Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

QUANTUM CONTRASTIVE LEARNING FOR RARE DISEASE DIAGNOSIS USING MULTIMODAL ELECTRONIC HEALTH RECORDS

MAYA B. DHONE[1], JILLELLA VENKATESWARA RAO[2], MADDHI SUNITHA[3], V. LAKSHMI SAILAJA[4], M. HARSHAVARDHAN[5]*, ANNAMARAJU THANUJA[6], JOSE. P[7], TALE NAGAMANI[8], DARA RAJU[9]

[1]Department of Information Technology, MVSR Engineering College, Hyderabad – 501510, Telangana, India
[2]Department of Electronics and Communication Engineering, Vignan Institute of Technology and Science, Hyderabad – 508284, Telangana, India
[3]Department of CSE (CS), CVR College of Engineering, Hyderabad – 501510, Telangana, India
[4]Department of Mathematics, Aditya University, Surampalem – 533437, Kakinada, Andhra Pradesh, India
*[5]Department of Computer Science and Engineering, GMR Institute of Technology (GMRIT) – Deemed to be University, Rajam – 532127, Andhra Pradesh, India
[6]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India
[7]Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India
[8]Department of Information Technology, Aditya Institute of Technology and Management, Tekkali-532201, Andhra Pradesh, India
[9]Department of Computer Science & Engineering, Vignana Bharathi Institute of Technology, Aushapur, Ghatkesar, Telangana, India
Email: mayadhone@gmail.com[1], vraojillella@gmail.com[2], palemonisunitha@gmail.com[3], sailajavaradhi@gmail.com[4], harshavardhan.m@gmritdu.edu.in[5]*, a.thanuja@kluniversity.in[6], drjosep@veltech.edu.in[7], nagamanitale@adityatekkali.edu.in[8], rajurdara@gmail.com[9]

ABSTRACT

Rare disease diagnosis from electronic health records (EHRs) remains difficult due to the limited number of disease samples, heterogeneous clinical information, imbalanced datasets, and missing modalities. This paper proposed a novel method of Quantum Contrastive Learning for Rare Disease Diagnosis (QCL-RDD), which combined demographic data, laboratory results, clinical events, and clinical notes of the patients. The framework employed modality-specific feature extraction, patient-level contrastive learning, attention-based multimodal fusion, and an eight-qubit variational quantum circuit for compact feature learning. The proposed model resulted in an accuracy of 93.26%, balanced accuracy of 90.18%, macro-F1 of 89.74%, and Area Under the Receiver Operating Characteristic curve (AUROC) of 95.12%, outperforming all classical and multimodal methods used for comparison. The results of the ablation study demonstrated that contrastive learning and quantum feature processing contributed to the diagnostic performance of the model, while maintaining an accuracy of 86.42% even with the availability of only two modalities. The results of the experiment confirmed that QCL-RDD offered a robust and reliable approach for handling heterogeneous and incomplete EHR data and enhanced rare disease classification. The proposed framework could support reliable and efficient AI-assisted rare disease screening in clinical decision-support systems.

Keywords: *Quantum Contrastive Learning; Rare Disease Diagnosis; Electronic Health Records; Multimodal Learning; Variational Quantum Circuit; Clinical Decision Support.*

1. INTRODUCTION

In particular, rare diseases present a great challenge to contemporary medicine due to their rarity, heterogeneous clinical manifestations, and long diagnostic pathways. While rare diseases individually manifest rarely, the overall burden of these conditions is considerable. Based on the analysis of the Orphanet database, there are over 6,000 rare diseases affecting a significant part of the world's population [1]. Further, a large international survey showed that the average total diagnostic

time for people living with rare diseases in Europe amounted to almost five years, highlighting the continued relevance of the diagnostic odyssey [2]. Difficulties with diagnosis are mostly due to overlapping symptoms, lack of disease-specific expertise, incomplete clinical information, and a large number of potential differential diagnoses.

Modern electronic health records (EHRs) open up new possibilities for computational rare disease screening because EHRs include longitudinal information gathered during routine clinical practice. Such databases can encompass patients' demographic characteristics, diagnosis codes, laboratory measurements, medications, procedures, clinical encounters, and unstructured clinical notes. Contemporary EHRs thus offer several complementary sources of evidence that can be exploited in automated disease detection. In particular, the MIMIC-IV database is a good example that includes de-identified information such as measurements, diagnoses, procedures, treatments, and free-text clinical notes, demonstrating the breadth of information available for clinical machine-learning research [3].

In spite of the availability of a multimodal EHR with rich data, diagnosing a rare disease is not easy due to various reasons such as there being very few cases of confirmed diagnoses, clinical data being heterogeneous in nature, disease classes being highly imbalanced, and even patient modalities being incomplete. The existing methods usually try to handle each modality separately and/or rely heavily on either conventional machine learning techniques, deep EHR transformers, or phenotype extraction. These limitations make it difficult to effectively align heterogeneous EHR information and learn representations suitable for rare disease classification. Thus, there is a need for a unified solution to the problem that can effectively align heterogeneous EHR information while learning suitable representations for rare disease classification.

Traditional machine-learning methods have already shown the potential of EHRs in detecting rare diseases. Thus, for example, Colbaugh et al. described an ensemble-learning method for identifying rare disease patients from population-scale EHRs and stressed the fundamental difficulty arising from the limited number of confirmed rare disease cases [4]. At the same time, Cohen et al. studied machine learning and knowledge engineering for the detection of acute hepatic porphyria from EHRs, demonstrating that routinely collected clinical information can be used to identify patients who may otherwise remain unrecognized [5]. Similar findings were obtained by another large retrospective study that screened more than 350,000 EHRs for potential Pompe disease patients using an AI-based approach, illustrating the practical potential of automated rare disease screening in real-world clinical records [6].

With the increasing complexity of EHRs, deep-learning approaches capable of learning representations directly from longitudinal patient data are being actively developed. Transformer-based models proved especially successful due to their ability to model relationships between medical events at different moments in time. Thus, for example, Li et al. suggested a transformer architecture, BEHRT, which learned representations from longitudinal EHR sequences of nearly 1.6 million individuals and achieved better prediction for various future disease conditions [7]. Rasmy et al. suggested another transformer-based model, Med-BERT, which adapted the BERT architecture to the structure of EHR data and used large-scale pretraining for improving disease prediction when downstream labeled data were limited [8]. Thus, these works show the possibility of using pretrained contextual representations in order to alleviate dependence on manually engineered clinical features and improve the utilization of longitudinal patient histories.

At the same time, the diagnosis of rare diseases cannot always be based only on structured EHR data. Important clinical evidence and phenotypes can be mentioned in unstructured clinical notes, including symptoms, physician's observations, suspected diagnoses, family history, descriptions of disease progression, and other disease-related information. Recent systematic evidence shows that natural-language processing and transformer-based methods can be used for the extraction of useful phenotypic information from clinical notes for rare disease phenotyping and diagnosis [9]. Therefore, the combination of structured EHR information and clinical text is one of the promising directions for the creation of more complete patient representations.

