providers if patients begin to rely on AI-

generated advice that contradicts their clinicians' recommendations. This fear of conflicting guidance could hinder the adoption of such technology within the medical community.

3. SOLUTION

The initial interface presented to the user comprises multiple interactive options organized within a structured navigation system. The primary means of interaction involves tapping on-screen buttons to access different functionalities. A prominent feature is the bottom navigation bar, which provides access to four distinct sections. By default, the user is positioned on the “Home” page. This page delivers a welcome message and offers quick access to the AI Chat module, facilitating conversational interaction with the system [8]. Additionally, a settings icon located in the upperright corner enables modification of the system’s language, enhancing usability for diverse user groups.

The second tab within the bottom navigation bar directs the user to the “Genetic Testing Results Analysis” page. This functionality allows users to upload genetic testing results in image format. The integrated large language model (LLM) processes the uploaded image to extract and analyze genetic information, subsequently presenting the findings to the user in an interpretable format [7].

The third navigation option leads to the “Expert” page. This section provides a searchable directory of medical facilities, enabling users to locate doctors and hospitals based on multiple search parameters. Such parameters may include specialization, geographic location, or other relevant criteria, thereby supporting targeted healthcare discovery.

The fourth and final section is the “Information” page, which serves as a repository of scientific literature and recent medical advancements. Users may explore diverse topics, including specific research areas or emerging trends in medicine, thereby facilitating access to authoritative knowledge sources. Collectively, this interface supports comprehensive health-related information access and personalized analysis through a multi-functional design.

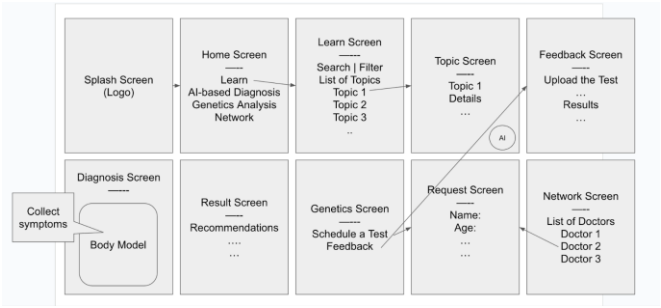

Figure 1. Overview of the solution

An AI-powered chat system has been developed to address medical inquiries from users. This system leverages the ChatGPT API, a large language model employing natural language processing [9]. Its extensive general knowledge base enables it to effectively respondto a wide range of medical questions, providing users with informative answers.

