ISSN 2689-8268| Volume 9

American Journal of Surgery and Clinical Case Reports

Case Report Open Access

A Digital Twin System to detect Ribose-5-Phosphate Isomerase Deficiency

Upadrista V[1]* and Isaacs R[2]

[1]Department of Computing, Glasgow Caledonian University, Glasgow G4 0BA, Scotland

[2]Former CEO The Permanente Medical Group

***Corresponding author:**

Venkatesh Upadrista,
Department of Computing, Glasgow Caledonian University, Glasgow G4 0BA, Scotland, UK

Received: 12 Jan 2026
Accepted: 22 Jan 2026
Published: 03 Feb 2026
J Short Name: AJSCCR

Citation:

Upadrista V. A Digital Twin System to detect Ribose-5-Phosphate Isomerase Deficiency. Ame J Surg Clin Case Rep. 2026; 9(2): 1-7

Keywords:

Digital Twin; Security and Privacy; Machine Learning; Ribose-5-Phosphate Isomerase Deficiency

1. Abstract

Adigital twin acts as a virtual model of a real-world entity or system, encompassing its entire lifecycle with updates in real time. Digital twins are emerging as powerful tools in healthcare, used for monitoring health metrics and predicting severe conditions. However, their use in predicting rare diseases such as Ribose-5-Phosphate Isomerase Deficiency (RPI deficiency) is still largely unexplored. Existing studies seldom focus on this rare metabolic disorder, and those that do often report limited success in predictive accuracy. Furthermore, current models inadequately address critical aspects of data security and privacy, which are paramount in healthcare.

To address these challenges, we developed a secure, ML-powered digital twin application. It is specifically designed to monitor health data in real-time and accurately predict the likelihood of RPI deficiency. This article proposes the use of machine learning algorithms (decision tree, naive Bayes, K-nearest neighbor, Random Forest, logistic regression) to create a prediction model for RPI deficiencies. The decision tree algorithm is deemed the best, achieving an overall accuracy of 95.0%. Early diagnosis significantly impacts recovery, and integrating machine learning into healthcare holds promise for RPI deficiency prediction and management. By integrating blockchain technology with advanced ML algorithms, our approach ensuring robust data security and privacy.

Our application demonstrates high prediction accuracy, setting a significant milestone in rare disease prediction literature. Additionally, system's architecture supports seamless integration with external medical devices, ensuring comprehensive and real-time data collection. The application's security measures have proven to be tamper-resistant, with built-in mechanisms to detect and correct any anomalies in the data.

2. Introduction

A digital twin is a digital representation of the physical state and behavior of a real-world entity, such as a person or a machine, kept updated through data from sensors and devices connected to the physical counterpart.

Ribose-5-Phosphate Isomerase Deficiency (RPI) is an extremely rare metabolic disorder characterized by the disruption of the pentose phosphate pathway, leading to neurological dysfunctions and developmental delays. The rarity and complexity of RPI deficiency make early diagnosis and intervention challenging, necessitating innovative approaches like digital twins for effective monitoring and prediction. Thought there is no cure for RPI, early diagnosis of Ribose-5-Phosphate Isomerase (RPI) deficiency can help prevent or slow the progression of severe leukoencephalopathy, seizures, and psychomotor regression. Prompt intervention allows for better management of visual and hearing impairments, reducing the risk of life-threatening complications and improving overall motor and cognitive outcomes, which significantly enhances the patient's quality of life.

The current literature shows a significant gap in using digital twins to predict rare diseases like RPI deficiency. Most existing digital twins in healthcare focus on common conditions and lack the extensibility to accommodate rare metabolic disorders. Additionally, data security and privacy concerns are often inadequately addressed, leaving sensitive health information vulnerable to breaches.

We propose a novel, secure digital twin application powered by machine learning and blockchain technology. This application not only predicts RPI deficiency with high accuracy at 95% level, but also ensures the security and privacy of patient data. Moreover, the system is designed to be versatile, allowing for its application to

other metabolic disorders with minimal modifications.