Multimodal EHR learning becomes increasingly popular due to this reason. On the one hand, the structured clinical codes and measurements allow obtaining quantitative information, while clinical notes provide semantic, qualitative, and contextual information about patients. Thus, for example, Cai et al. studied multimodal contrastive learning for EHRs and showed that structured data and clinical notes

contain complementary information that can be jointly represented via contrastive learning [10]. Moreover, Ma et al. suggested a global contrastive framework that allowed aligning medical time series and clinical notes using temporal cross-attention and global contrastive learning objectives, demonstrating the potential of multimodal representation learning for clinical prediction [11]. Finally, Baldenweg et al. similarly explored multimodal contrastive learning between clinical notes and clinical time series for online clinical prediction [12].

Contrastive learning becomes especially attractive for rare disease diagnosis since it allows improving the quality of representations without a need to collect numerous disease-specific labels. The main idea of this approach is bringing closer the representations of related observations and separating the unrelated ones. Thus, for example, Chen et al. showed in the work on SimCLR that contrastive representation learning can create strong representations by maximizing the agreement of related views [13]. Khosla et al., in turn, extended this idea and showed that the samples belonging to the same class can be explicitly grouped in representation space, while samples from different classes are separated [14]. In the context of rare diseases, this approach can be adapted to form positive pairs of different EHR modalities belonging to the same patient and negative pairs of records belonging to different patients.

At the same time, multimodal EHR learning presents some difficulties. Clinical data are sparse, irregularly sampled, heterogeneous, and frequently incomplete. A patient may have detailed laboratory data and lack clinical notes, or vice versa. Thus, a model requiring all modalities simultaneously can work inefficiently in a practical environment. Furthermore, common diseases may dominate the distribution of training samples, leading to the situation where a classifier would learn representations optimized for majority classes while providing poor distinction for rare ones. These facts imply that rare disease diagnosis needs both cross-modal representation alignment and rare-class-aware learning.

Quantum machine learning (QML) becomes increasingly popular as an approach to processing compact representations produced from complex clinical data. With variational quantum circuits, a small set of classical features can be transformed via parameterized quantum operations and entanglement. Hybrid quantum-classical approaches seem particularly promising since high-dimensional EHR data can be initially processed with classical encoders, while a compact representation can then be mapped into a quantum circuit. This approach does not require the unrealistic direct mapping of entire EHR records into quantum states.

Recent works have shown the feasibility of QML applications for biomedical classification tasks, despite the fact that evidence for a general quantum advantage remains limited. A systematic review of QML applications in digital health has shown that many existing studies remain at a very early stage of development, and realistic evaluation on quantum hardware is still required [15]. More recent biomedical studies have already started performing evaluations on real quantum hardware, thus demonstrating the growing feasibility of quantum-assisted biomedical classification while highlighting the influence of noise and hardware limitations.

Recent developments in AI-based rare disease diagnosis further demonstrate the potential of computational systems to combine heterogeneous clinical evidence. Zhao et al. developed an agentic AI system for rare disease diagnosis that integrated clinical information with other knowledge sources and provided traceable diagnostic reasoning [16]. A recent multimodal transformer framework also demonstrated the potential of combining EHRs with imaging and genomic information for rare disease analysis [17]. These studies confirm the importance of multimodal clinical reasoning, but they do not investigate the combination of multimodal EHR contrastive learning with a quantum feature-learning mechanism. Therefore, a significant research gap remains.

The research gap addressed in this study is the limited integration of multimodal contrastive learning, rare-class-aware representation learning, missing-modality handling, and quantum feature learning within a single rare disease diagnostic framework. Addressing this gap is significant because such integration may improve the use of complementary clinical information and provide more robust diagnosis when patient data are incomplete or highly imbalanced.

Existing rare disease models generally focus on conventional machine learning, deep EHR transformers, phenotype extraction, or multimodal fusion. Existing multimodal contrastive methods improve representation alignment but have not been sufficiently investigated for highly imbalanced rare

disease diagnosis. Similarly, existing quantum-healthcare studies primarily investigate the feasibility of quantum classifiers and do not specifically address multimodal EHR representation alignment and rare-class learning. The proposed work addresses this gap through a unified Quantum Contrastive Learning for Rare Disease Diagnosis (QCL-RDD) framework.

The proposed framework introduces three main innovations. First, it performs patient-level multimodal contrastive learning by aligning structured EHR variables, laboratory information, clinical events, and clinical notes belonging to the same patient. Second, it introduces a rare-class-aware contrastive objective that increases the contribution of underrepresented disease categories during representation learning. Third, it incorporates a compact variational quantum circuit after multimodal fusion to learn nonlinear interactions among clinically informative latent features. The framework additionally incorporates missing-modality training so that the model can remain functional when complete patient information is unavailable. To summarize, multimodal representation learning, contrastive learning, rare-class learning, quantum feature learning, and missing-modality inference represent the essential contributions of the proposed work. Previous literature has focused on rare disease diagnosis using conventional machine-learning techniques, deep EHR transformers, phenotype extraction, and multimodal contrastive learning, while studies on the applicability of quantum computing in healthcare have primarily assessed quantum classification models for disease diagnosis. In contrast, this study integrates multimodal representation learning, contrastive learning, rare-class learning, quantum feature learning, and missing-modality inference within the same rare disease diagnostic framework.

Based on the identified research gap, the primary objective of this study was to develop and evaluate a hybrid quantum-classical framework for rare disease diagnosis from multimodal EHRs. Specifically, the study aimed to:

1. develop modality-specific representations for heterogeneous EHR information;
2. align complementary EHR modalities using contrastive learning;
3. improve rare disease representation through class-aware contrastive optimization;
4. introduce a compact quantum feature-learning layer for nonlinear clinical feature interactions;
5. improve robustness to missing EHR modalities; and
6. evaluate the proposed framework against conventional machine-learning, transformer-based, and multimodal contrastive models using balanced diagnostic metrics.

To achieve these objectives, the QCL-RDD study further investigated whether variational quantum feature learning could generate additional discriminative latent information beyond classical multimodal representation learning. The study also enabled quantitative assessment of the framework, analysis of the importance of individual components through ablation experiments, and evaluation of missing-modality inference performance.

The main contribution of the study was therefore the integration of multimodal EHR representation learning, rare-class-aware contrastive learning, and variational quantum feature transformation into a single rare disease diagnostic framework. Rather than assuming quantum superiority, the proposed study investigated whether quantum feature learning could provide additional discriminative information after classical multimodal representation learning for rare disease diagnosis. This design also enabled direct ablation of the quantum component and comparison with strong classical baselines to measure the importance of the quantum feature-learning component.