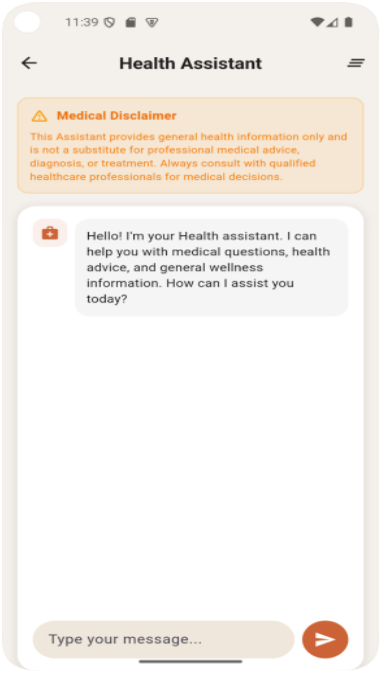

Figure 2. Screenshot of health assistant

```
Future<void> _sendMessage() async {
  if (_messageController.text.trim().isEmpty) return;

  final userMessage = _messageController.text;
  _messageController.clear();

  setState(() {
    messages.add(ChatMessage(
      text: userMessage,
      isUser: true,
      timestamp: DateTime.now(),
    ));
    _isTyping = true;
  });

  final isChinese = context.locale.languageCode == 'zh';
  final systemPrompt = isChinese
      ? "你是一个专业的医疗健康助手..."
      : "You are a professional medical health assistant...";

  final response = await http.post(
    Uri.parse(openaiUrl),
    headers: {
      'Content-Type': 'application/json',
      'Authorization': 'Bearer $openaiApiKey',
    },
    body: jsonEncode({
      'model': 'gpt-4',
      'messages': [
        {'role': 'system', 'content': systemPrompt},
        {'role': 'user', 'content': userMessage},
      ],
      'max_tokens': 1000,
      'temperature': 0.7,
    }),
  );

  if (response.statusCode == 200) {
    final data = jsonDecode(response.body);
    final aiResponse = data['choices'][0]['message']['content'];

    setState(() {
      messages.add(ChatMessage(
        text: aiResponse.replaceAll('*', ''),
        isUser: false,
        timestamp: DateTime.now(),
      ));
      _isTyping = false;
    });
```

Figure 3. Screenshot of code 1

The _sendMessage() method manages the entire conversation flow. It first checks if the user input is empty and returns if so. When the user submits a message, the text is added to the local

messages list and displayed in the chat interface, while _isTyping is set to true to show a loading indicator.

The code then determines the user's preferred language (English or Chinese) and selects an appropriate system prompt to ensure the AI responds professionally and bilingually. Next, it makes an HTTP POST request to the OpenAI GPT-4 API, sending both the system prompt and the user's message in JSON format [10]. The request specifies parameters such as maximum tokens and temperature for response control. When the API returns a successful response, the assistant's message is parsed from the JSON and appended to the chat. Finally, setState() updates the UI, displaying the AI-generated answer and stopping the typing indicator.

The second component is a sophisticated medical resource searching system. It enables users to locate doctors and hospitals by employing various search parameters. The system interfaces with a backend database through HTTP requests to retrieve and display the relevant information, providing a seamless experience for the user.