2.1. Objective

The objective of this study is to create a reliable and efficient prediction system for Ribose-5-Phosphate Isomerase (RPI) deficiency that can accurately identify the condition. To achieve this, we tested various machine learning models, including Random Forest, Decision Tree, Logistic Regression, and KNN, to determine which model offered the highest accuracy. Our findings revealed that the KNN model was the most effective, achieving a 95% accuracy rate. Additionally, we developed a user interface that indicates whether an individual has RPI deficiency and a separate interface for doctors to review detailed patient health data.

3. Literature Review

This limited collection of available studies was examined to explore the critical role of ribose-5-phosphate isomerase (RPI) in various diseases, emphasizing the challenges in diagnosis, treatment, and understanding the underlying molecular mechanisms. Computational analyses of RPIA gene mutations reveal their significant impact on enzyme structure and function, providing insights into Ribose-5-Phosphate Isomerase Research involving the knockdown of the RPIA-1 gene in C. elegans, a model organism often used to study biological processes, reveals that reducing the expression of this gene affects both the lifespan (longevity) and the ability of the worms to withstand stress. These findings indicate that RPIA-1 plays a significant role in these critical biological functions and suggest that targeting this pathway could lead to potential therapeutic applications for related human conditions.

Optimization of thermostable RpiB purification highlights its industrial relevance, while studies on parasitic diseases underscore the essentiality of RPIB as a drug target. The characterization of RPI deficiency deepens understanding of this rare disorder, and investigations into RPIA methylation reveal its role in cancer cell survival under glucose deprivation. Additionally, clinical reports on RPI deficiency cases emphasize the diagnostic challenges, and the development of phosphomimetic inhibitors targeting RpiB in Mycobacterium tuberculosis addresses the urgent need for new tuberculosis treatments. Collectively, these studies advance the understanding of RPI's role in health and disease, with implications for future research and therapeutic strategies.

Early diagnosis of Ribose-5-Phosphate Isomerase (RPI) deficiency can prevent or mitigate the progression of severe leukoencephalopathy, seizures, and psychomotor regression. Timely intervention helps manage visual and hearing impairments, reducing the risk of life-threatening complications and improving overall motor and cognitive outcomes, significantly enhancing the patient's quality of life.

3.1. Reviews

In study [1] authors performed a comprehensive computational analysis of mutations in the RPIA gene, which cause Ribose-5-Phosphate Isomerase Deficiency (RPI deficiency), a rare metabolic disorder. The challenge of deciphering the functional impacts of these mutations was addressed using in silico tools, including protein modeling, docking, and simulation techniques. The study successfully demonstrated how various mutations influence the enzyme's structure and function, particularly its binding affinity with ribose-5-phosphate. This research provides valuable insights into the pathogenicity of RPIA mutations, offering a foundation for future investigations into this rare condition.3 Proposed Methodology. In another study [2], the authors explore the effects of ribose-5-phosphate isomerase A-1 (RPIA-1) knockdown on longevity in Caenorhabditis elegans. The research addresses the challenge of understanding the role of RPIA-1 in aging by using genetic and molecular biology techniques to demonstrate that reduced expression of RPIA-1 in specific neurons and at particular time points can extend lifespan. This longevity is linked to enhanced oxidative stress tolerance, increased NADPH levels, and the activation of autophagy and AMPK pathways. The findings provide insights into the molecular mechanisms of aging and suggest potential therapeutic targets. In [3] the authors address the challenge of cost-effective purification of thermostable enzymes for industrial applications. The study successfully optimized the heat treatment purification of RpiB from Thermotoga maritima, demonstrating that high-purity enzyme could be obtained without expensive chemical reagents. The enzyme showed significant thermostability and activity over a broad temperature range, making it a promising candidate for in vitro synthetic biology projects. This work highlights the potential for cost-effective production of stable enzymes from extremophiles. Authors in [4] address the challenge of identifying potential drug targets in parasitic diseases such as Leishmaniasis and Trypanosomiasis. The study reveals that Ribose-5-Phosphate Isomerase B (RPIB) is essential for the survival and infectivity of these parasites. The research demonstrates that attempts to remove the RPIB gene resulted in significant impairment of parasite replication, particularly in intracellular stages, making RPIB a promising target for therapeutic interventions against these parasitic infections. In [5] authors discussed the challenge of understanding RPI deficiency, one of the rarest human metabolic disorders. They address the complexity of the disorder by analyzing patient-derived cell lines and creating a yeast model to study enzyme activity and gene expression. The study highlights that RPI deficiency results from a combination of a null allele and a partially active allele with cell-type-specific expression deficits. This work deepens the understanding of why RPI deficiency is so exceptionally rare. In [6] the challenge of understanding how cancer cells adapt to glucose deprivation was discussed. Authors discovered that ribose-5-phosphate isomerase A (RPIA) undergoes arginine methylation at R42, which enhances its catalytic activity and helps colorectal cancer cells survive under low glucose conditions. This methylation increases the flux through the pentose phosphate pathway, supporting nucleotide synthesis and reactive