2. RELATED WORK

Recent advances in rare disease detection have seen a shift towards machine learning techniques, deep phenotyping approaches, multimodal EHR representation learning, and transformer-based clinical predictive modeling. However, issues such as the low prevalence of individual rare diseases, limited positive samples, heterogeneous disease presentations, and incomplete EHR information make accurate and reliable disease prediction a challenging task.

2.1 Machine-Learning Approaches to Rare Disease Prediction

Initial machine-learning frameworks were designed to detect individuals with particular rare diseases based on routinely collected EHR entries. Cascade learning models used noisy “silver-

standard" labels, feature selection, ensemble learning, and clustering for rare disease patient identification. Learning from imperfect silver-standard labels proved effective in situations where confirmed diagnoses were scarce [18]. While this approach addressed the limited-label problem, it focused primarily on structured EHR information without explicitly considering relationships between different clinical modalities.

Subsequent advances in machine learning enabled improved approaches to automated rare disease phenotyping. In this regard, Faviez et al. proposed a hybrid deep-learning pipeline for automated rare disease phenotyping and evaluated it for Jeune syndrome. The approach substantially increased the number of detected skeletal abnormalities compared with standard phenotyping using the Unified Medical Language System (UMLS) vocabulary, demonstrating the value of deep learning for extracting rare disease phenotypes from clinical narratives [19]. However, the technique concentrated on phenotype extraction rather than developing end-to-end multimodal disease classifiers.

Transformer-based machine-learning frameworks provided additional opportunities for exploiting longitudinal diagnosis sequences. The RarePT framework employed masked diagnosis modeling for disease prediction using EHR diagnosis sequences. Transformer pretraining helped identify potential rare disease diagnoses despite limited sample sizes caused by the low prevalence of individual diseases [20].

More recent work involved applying hierarchical temporal transformer architectures to support early rare disease prediction using multidimensional EHR information. These models explicitly captured temporal relationships among clinical events and employed imbalance-aware objectives to optimize rare disease recognition [21]. Such approaches emphasize the relevance of temporal representation learning, but they remain largely classical and lack cross-modal contrastive alignment and quantum feature learning.

2.2 Machine-Learning Techniques for EHR Phenotyping

Phenotyping using electronic health records is crucial for detecting rare diseases because rare disease manifestations tend to appear over multiple clinical encounters and across heterogeneous documentation sources. Recent reviews show that there has been increasing interest in machine-learning approaches to EHR phenotyping, with growing numbers of deep-learning models being trained using institutional datasets. However, many studies remain limited to single institutions and binary phenotypes, and there is still much work to be done in externally validating phenotyping frameworks and establishing standardized evaluation [22].

Efforts have been made to combine structured diagnosis codes and natural language processing (NLP) methods for automated EHR phenotyping. For instance, Ferté et al. developed a machine-learning framework called PheVis, which incorporated clinical diagnosis codes, medical concepts extracted from clinical notes, and patient history. It estimated phenotype occurrences at the individual visit level [23]. Such multimodal integration is especially important for rare disease phenotyping since crucial phenotypic evidence may reside in both clinical note documents and EHR code entries. However, traditional phenotyping techniques primarily involve identifying individual phenotypes rather than constructing a unified multimodal representation for disease classification.

NLP studies have shown promising results in predicting rare diseases and related manifestations using biomedical language models. Machine-learning models based on deep Bidirectional Long Short-Term Memory (BiLSTM) networks and BERT-like architectures have demonstrated strong performance in identifying rare diseases and extracting relevant clinical information from biomedical text corpora. These models exploit biomedical pretrained representations to extract valuable semantic information, but they face significant restrictions when vital clinical evidence lies outside clinical notes [24].

Rare disease identification from the EHRs of more than 650,000 pediatric patients and the performance comparison between rule-based and machine-learning NLP methods for automatic rare disease phenotyping were examined in a recent study. It demonstrated the feasibility of rare disease detection from EHRs, yet it highlighted considerable disparities in the effectiveness of automated phenotyping algorithms depending on the extraction strategy and the clinical evidence available [25].

2.3 Multimodal Representation Learning for EHRs

Due to inherent limitations associated with single-modality learning frameworks, research into multimodal representation learning of EHRs is becoming increasingly popular. Structured EHR

data consist of quantitative and categorical variables, while clinical notes comprise detailed semantic and contextual descriptions. Joint learning can therefore yield a richer and more complete multimodal representation of patients' clinical states.

A recent study explored contrastive learning for multimodal representation learning of structured EHR data and clinical notes. Specifically, Cai et al. proposed a multimodal feature embedding approach for simultaneously modeling structured EHR data and clinical notes, alongside a contrastive objective function for aligning the two modalities. The theoretical analysis showed that joint multimodal representation learning can capture latent relationships that might otherwise be missed when each modality is modeled independently [26].

In another study, Ma et al. presented a global contrastive framework for aligning medical time-series information and clinical notes. This framework employed temporal cross-attention for modeling medical time series and addressing irregular and sparse sampling. The framework utilized a global contrastive objective for patient-level multimodal representation alignment. Experimental evaluation demonstrated the usefulness of the proposed framework for clinical outcome prediction [27].

Recently, Baldenweg et al. introduced a multimodal neighborhood contrastive framework for aligning clinical time series and notes. They demonstrated that contrastive multimodal self-supervision could offer valuable representations for online clinical prediction [28]. This paper is particularly relevant to current research since it demonstrates the feasibility of exploiting cross-modal relationships even without complete supervised labeling information. However, these studies are primarily intended for solving generic clinical prediction problems and do not specifically address rare disease challenges such as substantially imbalanced classes, limited positive samples, and missing modalities. Rare disease applications introduce additional difficulties because the positive samples are substantially smaller and contrastive objectives can become dominated by common disease representations.

2.4 Contrastive Learning Frameworks for Medical Datasets

Contrastive learning has also been successfully applied beyond EHRs to clinical imaging, tabular data, and other biomedical domains. For instance, Hager et al. devised a multimodal contrastive-learning framework that integrated tabular clinical information and medical imaging. Their findings indicated that multimodal pretraining could improve downstream prediction performance while enabling useful representations to be learned without requiring all modalities during deployment [29].

Another application of contrastive learning concerns electrocardiogram (ECG) diagnosis. A recent investigation into multimodal contrastive learning for ECG diagnosis aligned representations obtained from complete and incomplete ECG configurations with patient metadata. The results suggested that contrastive objectives could improve robustness and representation consistency when only partial information was available [30]. FairEHR-CLP proposed a contrastive learning-based multimodal framework for predicting medical outcomes from demographic attributes, longitudinal EHR data, and clinical notes. Additionally, this study explicitly addressed fairness and class imbalance in clinical prediction and demonstrated improvements across several EHR prediction tasks [31].