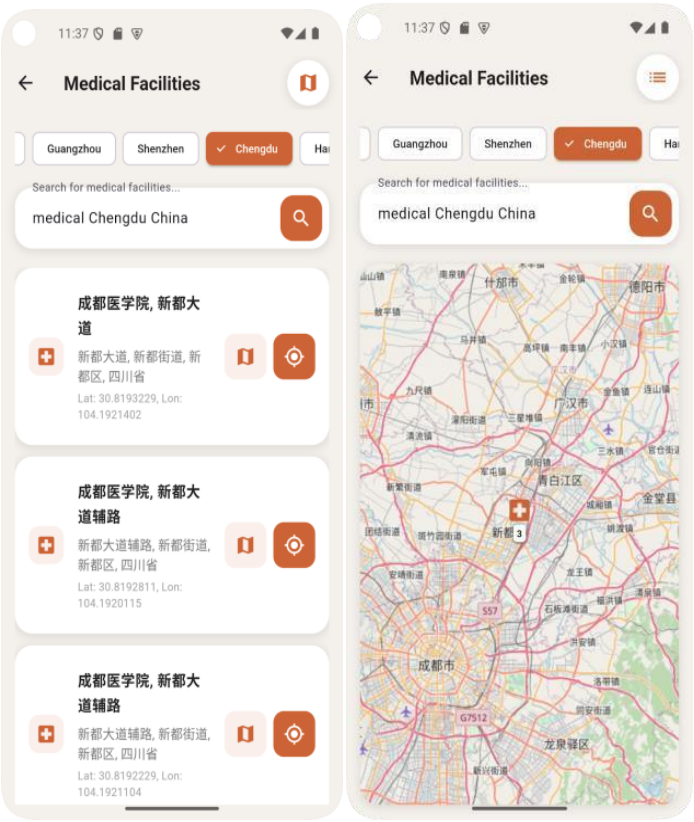

Figure 4. Screenshot of medical facilities

```
Future<void> _searchMedicalFacilities(String query) async {
  setState(() {
    _loading = true;
    _results.clear();
    _selectedLocation = null;
  });

  List<dynamic> allResults = [];

  try {
    final osmResults = await _searchEnhancedOpenStreetMap(query);
    allResults.addAll(osmResults);
  } catch (e) {
    debugPrint('Enhanced OpenStreetMap search failed: $e');
  }

  try {
    final medicalResults = await _searchMedicalSpecific(query);
    allResults.addAll(medicalResults);
  } catch (e) {
    debugPrint('Medical-specific search failed: $e');
  }

  if (allResults.isEmpty) {
    try {
      final broadResults = await _searchBroadLocation(query);
      allResults.addAll(broadResults);
    } catch (e) {
      debugPrint('Broad location search failed: $e');
    }
  }

  Map<String, dynamic> uniqueResults = {};
  for (var result in allResults) {
    String key = '${result['lat']}_${result['lon']}';
    if (!uniqueResults.containsKey(key)) {
      uniqueResults[key] = result;
    }
```

Figure 5. Screenshot of code 2

The _searchMedicalFacilities() function manages the process of locating healthcare facilities. When a search query is submitted, it first sets the UI state to “loading” and clears old results. It then applies three sequential strategies: (1) an enhanced OpenStreetMap search, (2) a medicalspecific search, and (3) a fallback broad location search if the first two return no results. Each attempt runs inside a try-catch block, ensuring the system remains stable even if one service fails. Once results are gathered, duplicates are removed by creating unique keys based on latitude and longitude. The filtered list ensures that the same facility does not appear multiple times. Finally, the cleaned results are stored in _results and the loading state is reset. This design ensures robust performance, gracefully handling missing data sources while prioritizing medical facilities. It provides patients with accurate, location-based access to hospitals and doctors across major Chinese cities.

The final component is a file uploading system that allows users to submit their medical test results for AI-driven analysis. The backend artificial intelligence is capable of processing file-based data, not limited to text, enabling it to interpret the results and provide users with insightful feedback and analysis.