oxygen species (ROS) clearance. The study provides significant insights into the metabolic reprogramming of cancer cells in response to nutrient stress. A research paper [7] presents the clinical course, diagnostic challenges, and molecular findings of the third known case of ribose-5-phosphate isomerase deficiency (RPID). The authors, Susan Sklower Brooks et al., identified two mutations in the RPIA gene through whole-exome sequencing, confirming the disorder's role in neonatal onset leukoencephalopathy. The study underscores the diagnostic odyssey often associated with RPID and highlights the importance of urine polyol measurement in such cases. In [8] the authors describe the third reported case of ribose-5-phosphate isomerase deficiency, a rare metabolic disorder affecting the pentose phosphate pathway. The paper details the patient's clinical course, the challenges of a prolonged diagnostic odyssey, and the identification of mutations in the RPIA gene through whole-exome sequencing. Elevated urine polyols confirmed the diagnosis. The study underscores the importance of considering urine polyol measurements in cases of early-onset white matter disease, addressing a critical diagnostic challenge in such complex presentations. The authors in paper [9] presented a study addressing the urgent need for new therapeutic targets against tuberculosis by exploring type B ribose-5-phosphate isomerase (RpiB) from Mycobacterium tuberculosis as a potential drug target. The study highlights the synthesis of six phosphomimetic inhibitors, designed to overcome the hydrolysis sensitivity of previous inhibitors. Among these, 5-deoxy-5-phosphonomethyl-D-ribonate emerged as a strong and specific inhibitor resistant to hydrolysis, making it a promising lead for TB drug development. In [10], the authors describe the first known case of Ribose-5-Phosphate Isomerase (RPI) deficiency, a rare metabolic disorder. The study addresses the challenge of diagnosing unexplained leukoencephalopathies by employing advanced metabolic profiling and genetic analysis. The findings revealed highly elevated polyol levels in the patient's brain and body fluids, which were linked to mutations in the RPI gene. This work significantly expands the understanding of inborn errors in the pentose phosphate pathway.