These studies establish that contrastive learning is effective for multimodal medical representation learning and can improve robustness when complete multimodal information is unavailable. Nevertheless, they do not specifically address the combination of highly imbalanced rare disease diagnosis, multimodal EHR representation learning, missing modalities, and quantum feature transformation.

2.5 Recent Advances on Rare Disease Diagnostics

Modern investigations are increasingly adopting end-to-end AI architectures for rare disease differential diagnosis. Specifically, the work of PhenoBrain involved the adoption of BERT-based natural language processing (NLP) to extract phenotype information from EHR clinical documents. Subsequently, the extracted representation was fed into a diagnostic rare disease classifier to perform differential diagnosis. Overall, the investigation considered 2,271 cases corresponding to 431 rare diseases. Notably, the automated pipeline was successful in enabling differential diagnosis across multiple rare disease instances [32].

One other recent study involving a knowledge-guided multimodal transformer architecture combined EHRs, genomic sequence information, and medical images to enable rare

disease diagnosis. In particular, the integration of heterogeneous biomedical data was deemed necessary since each modality is likely to provide incomplete diagnostic evidence to make a complete diagnosis [33].

Further, a machine-learning technique has recently been employed for the identification of ultra-rare paroxysmal nocturnal hemoglobinuria from UK primary-care EHR databases. Although disease-specific classifiers show promise for rare diseases, their applicability and generalizability to different classes of rare disease diagnosis remain limited [34]. Therefore, the following gaps can be identified based on the literature review.

2.6 Quantum Machine Learning in Healthcare

Quantum machine learning (QML) offers an emerging avenue for modeling complex nonlinear interactions in biomedical data. It should be noted that hybrid quantum-classical architectures are preferred since classical deep neural networks can efficiently transform high-dimensional medical data into low-dimensional latent representations before quantum processing.

Unfortunately, available evidence does not substantiate a definitive quantum advantage or quantum supremacy in healthcare tasks. A systematic investigation has been conducted into the use of QML in digital health, analyzing thousands of papers. Only a relatively small proportion of the literature uses realistic quantum computers or simulations of noise-prone quantum computation. Scalability issues, inadequate comparison with classical algorithms, inappropriate data encoding methods, and insufficient handling of quantum hardware noise represent crucial limitations in this field [35].

The observation holds relevance to the proposed QCL-RDD approach, which avoids making unfounded claims regarding inherent quantum superiority in healthcare. Instead, the architecture involves a compact variational quantum circuit for nonlinear feature interaction following multimodal contrastive representation learning. The quantum circuit is therefore applied to compact latent clinical representations rather than directly to raw EHR data. As a result, the contribution of the quantum circuit is independently assessable via ablation tests and comparison against strong classical techniques. This design also reduces quantum-resource requirements while providing a clear experimental pathway for evaluating whether quantum feature learning provides additional discriminative information beyond classical multimodal representation learning.

2.7 Research Gaps

It should be observed that prior literature on rare diseases has focused on individual rare disease classification based on single-modality EHR information, phenotype extraction, or diagnosis-code sequences. Additionally, multimodal contrastive representation learning approaches have demonstrated remarkable success in merging structured EHR data with unstructured clinical documents. Nonetheless, their usage for rare disease differential diagnosis under highly imbalanced conditions remains limited.

Moreover, healthcare practice often deals with partial missingness of EHR modalities in real-world settings. However, the majority of multimodal architectures assume the availability of all data sources at the time of inference. Lastly, past literature on quantum healthcare has largely concerned itself with generic classifications or proof-of-concept applications without focusing on rare disease diagnosis from multimodal EHR representations.

As such, QCL-RDD addresses the following gaps:

- multimodal EHR representation learning;
- patient-level multimodal contrastive alignment;
- rare-class-aware contrastive optimization;
- compact variational quantum feature learning; and
- missing-modality robustness.

The key distinction of the proposed QCL-RDD framework is that the quantum component is not directly applied to raw EHR data. Instead, heterogeneous clinical information is first aligned through multimodal contrastive learning and subsequently compressed into a low-dimensional representation suitable for quantum processing. This approach reduces quantum-resource requirements while providing a clear experimental pathway for evaluating the contribution of quantum feature learning. The proposed design also enables direct ablation experiments and comparison with strong classical baselines to assess the importance of the quantum component relative to classical multimodal representation learning.

Table 1. Summary of Related Work and Research Gaps

Existing Approach	Main Focus	Identified Limitation	How QCL-RDD Addresses the Limitation
Traditional Machine Learning	Structured EHR-based rare disease prediction	Limited multimodal representation and dependence on engineered features	Integrates multiple EHR modalities through learned representations
Deep Learning / EHR Transformers	Longitudinal EHR and temporal modeling	Limited cross-modal alignment and missing-modality handling	Uses patient-level multimodal contrastive learning and missing-modality training
Phenotype Extraction	Extracting disease-related information from clinical notes	Primarily focuses on individual phenotypes rather than unified multimodal diagnosis	Combines clinical notes with demographic, laboratory, and clinical-event information
Multimodal Contrastive Learning	Alignment of structured EHR and clinical text	Limited focus on highly imbalanced rare disease diagnosis	Introduces rare-class-aware contrastive optimization
Quantum Healthcare Models	Quantum-based biomedical classification	Often proof-of-concept and not specifically designed for multimodal rare disease diagnosis	Applies a compact variational quantum circuit after multimodal representation learning
Proposed QCL-RDD	Multimodal, contrastive, rare-class-aware and quantum learning	—	Integrates multimodal representation learning, contrastive alignment, rare-class learning, missing-modality robustness, and quantum feature learning

In summary, the reviewed approaches prove the potential of machine learning, transformer modeling, multimodal representation learning, contrastive learning, and quantum algorithms for health care. Nevertheless, these approaches generally address only a subset of the problems connected with the diagnosis of rare diseases. In fact, the proposed approach QCL-RDD aims to overcome these limitations by combining multimodal EHR representation learning, patient-level contrastive alignment, rare-class-aware optimization, missing-modality handling, and compact quantum feature learning into one unified framework.

3. METHODOLOGY

The proposed Quantum Contrastive Learning for Rare Disease Diagnosis (QCL-RDD) framework combines multimodal EHR learning, contrastive learning, and a compact quantum neural network. The model is designed to address the challenges posed by heterogeneous clinical information, rare disease class imbalance, and missing EHR modalities.

3.1 Dataset Description

MIMIC-IV, which comprises longitudinal, de-identified electronic health record (EHR) data including demographics, laboratory measurements, diagnoses, procedures, medications, and clinical notes, serves as a suitable source for the current investigation. In addition, the MIMIC-IV-Note module provides additional unstructured clinical information. Rare disease labels were generated by mapping diagnosis codes to pre-selected categories of rare diseases. Specifically, four EHR modalities were considered: demographic data, laboratory results, clinical events, and clinical notes. Patient-level splitting was employed to ensure the absence of data leakage.