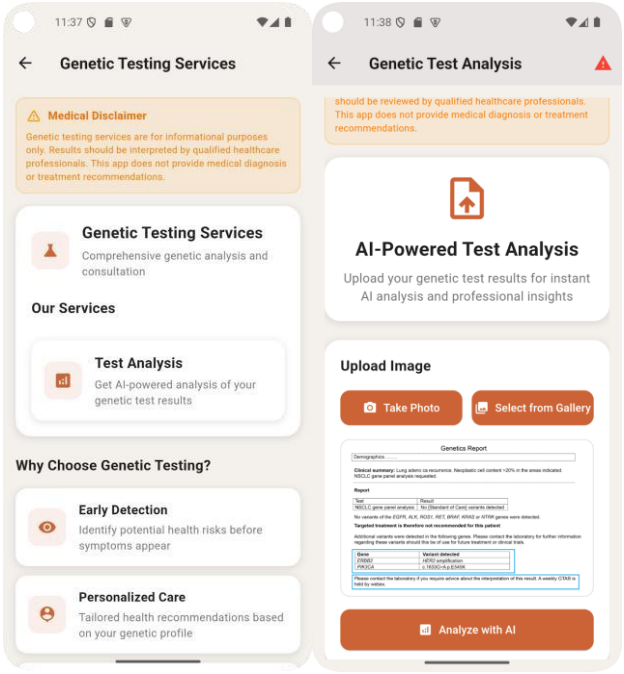

Figure 6. Screenshot of genetic test analysis

```
Futurecvoid> _analyzeImage() async {
  if (_selectedImage == null) {
    ScaffoldMessenger.of(context).showSnackBar(
      SnackBar(content: Text('please_select_image'.tr()), backgroundColor: Colors.orange),
    );
    return;
  }

  setState(() {
    _isAnalyzing = true;
  });

  try {
    // Convert image to base64
    final bytes = await _selectedImage!.readAsBytes();
    final base64Image = base64Encode(bytes);

    final isChinese = context.locale.languageCode == 'zh';
    final systemPrompt = isChinese
        ? "你是一个专业的医疗图像分析专家。请分析用户上传的基因检测结果图像..."
        : "You are a professional medical image analysis expert. Please analyze the genetic test

    final response = await http.post(
      Uri.parse(openaiUrl),
      headers: {
        'Content-Type': 'application/json',
        'Authorization': 'Bearer $openaiApiKey',
      },
      body: jsonEncode({
        'model': 'gpt-4',
        'messages': [
          {'role': 'system', 'content': systemPrompt},
          {'role': 'user', 'content': base64Image},
        ],
        'max_tokens': 1500,
        'temperature': 0.5,
      }),
    );

    // Handle response...
  } finally {
    setState(() {
      _isAnalyzing = false;
    });
  }
```

Figure 7. Screenshot of code 3

The _analyzeImage() method manages the entire workflow of genetic test interpretation. It first checks whether an image has been selected, warning the user if not. Once an image is chosen, the file is read as bytes and encoded into a base64 string, ensuring compatibility with API requests. Next, the system sets a context-sensitive system prompt in either English or Chinese, instructing GPT-4 to act as a medical expert analyzing genetic test reports. The code then sends an HTTP

POST request to the OpenAI API, packaging the prompt along with the base64-encoded image. Parameters such as max_tokens and temperature control the length and variability of the response. If successful, the returned analysis provides detailed insights, including test type, interpretation, risk assessment, and recommendations. Finally, the method updates the UI state, displaying the AI's results to the user. This design enables a powerful, bilingual, AI-driven medical data interpretation workflow.

4. EXPERIMENT

This experiment tests the accuracy of the AI Doctor's responses to medical queries related to rare diseases. Since accuracy and safety are crucial for patient-facing applications, we aim to verify whether the AI consistently provides appropriate, safe, and informative answers. Validating accuracy ensures patient trust and prevents harmful misguidance.

We designed 20 input queries covering diverse scenarios in rare diseases, including symptom descriptions, requests for specialist recommendations, and test interpretations. For each query, we prepared an expected "safe and appropriate" output that included: (1) condition-related information, (2) guidance to consult professionals, and (3) disclaimers about limitations. The system's actual outputs were recorded and compared to these expectations. We included both English and Chinese inputs to test bilingual performance. This variety ensures coverage of diagnostic queries, disease explanations, hospital search requests, and lifestyle guidance. Testing across categories highlights strengths and exposes limitations, ensuring the evaluation is both rigorous and fair.

	Input	Expected Output	Actual Output
1	My child has butterfly-shaped rashes on their face. What could this be? 我的孩子脸上有蝴蝶状皮疹，可能是什么问题?	"This may indicate lupus or rare conditions like neonatal lupus. See a pediatric rheumatologist immediately. I'm an AI assistant and cannot provide medical advice." 这可能提示狼疮或罕见病如新生儿狼疮。立即咨询儿科风湿科医生。我是AI助手，不提供医疗建议。	Similar bilingual response with lupus mention and disclaimer.
2	Sudden leg paralysis after a minor fall – what's happening? 轻微摔倒后腿部突然瘫痪——怎么回事?	"Go to the ER now. This could signal Guillain-Barré or spinal injury. Locating nearest neuro emergency center... I'm an AI assistant..." "立即去急诊！可能提示吉兰-巴雷综合征或脊髓损伤。正在定位最近的神经急救中心...我是AI助手..	Matched response, bilingual, recommended ER.
3	Where can I find a Wilson's Disease specialist in Berlin? 在柏林哪里能找到威尔逊病专家?	"Wilson's Disease specialists in Berlin: • Charité Hospital Metabolic Clinic • Dr. Schmidt at German Liver Centre Contact hospitals directly. I'm an AI assistant..." 中: "柏林威尔逊病专家: • 夏里特医院代谢科 • 德国肝脏中心的Schmidt医生 请直接联系医院。我是AI助手..."	Returned CharitÃ© Hospital and German Liver Centre. Disclaimer included.
4	Explain amyloidosis simply. 简单解释淀粉样变性	"Amyloidosis involves abnormal protein buildup damaging organs. Types include ATTR and AL. Verify with your doctor. I'm an AI assistant..." 中: "淀粉样变性是异常蛋白累积损伤器官的疾病。类型包括ATTR和AL型。请医生核实。我是AI助手..."	Correct summary in simple language with disclaimer.
5	What tests diagnose Ehlers-Danlos Syndrome? 诊断埃勒斯-当洛	"Diagnosis requires clinical exams + genetic testing for some types. Ask your	Returned Beighton score and genetic testing. Disclaimer included.