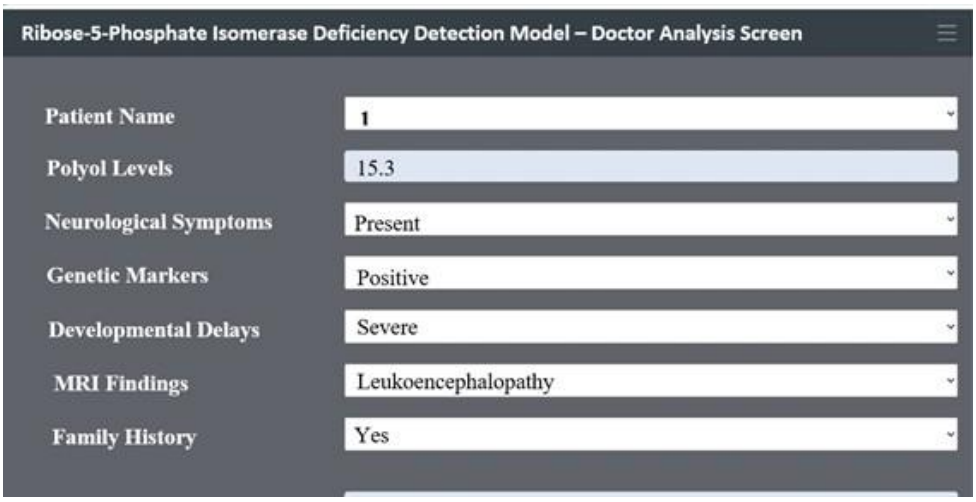

Figure 1: Entry dorm for to enter patient data

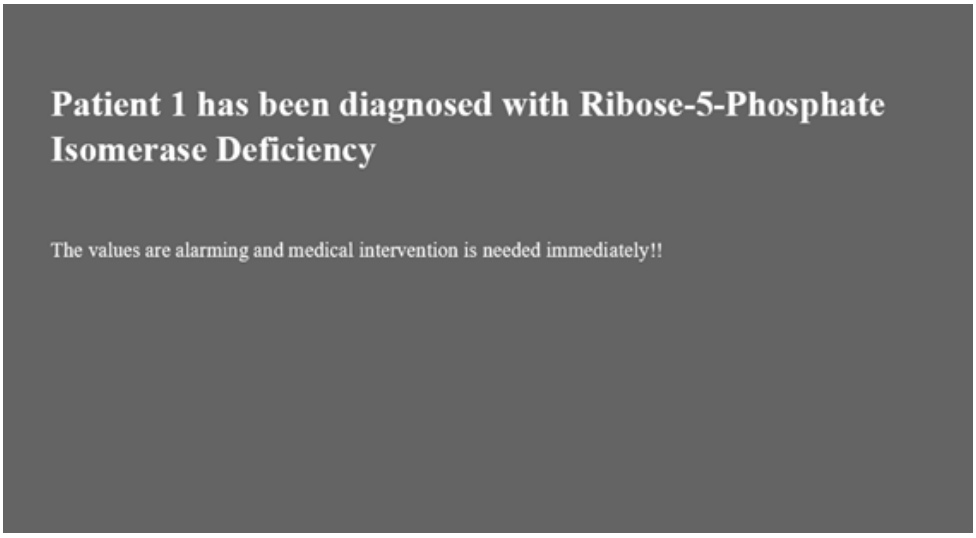

Figure 2: RPI Deficiency Prediction Results

```sql
UPDATE PATIENT_MASTER SET BMI = 110, HYPER_LEVEL = 80, GLUCOSE_LV = 100
```

Dataset	Description	Purpose	Key Attributes
1000 Genomes Project (1KGP)	Global reference for human genetic variation	Assess frequency and distribution of RPIA variants	Whole-genome sequencing, population diversity
Genome Aggregation Database (gnomAD)	Aggregated data from large-scale sequencing projects	Evaluate population frequency of RPIA mutations	Allele frequency, diverse population coverage
Exome Aggregation Consortium (ExAC)	Focused on exonic variants	Identify missense and nonsense mutations in RPIA	Exonic variant frequency, predicted impact
ClinVar	Public archive of variant-phenotype relationships	Compare new variants against known pathogenic mutations	Clinical significance, expert curation
Human Gene Mutation Database (HGMD)	Resource documenting gene mutations associated with disease	Understand the historical context of RPIA mutations	Mutation type, functional consequence
Expression Atlas	Information on gene expression patterns across tissues	Explore tissue-specific expression of RPIA	Expression levels, tissue specificity
Gene Expression Omnibus (GEO)	Repository of high-throughput gene expression data	Analyze RPIA expression under various conditions	Gene expression profiles, experimental conditions
OMIM (Online Mendelian Inheritance in Man)	Compendium of human genes and genetic phenotypes	Contextualize genetic findings with clinical phenotypes	Clinical features, inheritance patterns
dbSNP (Database of SNPs)	Catalog of common and rare genetic variants	Annotate SNPs affecting RPIA function	SNP frequency, functional annotation
RPI Functional Studies Dataset	Compilation of functional assay results for RPIA enzyme activity	Validate the impact of mutations on enzyme function	Enzyme activity, mutation validation