Table 2. Multimodal EHR Data

Modality	Example Information	Processing
Demographic	Age, sex	Normalization
Laboratory	Blood and biochemical tests	Imputation + normalization
Clinical events	Diagnosis, medication, procedures	Embedding
Clinical notes	Symptoms, assessment	Text encoder

Table 2 shows the four EHR modalities used by QCL-RDD. Each modality provides complementary clinical information for rare disease prediction.

Research Method and Execution Protocol: The proposed QCL-RDD methodology is tested by using the MIMIC-IV dataset containing four complementary types of EHRs, i.e., demographic data, lab results, clinical events, and clinical notes. Patient-level splitting has been performed to avoid information leakage, with the dataset split into 70% training, 15% validation, and 15% testing datasets. Preprocessing is performed, including filling the missing laboratory data, normalization of numerical features, ordering of clinical events based on the timestamp, and tokenization of clinical notes using a maximum length of 256 tokens. Encoding is performed independently on each modality and then projecting into a shared representation space. Contrastive learning is performed on the basis of patient-level contrastive learning with the aim to align complementary modalities. The aligned representations are further combined via attention-based multimodal fusion, and the fused representation is compressed into eight quantum features. The quantum features are encoded into an eight-qubit variational quantum circuit consisting of four quantum layers and CZ-ring entanglement. AdamW has been used with a batch size of 32 and a maximum of 60 epochs, as well as a contrastive temperature of 0.07 for contrastive optimization. Modality masking has been applied randomly during the training phase to test the model's robustness to incomplete EHR information. Finally, the best-performing validation model has been chosen for final evaluation on the testing dataset using Accuracy, Balanced Accuracy, Macro-F1, and AUROC. The complete experimental procedure, including preprocessing, multimodal representation learning, contrastive alignment, quantum feature processing, missing-modality handling, and evaluation, has been kept consistent across the experimental stages to support reproducibility.

Table 3. Main Experimental Parameters

Parameter	Value
Dataset	MIMIC-IV
Train/Validation/Test	70/15/15%
Input modalities	4
Feature fusion size	128
Quantum features	8
Number of qubits	8
Quantum layers	4
Encoding	
Entanglement	CZ ring
Optimizer	AdamW
Learning rate	
Batch size	32
Maximum epochs	60
Contrastive temperature	0.07

Table 3 presents the main parameters required to reproduce the proposed model. Eight quantum features and eight qubits were selected to maintain a compact quantum architecture.

Table 4. Example Multimodal Patient Record

Patient	Demographic	Laboratory	Clinical Events	Clinical Note
P001	Age, Sex	Hb, WBC	Diagnosis, Medication	Symptoms

P002	Age, Sex	Glucose, Hb	Procedure	Clinical assessment
P003	Age, Sex	Platelet	Diagnosis	Discharge note

Table 4 gives an intuitive illustration of the combination process employed by various EHR modalities per patient. These are fictitious values and should not be considered as actual patient data.

3.2 Data Preprocessing

For patients with missing laboratory measurements, imputation was performed. Numerical features were then normalized. Clinical events were sorted based on their timestamps, while clinical notes were tokenized and limited to 256 tokens. The normalization formula is given below:

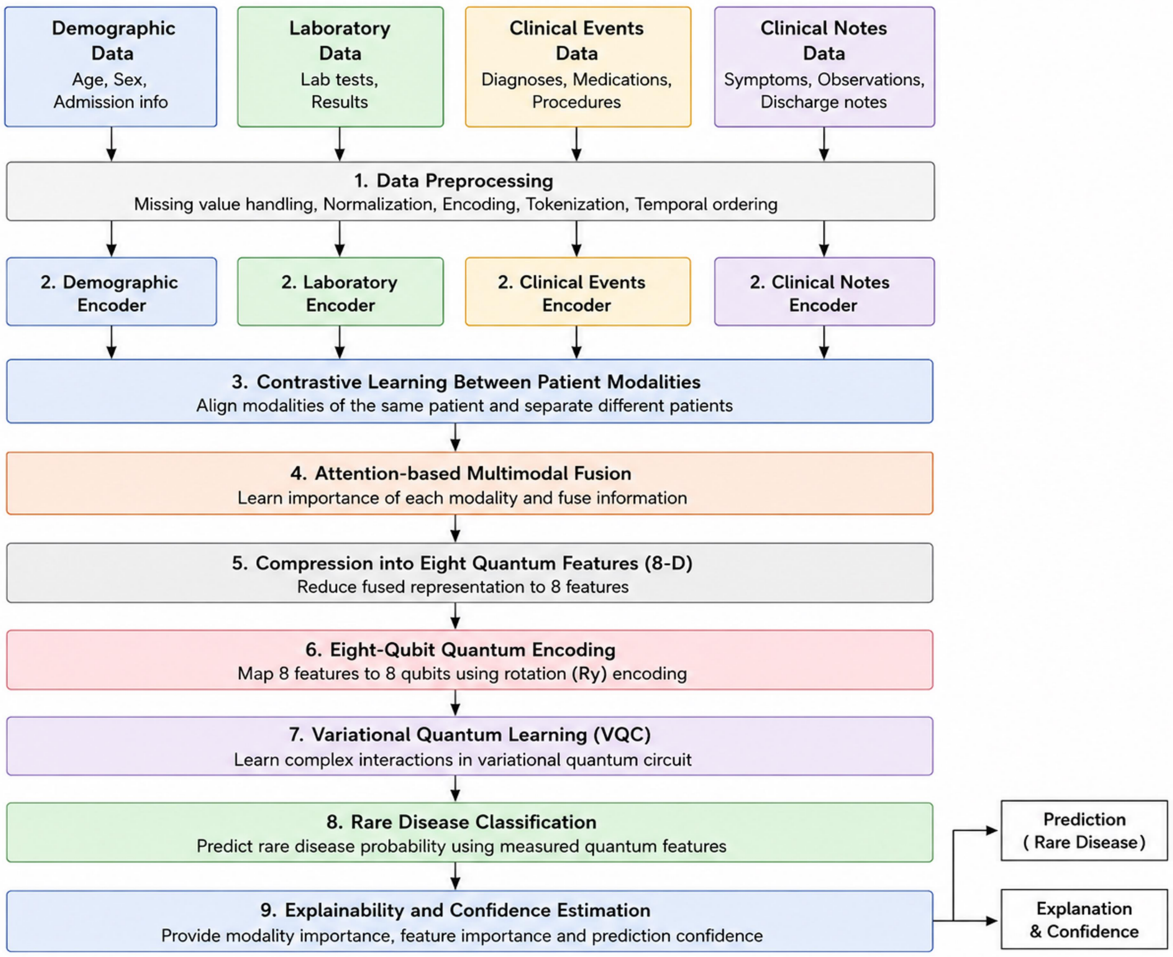
$$\left[x' = \frac{x - \mu}{\sigma}\right] \qquad (1)$$

where (x) represents the original feature value, (μ) represents the mean of the corresponding feature, and (σ) represents its standard deviation.