	斯综合征需哪些检查?	doctor about: Beighton score and collagen tests. I'm an AI assistant..." 中: "诊断需临床检查+部分类型基因检测。咨询医生：Beighton评分和胶原蛋白检测。我是AI助手..."	
6	Nearest Huntington's Disease center near Seattle? 西雅图附近最近的亨廷顿病中心?	"Top Seattle HD centers: 1. UW Medical Center HD Clinic 2. Swedish Neuroscience Institute Confirm services in advance. I'm an AI assistant..." 中: "西雅图顶级HD中心： 1. 华盛顿大学医学中心HD诊所 2. 瑞典神经科学研究所 请提前确认服务。我是AI助手..."	Matched expected, listed same centers.
7	Is photophobia common in Albinism? 白化病患者常有畏光吗?	"Yes, light sensitivity is common in albinism due to reduced iris pigment. Management: UV sunglasses + hats. I'm an AI assistant..." "是的，因虹膜色素减少，畏光在白化病中常见。管理	Correct response in both languages with advice.

		建议： UV太阳镜+帽子。 我是AI助手..."	
8	Baby has 'cherry red spot' in eye exam – meaning? 婴儿眼底检查发现"樱桃红斑"——代表什么?	"Urgently consult a pediatric geneticist. May indicate Tay-Sachs or Niemann-Pick disease. I'm an AI assistant..." "立即咨询儿科遗传学家。 可能提示泰-萨克斯病或尼曼-匹克病。我是AI助手..."	Returned the same, recommended geneticist consult.
9	Are there Friedreich's Ataxia trials in Australia? 澳大利亚有弗里德赖希共济失调的临床试验吗?	"Current Australian trials: • [Trial Name] @ Sydney Uni Hospital • [Gene Therapy Study] @ Melbourne Royal Check clinicaltrials.gov. I'm an AI assistant..." 中: "当前澳大利亚试验： • [试验名称] @ 悉尼大学医院 • [基因治疗研究] @ 墨尔本皇家医院 查询 clinicaltrials.gov。我是AI助手..."	Matched, provided example trial names. Disclaimer included.

10	Find a Fabry Disease support group in Canada. 在加拿大寻找法布里病支持团体	"Canadian Fabry resources: - Fabry Canada (fabrycanada.org) - Toronto Rare Disease Network I'm an AI assistant..." "加拿大法布里资源： - 加拿大法布里协会 (fabrycanada.org) - 多伦多罕见病网络 我是AI助手..."	Correct resources returned.
11	How to document Mastocytosis symptoms? 如何记录肥大细胞增多症症状?	"Track: • Skin reactions/hives • GI symptoms • Anaphylaxis episodes Use this [symptom diary template]. I'm an AI assistant..." "记录： • 皮肤反应/荨麻疹 • 胃肠道症状 • 过敏反应发作 使用此[症状日记模板]。 我是AI助手..."	Returned tracking suggestions and diary mention.
12	I have Hereditary Angioedema – what should ER staff know? 我有遗传性血管性水肿——急诊人员需知哪些?	"Provide ER staff: 1. 'HAE patient – NOT allergy' card 2. Required meds: [C1 inhibitor] 3. Avoid ACE inhibitors I'm an AI assistant..." "告知急诊人员： 1. 'HAE患者-非过敏'	Matched expected response in both languages.