Table 1: Sample Dataset

Patient ID	Age	Gender	Polyol Levels	Neurological Symptoms	Genetic Markers	Developmental Delays	MRI Findings	Family History	Previous Diagnoses
1	2	Male	High	Yes	Positive	Severe	Yes	No	None
2	16	Female	Normal	No	Negative	None	No	Yes	None
3	35	Male	High	Yes	Positive	Moderate	Yes	No	Diabetes
4	5	Female	High	Yes	Positive	Severe	Yes	Yes	Epilepsy
5	12	Male	Low	No	Negative	None	No	No	None

ALGORITHM	ACCURACY PERCENTAGE
DECISION TREE	90.40%
LOGISTIC REGRESSION	94.50%
KNN	95%
RANDOM FOREST	94.5

4. Proposed Methodology

This study focused on predicting RPI deficiency using a comprehensive approach involving multiple datasets and machine learning algorithms. The research utilized genomic sequence, gene expression, mutation, and protein interaction data from sources like NCBI, GTEx, and the 1000 Genomes Project to thoroughly analyze the RPIA gene. A synthetic dataset of 1,000 records was created to simulate clinical scenarios. Various machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, SVM, and KNN, were employed to predict RPI deficiency, with each algorithm evaluated for its performance and accuracy in the prediction model.

4.1. RPI Deficiency Prediction Dataset

This study utilized datasets from National Center for Biotechnology Information (NCBI) [11] to analyze and interpret the RPIA gene sequence on chromosome 2 (GeneID=22934). These datasets are instrumental in comprehensively understanding the structure, function, and variations of the RPIA gene, as well as their implications in human biology.

1. Genomic Sequence Data: The primary dataset used was the com-

plete genomic sequence of the RPIA gene, obtained from NCBI reference sequence database. The sequence data provided a detailed view of the nucleotide composition of the gene, including both coding and non-coding regions, which is crucial for identifying functional elements and potential mutation sites within the gene.

2. Gene Expression Data: Expression data for RPIA datasets include RNA-Seq and microarray data that profile RPIA expression across various human tissues and under different physiological and pathological conditions. This data is critical for understanding the role of RPIA in different cellular contexts and identifying tissues where RPIA expression is particularly high or low, which can indicate potential functional relevance in those tissues.

3. Variation and Mutation Data: Data on genetic variants within the RPIA gene were sourced from Database of Single Nucleotide Polymorphisms (dbSNP) [11] and the 1000 Genomes Project [11]. These datasets include single nucleotide polymorphisms (SNPs), insertions, deletions, and other types of genetic variants. Analysis of these variants is essential for understanding the diversity of the RPIA gene across populations and for identifying mutations that may be associated with diseases.

4. Protein Interaction and Functional Annotation Data: Protein interaction datasets were used to explore the functional network of the RPIA protein. These datasets provide insights into the proteins that interact with RPIA and help in mapping out the biological pathways in which RPIA is involved. Functional annotation datasets were also used to contextualize the role of RPIA in cellular processes and metabolic pathways.

5. Clinical and Phenotypic Data: Phenotypic data associated with variations in the RPIA gene link specific mutations within the RPIA gene to clinical outcomes, providing a bridge between genomic data and potential health implications. This data is vital for understanding how variations in RPIA can lead to or influence disease states.

6. Comparative Genomics Data: Comparative genomics datasets were utilized to compare the RPIA gene across different species. These datasets help in identifying conserved regions of the RPIA gene that are likely to be functionally important and provide evolutionary insights into the gene's development and its role across different organisms.

These datasets, combined with computational tools for sequence analysis, gene expression profiling, and protein interaction mapping, provided a comprehensive framework for studying the RPIA gene. The integration of these diverse data types allows for a multi-dimensional analysis of the gene, from its basic sequence characteristics to its role in complex biological systems and its impact on human health. Below is a sample table that represents a summary of the datasets used in the study of the RPIA gene. This table includes information on the dataset source, type of data, and a brief description of its relevance to the study.