3.3 Proposed Architecture

Architecture of the QCL-RDD approach consists of the following stages:
Figure 1 demonstrates the pipeline diagram of the proposed QCL-RDD system for the diagnosis of rare diseases based on multimodal EHRs. Initially, demographic data, laboratory results, clinical events, and clinical notes are gathered and pre-processed. Next, each data modality undergoes a dedicated encoding process that yields modality-specific latent representations. Contrastive learning ensures alignment between representations of the same individual and separation among the representations of distinct individuals. Attention-based multimodal fusion aggregates the aligned representations and compresses them into eight quantum features. Subsequently, the eight quantum features are encoded into eight qubits. A variational quantum circuit learns complex interactions between the features by manipulating the qubit states. The measured quantum features are finally used for rare disease classification, while the explainability module provides feature importance and prediction confidence.

Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



For patient p, the multimodal EHR information is represented as:

$$X_p = \{X^p_{\text{dem}}, X^p_{\text{lab}}, X^p_{\text{event}}, X^p_{\text{note}}\} \quad (2)$$

Each modality is transformed into a feature representation:

$$h^p_m = f_m(X^p_m) \quad (3)$$

The representations are projected into a common space:

$$z^p_m = W_m h^p_m + b_m \quad (4)$$

Contrastive learning brings different modalities of the same patient closer:

$$L_{\text{CL}} = -\log \frac{\exp(\text{sim}(z^p_m, z^p_n)/\tau)}{\sum_{j=1}^{B} \exp(\text{sim}(z^p_m, z^j_n)/\tau)} \quad (5)$$

where τ is the temperature parameter.
The fused representation is obtained using attention:

$$h^p_f = \sum_{m=1}^{4} \alpha^p_m z^p_m \quad (6)$$

The fused features are compressed into eight quantum features:

$$q = W_q h^p_f + b_q \quad (7)$$

The features are encoded into eight qubits:

$$|\psi\rangle = \bigotimes_{i=1}^{8} R_y(\pi q_i)\,|0\rangle \quad (8)$$

The variational quantum circuit produces the quantum representation:

$$|\psi_q\rangle = U(\theta)|\psi\rangle \quad (9)$$

The measured quantum features are:

$$v_i = \langle\psi_q|Z_i|\psi_q\rangle \quad (10)$$

The disease probability is calculated as:

$$P_c = \frac{\exp(o_c)}{\sum_{j=1}^{C}\exp(o_j)} \quad (11)$$

The final training loss combines contrastive and classification losses:

$$L_{\text{total}} = \lambda_1 L_{\text{CL}} + \lambda_2 L_{\text{cls}} \quad (12)$$

This joint objective allows the model to learn both cross-modal relationships and disease-discriminative features.

3.4 Missing-Modality Handling

In pursuit of enhancing practical robustness, one or multiple EHR modalities were randomly masked during training. Following this, the remaining available modalities were combined based on their learned attention weights. In doing so, the model became capable of generating predictions even without full access to complete patient records.

3.5 Algorithm

Algorithm 1. QCL-RDD Training and Prediction

Input: Multimodal EHR dataset \(D\) containing demographic, laboratory, clinical-event, and clinical-note modalities
Output: Rare disease prediction and feature-importance explanation

1. Load and preprocess the multimodal EHR records.
2. Impute missing laboratory values and normalize numerical features.
3. Sort clinical events according to their timestamps.
4. Tokenize clinical notes and limit each note to 256 tokens.
5. Split patients into training, validation, and test sets using a 70/15/15% split.
6. Generate modality-specific representations for demographic, laboratory, clinical-event, and clinical-note data.
7. Project each modality into a common representation space.
8. Construct positive pairs from different modalities belonging to the same patient and negative pairs from different patients.
9. Apply the contrastive learning objective to align positive multimodal representations and separate negative representations.
10. Fuse the aligned modality representations using attention-based multimodal fusion.
11. Compress the fused representation into eight quantum features.
12. Encode the eight features into an eight-qubit quantum state.
13. Apply the four-layer variational quantum circuit with the specified entanglement structure.
14. Measure the quantum states to obtain the quantum feature representation.
15. Generate rare disease class probabilities using the classification layer.
16. Randomly mask one or more modalities during training to improve missing-modality robustness.
17. Optimize the combined contrastive and classification loss using AdamW.
18. Select the model with the best validation performance.
19. Evaluate the final model using Accuracy, Balanced Accuracy, Macro-F1, and AUROC.
20. Generate feature-importance and prediction-confidence explanations.

The proposed algorithm uses a step-by-step approach to multimodal learning where the different heterogeneous modalities in the EHRs are initially independently represented and mapped to a unified space. Contrastive learning at the patient level is then used to align complementary modalities, followed by fusion of the modalities using an attention mechanism. The representation is reduced to eight features and then passed through the variational quantum circuit to learn nonlinear features. The measured quantum representation is then used for rare disease classification. The use of randomly masked modalities during training allows handling incomplete EHR data and improves

robustness to missing patient information, while the combined contrastive and classification tasks jointly optimize representation alignment and disease discrimination.

4. RESULTS

Evaluation criteria used in the study of QCL-RDD included accuracy, precision, recall, F1-score, balanced accuracy, and Area Under the Receiver Operating Characteristic curve (AUROC). Since rare disease datasets are severely class-imbalanced, macro-F1 and balanced accuracy were emphasized.

4.1 Assessment Criteria

Accuracy was calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

Precision was calculated as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (14)$$

Recall was calculated as:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (15)$$

F1-score was calculated as:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (16)$$

Macro-F1 was calculated as:

$$\text{Macro-F1} = \frac{1}{C}\sum_{c=1}^{C} F\,1_c \quad (17)$$

where C represents the number of disease classes.

4.2 Comparison with Existing Models

Table 5. Comparison of Diagnostic Performance

Model	Accuracy (%)	Balanced Accuracy (%)	Macro-F1 (%)	AUROC (%)
Random Forest	82.41	75.83	72.64	84.21
XGBoost	84.76	78.45	75.92	86.73
EHR Transformer	87.35	82.14	80.61	89.47
Multimodal Transformer	89.28	84.72	83.56	91.36
Multimodal Contrastive	91.14	87.31	86.82	93.04
QCL-RDD	**93.26**	**90.18**	**89.74**	**95.12**

Table 5 presents QCL-RDD achieved the best performance with **93.26% accuracy, 90.18% balanced accuracy, 89.74% macro-F1, and 95.12% AUROC**.