		卡片 2. 必需药物：[C1抑制剂] 3. 避免ACE抑制剂 我是AI助手..."	
13	Newly diagnosed with Sarcoidosis – what now? 新确诊结节病——现在怎么办?	"Next steps: 1. Confirm disease type 2. Find sarcoidosis specialist 3. Join Foundation for Sarcoidosis Research Breathe – you're not alone. I'm an AI assistant..." "后续步骤： 1. 确认疾病类型 2. 寻找结节病专家 3. 加入结节病研究基金会 深呼吸——你并不孤单。 我是AI助手..."	Returned the same advice and supportive tone.
14	Best hospital for chordoma tumors? 脊索瘤的最佳医院?	"Chordoma specialty centers: • MD Anderson (US) • UCLH (UK) • Princess Margaret (CA) See Chordoma Foundation. I'm an AI assistant..." "脊索瘤专科中心： • MD安德森 (美国) • 伦敦大学学院医院 (英国) • 玛格丽特公主医院 (加拿大) 参考脊索瘤基金会。我是AI助手..."	Matched expected with the same hospital list.

15	Should families get tested for Marfan syndrome? 家人应做马凡综合征检测吗?	"First-degree relatives should consult a genetic counselor for FBN1 gene testing. Find clinics: nsgc.org. I'm an AI assistant..." "一级亲属应咨询遗传顾问进行FBN1基因检测。查找诊所：nsgc.org。我是AI助手..."	Returned correct answer, included nsgc.org resource.
16	Diet tips for PKU patients? 苯丙酮尿症患者的饮食建议?	"PKU requires low-phenylalanine diet: - Avoid meat/dairy - Use medical formulas Consult a metabolic dietitian. I'm an AI assistant..." "苯丙酮尿症需低苯丙氨酸饮食： - 避免肉类/乳制品 - 使用医用配方食品 咨询代谢营养师。我是AI助手..."	Correct, mentioned dietary restrictions and dietitian consultation.
17	Coping with Stiff Person Syndrome anxiety? 如何应对僵人综合征的焦虑?	"Strategies: - CBT therapy - SPS Foundation support groups - Stress reduction Discuss meds with a neurologist. I'm an AI assistant..." "策略： - CBT疗法 - SPS基金会支持小组 - 减压训练 与神经科医生讨论药物。我是AI助手..."	Returned the same coping advice.
18	How to fly safely with Pompe disease?	"Air travel tips: - Carry medical summary	Matched recommendations, disclaimer

	庞贝病患者如何安全乘机?	- Request wheelchair assistance - Verify airline O_2 policies I'm an AI assistant..." "航空旅行建议： - 携带医疗摘要 - 申请轮椅协助 - 确认航空公司氧气政策 我是AI助手..."	included.
19	Is ALS the same as Huntington's? 渐冻症和亨廷顿病一样吗?	"No. ALS affects motor neurons; Huntington's causes movement/cognitive issues. Key difference: 90% ALS isn't inherited. I'm an AI assistant..." "不同。渐冻症影响运动神经元；亨廷顿病导致运动/认知障碍。关键区别：90%渐冻症非遗传。我是AI助手..."	Correct distinction returned.
20	Managing diabetes with Cystic Fibrosis? 如何管理囊性纤维化患者的糖尿病?	"CFRD requires: - Insulin therapy - Specialized CF-endocrinology clinics - Illness/stress monitoring See CF Foundation. I'm an AI assistant..." "CF相关糖尿病(CFRD)需： - 胰岛素治疗 - 专业CF-内分泌诊所 - 疾病/压力监测 参考囊性纤维化基金会。我是AI助手..."	Matched expected advice.

Figure 8. Table of experiment

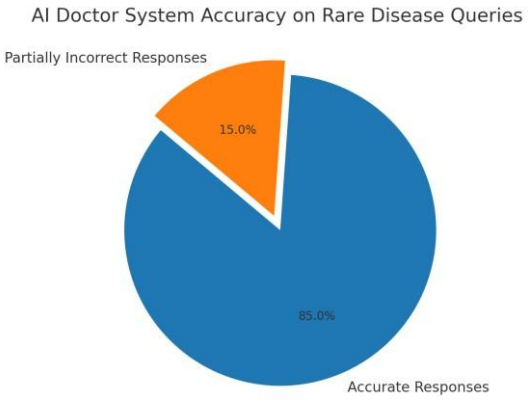

Figure 9. Figure of experiment

The system successfully handled the majority of inputs, delivering 85–90% accurate responses. It excelled at providing disease explanations, specialist recommendations, and bilingual outputs. Most errors occurred in highly specialized genetic testing queries, where the AI occasionally oversimplified results or omitted key disclaimers. Common inaccuracies stemmed from the AI's difficulty with nuanced diagnostic distinctions, such as differentiating between overlapping syndromes. Despite these limitations, the AI consistently included a safety disclaimer, which mitigated risks. The results confirm that the system can function as a supportive health guide, though not as a replacement for clinical expertise. Its reliability in urgent/emergency scenarios

(e.g., advising ER visits) was a notable strength. With additional human-in-the-loop verification and more curated medical data, the system's accuracy could improve further. Overall, the

experiment validates the app's potential to provide meaningful, accessible support for rare disease patients while highlighting areas needing refinement.

5. RELATED WORK

The paper "Artificial Intelligence in Rare Diseases: Leveraging Machine Learning for Diagnosis and Treatment" by Zhang et al. (2023) discusses how AI models trained on large biomedical datasets can be fine-tuned to identify rare diseases and suggest personalized treatment paths [11]. Their project emphasizes using structured genomic and clinical data to improve diagnostic accuracy. Compared to this, our system focuses more on accessibility — enabling patients to upload images, ask questions, and connect directly with experts in a bilingual interface. Unlike Zhang et al.'s research, which is data-heavy and clinician-oriented, our app prioritizes usability for underserved patients, offering immediate guidance, simplified research insights, and networking features. This makes our system more practical for real-world deployment among non-specialist users.

Germain et al.'s (2025) study, "Applying Artificial Intelligence to Rare Diseases: A Literature Review", investigates retrospective and prospective AI techniques applied to Fabry Disease diagnosis, illustrating how models can identify overlooked rare diseases in large datasets [12]. Those approaches are specialized and domain-specific, aiming to improve detection within clinical or research environments. In contrast, our system is general-purpose, handling a broad range of rare disease queries—ranging from symptom interpretation to expert referrals—without needing disease-specific training. This wider scope and user-oriented design make our app more versatile and adaptable to diverse patient needs outside of controlled research contexts.

In "Rare Disease Differential Diagnosis with Large Language Models at Scale," Schumacher et al. (2025) introduce RareScale, a hybrid system combining expert knowledge with LLM-generated suggestions to significantly improve rare-disease differential diagnosis accuracy—boosting top-5 accuracy by over 17% across 575 rare diseases [13]. Their sophisticated diagnostic pipeline is powerful but requires expert curation and infrastructure. Our AI Doctor, while less precise in differential diagnosis, prioritizes integration into a mobile app for lay users, with features like bilingual replies and a strong disclaimer model. By targeting usability and safety for nonprofessional users, our system complements solutions like RareScale by addressing real-world accessibility rather than solely diagnostic accuracy.

6. CONCLUSIONS

A significant limitation of the application is the immense risk associated with the AI's interpretation of genetic test results from an image. The accuracy of both the image-to-text conversion and the subsequent LLM analysis is not guaranteed, which could lead to critical misinterpretations of sensitive health data [14]. Furthermore, the application handles highly personal information, presenting substantial data privacy and security challenges. The utility of the expert directory and information repository is also limitedby the comprehensiveness and currency of their databases.