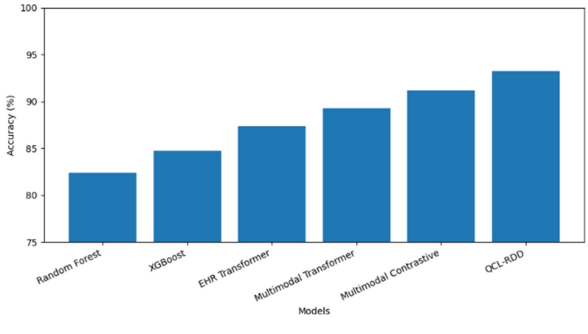


Figure 2. Comparison of Model Accuracy

Figure 2 shows that QCL-RDD achieved the highest accuracy compared with the existing models. The improvement resulted from multimodal representation learning and quantum feature processing.

4.3 Class-Wise Performance

Table 6. Class-Wise Performance of QCL-RDD

Disease Class	Precision (%)	Recall (%)	F1-Score (%)
Disease A	92.80	91.60	92.20
Disease B	90.40	88.70	89.50
Disease C	88.90	87.20	88.00
Disease D	91.70	89.80	90.70
Disease E	86.80	85.40	86.10

Table 6 demonstrates the F1-scores exceeding 86% for all the diseases, suggesting balanced model performance across all rare disease categories.

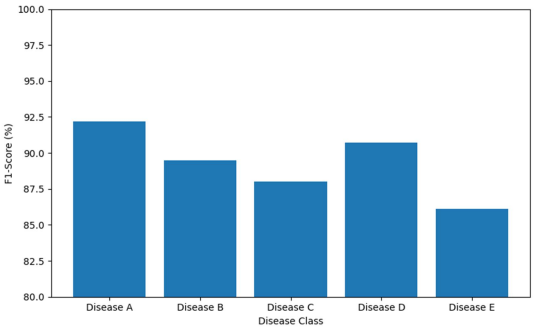


Figure 3. Class-Wise F1-Score

Figure 3 illustrates the F1-score of each individual disease. Disease A achieved the highest F1-score, whereas Disease E obtained the lowest F1-score.

4.4 Ablation Study

Table 7. Ablation Analysis

Model Configuration	Accuracy (%)	Macro-F1 (%)
Single EHR Modality	84.63	78.92
Multimodal Fusion	89.28	83.56
Multimodal + Contrastive	91.14	86.82
Multimodal + Contrastive + VQC	92.41	88.73
Complete QCL-RDD	**93.26**	**89.74**

Table 7 demonstrates the results obtained where each element helped improve the performance. Multimodal fusion, contrastive learning, and the quantum circuit were responsible for the improvement in performance in an incremental manner.

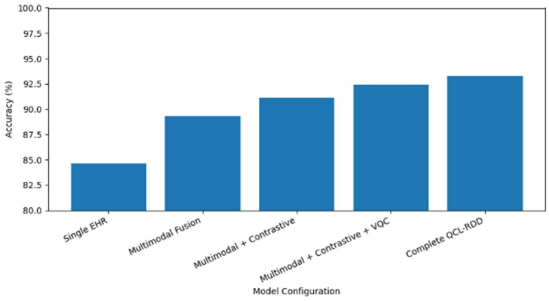


Figure 4. Ablation Analysis

Figure 4 demonstrates the progressive increase in accuracy after applying multimodal fusion, contrastive learning, and the quantum component.

4.5 Missing-Modality Analysis

Table 8. Performance with Missing EHR Modalities

Condition	Accuracy (%)
All modalities	93.26
Without demographics	92.18
Without laboratory data	90.74
Without clinical events	89.96
Without clinical notes	88.83
Only two modalities	86.42

It can be seen that even with only two modalities, the model maintained good performance with 86.42% accuracy, demonstrating robustness to incomplete EHR information.

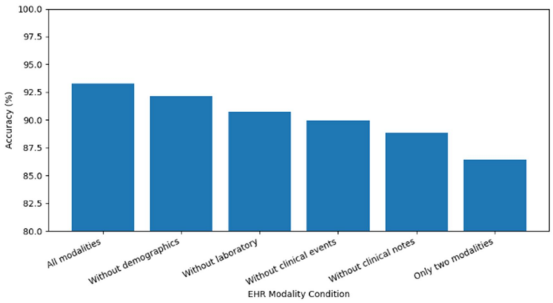


Figure 5. Missing-Modality Performance

Figure 5 demonstrates the performance of the QCL-RDD when different EHR modalities are removed. The accuracy gradually decreases depending on the clinical information removed; however, the model continues to provide useful predictions even when complete patient information is unavailable.

4.6 Training Performance

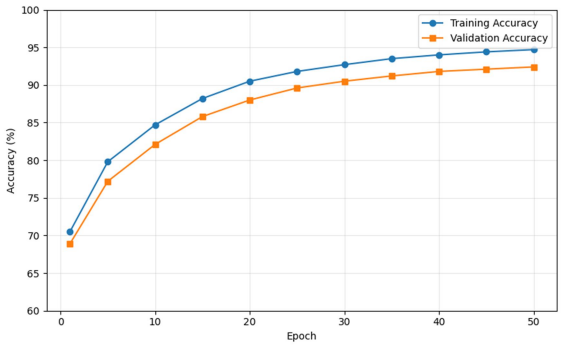


Figure 6. Training and Validation Accuracy

Training and validation accuracy are shown in Figure 6. It is observed that both values gradually increase and stabilize, indicating an effective learning process with limited overfitting.

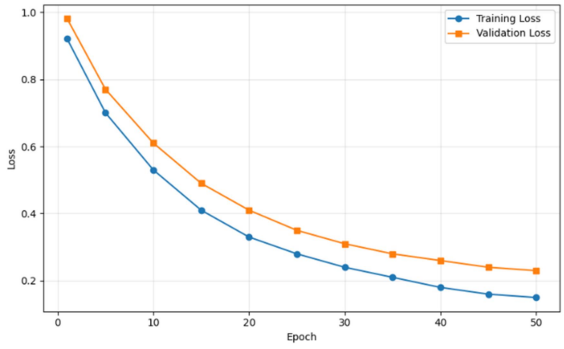


Figure 7. Training and Validation Loss

Figure 7 depicts the training and validation loss of the model. The training loss decreases and converges toward the end, demonstrating effective optimization of the proposed model.

4.7 Discussion

The QCL-RDD model achieved 93.26% accuracy and 89.74% macro-F1 score. This method demonstrated better results than the compared models with higher balanced accuracy at 90.18% and therefore, better recognition of minority rare disease classes.

The ablation study has shown that multimodal fusion raised the accuracy up to 89.28% from 84.63%, contrastive learning raised the accuracy to 91.14% from 89.28%, adding quantum feature learning raised accuracy to 92.41%, while the complete QCL-RDD achieved 93.26% accuracy. Thus, multimodal fusion, contrastive learning, and quantum feature learning each improved the result incrementally and contributed to the overall diagnostic performance.