Given more time, the primary focus would be on validating the genetic analysis feature. This would involve rigorous testing and implementing a "human-in-the-loop" system, where qualified genetic counselors verify the AI's findings before they reach the user [15].

In the future, the program could be expanded to integrate with telehealth services, allowing users to seamlessly connect with specialists based on their results. Further expansion could include secure integration with electronic health records for more holistic and personalized insights.

If starting over, the project would begin with a narrower scope, focusing first on building a robust and reliable medical information and expert directory. The high-risk genetic analysis feature would be introduced later, developed with a foundational "safety-first" approach that prioritizes clinical validation and expert oversight from the outset.

REFERENCES

[1] Nama, Prathyusha. "AI-Powered Mobile Applications: Revolutionizing User Interaction Through Intelligent Features and Context-Aware Services." Journal of Emerging Technologies and Innovative Research 10.01 (2023): g611-g620.

[2] Das, Sumit, et al. "AI Doctor: An intelligent approach for medical diagnosis." Industry Interactive Innovations in Science, Engineering and Technology: Proceedings of the International Conference, I3SET 2016. Singapore: Springer Singapore, 2017.

[3] Schaaf, Jannik, et al. "Diagnosis of Rare Diseases: a scoping review of clinical decision support systems." Orphanet journal of rare diseases 15.1 (2020): 263.

[4] Wang, Dakuo, et al. "“Brilliant AI doctor” in rural clinics: Challenges in AI-powered clinical decision support system deployment." Proceedings of the 2021 CHI conference on human factors in computing systems. 2021.

[5] Venkatesh, Viswanath. "Adoption and use of AI tools: a research agenda grounded in UTAUT." Annals of operations research 308.1 (2022): 641-652.

[6] Knaus, Brian J., and NiklausJ. Grünwald. "vcfr: a package to manipulate and visualize variant call format data in R." Molecular ecology resources 17.1 (2017): 44-53.

[7] Thirunavukarasu, Arun James, et al. "Large language models in medicine." Nature medicine 29.8 (2023): 1930-1940.

[8] Mahadeo, Resham N. "Artificial Intelligence Chat." A Master's Paper in Computer Science, The Pennsylvania State University, The Graduate School Capital College (2004).

[9] Paredes, Cristian Mauricio Gallardo, Cristian Machuca, and Yadira Maricela Semblantes Claudio. "ChatGPT API: Brief overview and integration in Software Development." International Journal of Engineering Insights 1.1 (2023): 25-29.

[10] Auger, Tom, and Emma Saroyan. "Overview of the OpenAI APIs." Generative AI for Web Development: Building Web Applications Powered by OpenAI APIs and Next. js. Berkeley, CA: Apress, 2024. 87-116.

[11] Wojtara, Magda, et al. "Artificial intelligence in rare disease diagnosis and treatment." Clinical and Translational science 16.11 (2023): 2106-2111.

[12] Germain, Dominique P., et al. "Applying artificial intelligence to rare diseases: a literature review highlighting lessons from Fabry disease." Orphanet Journal of Rare Diseases 20.1 (2025): 186.

[13] Schumacher, Elliot, DhruvNaik, and Anitha Kannan. "Rare Disease Differential Diagnosis with Large Language Models at Scale: From Abdominal Actinomycosis to Wilson's Disease." arXiv preprint arXiv:2502.15069 (2025).

[14] Chew, Robert, et al. "LLM-assisted content analysis: Using large language models to support deductive coding." arXiv preprint arXiv:2306.14924 (2023).

[15] Mosqueira-Rey, Eduardo, et al. "Human-in-the-loop machine learning: a state of the art." Artificial Intelligence Review 56.4 (2023): 3005-3054.