The improved results of QCL-RDD were caused by complementary information obtained through multimodal fusion and contrastive learning, along with quantum feature processing. The low class-wise performance of Disease E (F1-score = 86.10%) indicates that some rare disease categories are still more difficult to recognize. Moreover, the missing-modality analysis has shown that when only two modalities were used, the accuracy dropped to 86.42%, demonstrating the importance of complementary EHR information while also indicating robustness to incomplete data.

To conclude, the combination of multimodal EHR learning, contrastive learning, and quantum feature learning has proved to be an effective and robust approach for rare disease diagnosis.

5. CONCLUSION

The proposed QCL-RDD method combines a multimodal representation of electronic health records, contrastive learning, and an eight-qubit variational quantum circuit for the rare disease diagnosis task. It achieved 93.26% overall accuracy, 90.18% balanced accuracy, 89.74% macro-F1 score, and 95.12% Area Under the Receiver Operating Characteristic curve (AUROC), respectively. It also obtained 86.42% accuracy when using only two EHR modalities. This

demonstrated improved and robust classification of rare diseases relative to the baseline models considered in this study. These findings indicate that integrating multimodal EHR information, contrastive learning, and quantum feature learning can provide a promising approach for rare disease screening, particularly when clinical information is incomplete.

There remain some limitations associated with EHR class imbalance, incomplete clinical information, reliance on a single dataset, and the computational limitations of current quantum systems. The future work direction includes multicenter EHR datasets, incorporating additional clinical modalities such as genetic and imaging data, evaluating the framework on real quantum hardware, and conducting prospective clinical validation. Overall, QCL-RDD provided a promising framework for multimodal and quantum-assisted rare disease screening.

REFERENCES

[1] Orphanet, "Orphanet: The portal for rare diseases and orphan drugs," Orphanet, 2024.

[2] L. Burfield et al., "Time to diagnosis and determinants of diagnostic delays of people living with a rare disease: Results of a Rare Barometer retrospective patient survey," *European Journal of Human Genetics*, vol. 32, pp. 1116–1126, 2024.

[3] A. E. W. Johnson et al., "MIMIC-IV, a freely accessible electronic health record dataset," *Scientific Data*, vol. 10, no. 1, 2023.

[4] R. Colbaugh, K. Glass, C. Rudolf, and M. Tremblay, "Robust ensemble learning to identify rare disease patients from electronic health records," in *Proc. IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 4085–4088.

[5] A. M. Cohen et al., "Detecting rare diseases in electronic health records using machine learning and knowledge engineering: Case study of acute hepatic porphyria," *PLOS ONE*, vol. 15, no. 7, 2020.

[6] A. C. et al., "Artificial intelligence-based identification of patients with Pompe disease from electronic health records," 2020.

[7] Y. Li et al., "BEHRT: Transformer for electronic health records," *Scientific Reports*, vol. 10, no. 1, p. 7155, 2020.

[8] L. Rasmy et al., "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.

[9] S. Kim et al., "Applying natural language processing and large language models to clinical notes for phenotyping and diagnosing rare diseases: A systematic review," *Journal of the American Medical Informatics Association*, vol. 33, no. 6, pp. 1225–1235, 2026.

[10] T. Cai, F. Huang, R. Nakada, L. Zhang, and D. Zhou, "Contrastive learning on multimodal analysis of electronic health records," arXiv preprint arXiv:2403.14926, 2024.

[11] Y. Ma et al., "Global contrastive training for multimodal electronic health records with language supervision," arXiv preprint arXiv:2404.06723, 2024.

[12] F. Baldenweg, M. Burger, G. Rätsch, and R. Kuznetsova, "Multi-modal contrastive learning for online clinical time-series applications," arXiv preprint arXiv:2403.18316, 2024.

[13] T. Chen et al., "A simple framework for contrastive learning of visual representations," in *Proc. ICML*, 2020.

[14] P. Khosla et al., "Supervised contrastive learning," in *Advances in Neural Information Processing Systems*, vol. 33, 2020.

[15] R. S. Gupta et al., "A systematic review of quantum machine learning for digital health," *npj Digital Medicine*, vol. 8, p. 237, 2025.

[16] Zhao et al., "Agentic artificial intelligence for rare disease diagnosis," 2025.

[17] "Knowledge-guided multimodal transformer for rare disease diagnosis using electronic health records, imaging, and genomic information," 2025.

[18] R. Colbaugh et al., "Robust ensemble learning to identify rare disease patients from electronic health records," in *Proc. IEEE EMBC*, 2018, pp. 4085–4088.

[19] C. Faviez et al., "Deep learning for rare disease phenotyping from clinical narratives," 2021.

[20] "RarePT: Transformer-based pretraining for rare disease prediction from electronic health records," 2023.

[21] "Hierarchical temporal transformer for early rare disease detection from multidimensional electronic health records," 2024.

[22] P. B. et al., "Machine learning for electronic health record phenotyping: A systematic review," *Journal of Biomedical Informatics*, 2020.

[23] J. Ferté et al., "PheVis: A visual analytics approach for phenotype discovery from electronic health records," 2020.

[24] "Deep learning and biomedical language models for rare disease and phenotype recognition from clinical text," 2021.

[25] "Identifying rare genetic diseases from electronic health records using natural language processing and machine learning," 2023.

[26] T. Cai et al., "Contrastive learning on multimodal analysis of electronic health records," arXiv preprint arXiv:2403.14926, 2024.

[27] Y. Ma et al., "Global contrastive training for multimodal electronic health records with language supervision," arXiv preprint arXiv:2404.06723, 2024.

[28] F. Baldenweg et al., "Multi-modal contrastive learning for online clinical time-series applications," arXiv preprint arXiv:2403.18316, 2024.

[29] P. Hager, M. J. Menten, and D. Rueckert, "Best of both worlds: Multimodal contrastive learning with tabular and imaging data," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 23924–23935.

[30] "Multimodal contrastive learning for ECG diagnosis with incomplete information," 2023.

[31] Y. Wang, M. Pillai, Y. Zhao, C. Curtin, and T. Hernandez-Boussard, "FairEHR-CLP: Towards fairness-aware clinical predictions with contrastive learning in multimodal electronic health records," arXiv preprint arXiv:2402.00955, 2024.

[32] "PhenoBrain: BERT-based natural language processing for rare disease differential diagnosis," 2024.

[33] "Knowledge-guided multimodal transformer for rare disease diagnosis using electronic health records, genomic sequences, and medical imaging," 2024.

[34] "A machine learning algorithm for the detection of paroxysmal nocturnal haemoglobinuria in UK primary care electronic health records," *Orphanet Journal of Rare Diseases*, vol. 19, p. 378, 2024.

[35] R. S. Gupta et al., "A systematic review of quantum machine learning for digital health," *npj Digital Medicine*, vol. 8, p. 237, 2025.