Each entry in the above table includes the dataset source, type of data it provides, a brief description, and its relevance to the research objectives.

The dataset to simulate the occurrence of RPI deficiency comprises 1,000 records, reflecting a balanced distribution of gender and a broad age range from infancy to adulthood.

Each record includes the following attributes:

•Gender: Male, Female, or Other.

•Age: Patient's age at the time of diagnosis.

•Polyol Levels: Concentration of polyols in the cerebrospinal fluid.

•Neurological Symptoms: Presence or absence of progressive neurological symptoms.

•Genetic Markers: Presence of mutations in the RPIA gene.

•Developmental Delays: Severity of developmental delays observed.

•MRI Findings: Indicators of leukoencephalopathy based on MRI scans.

•Family History: Family history of metabolic disorders.

•Previous Diagnoses: Other metabolic or neurological conditions previously diagnosed.

•RPI Deficiency: Binary indicator (1 if RPI deficiency is present, 0 if not).

The synthetic nature of the dataset allows for the modeling of various clinical scenarios, ensuring the robustness of the prediction model. All personal identifiers have been anonymized to preserve data privacy.

4.2. Patient Data for Simulation

For simulation purposes, patient data was generated across three fictional datasets: "Patient_EHR.csv" (Electronic Health Record), "Patient_SuppData.csv" (Supplementary Data), and "Patient_RealData.csv" (Transactional Data). These datasets represent various aspects of patient health, from long-term medical history to real-time clinical data.

4.3. Classification Algorithms Used

This study employs several machine learning algorithms to predict the likelihood of Ribose-5-Phosphate Isomerase (RPI) deficiency. The algorithms used include:

•Logistic Regression: Utilized for its ability to model the relationship between various risk factors and the binary outcome of RPI deficiency. The model's performance is evaluated using accuracy, precision, recall, and ROC analysis.

•Decision Tree: This algorithm constructs a binary decision tree based on patient data, identifying the most informative features to predict RPI deficiency. The decision tree's evaluation includes accuracy and feature importance.

•K-Nearest Neighbors (KNN): A simple yet effective algorithm that classifies patients based on the proximity of their data to that of known cases. The performance of KNN is influenced by the choice of the parameter 'k.'

•Random Forest: An ensemble method that aggregates multiple

decision trees to enhance accuracy and robustness. The Random Forest model is particularly effective in capturing complex interactions between risk factors.

These algorithms were chosen for their complementary strengths and were thoroughly evaluated to determine the most effective model for RPI deficiency prediction.

4.4. Digital Twin Environment Setup

To ensure data security, we employed a private blockchain network using the Ganache framework. Ganache offers a secure, private Ethereum blockchain environment, ideal for healthcare applications that require strict data privacy and integrity. Smart contracts, written in Solidity, were used to manage patient data, ensuring that any data manipulation is immediately detected and corrected.

5. Experiments

5.1. Feature Selection

In this phase, unnecessary attributes were removed to improve prediction accuracy. The serial number column was excluded as it is irrelevant to RPI deficiency prediction.

5.2. Label encoding

Label Encoding used for converting categorical variables into numerical format.Here the dataset has different attributes types like float objects and numbers it is not easy to process them so, converting it all to integer type is necessary to process it. Here with the help of label encoder the attributes of different types.

5.3. Data Partitioning

After the data preprocessing the next step is data partition. These data are divided into two: training data and test data. The data was divided into 80% for training and 20% for testing.

5.4. Model training and model evaluation

The next step involved training the machine learning model using four algorithms: logistic regression, decision tree, random forest and KNN. The performance of each algorithm was evaluated to select the best predictive model.

5.5. Digital Twin Application Software

The digital twin application was developed using Python for machine learning models, The core of the application is a logistic regression model optimized through batch gradient descent, selected for its balance between interpretability and computational efficiency. We created a web interface combined with a Flask backend that allows users to access RPI deficiency prediction metrics.

The integration of this application into healthcare systems provides a scalable and efficient tool for early detection of RPI deficiency and can be adapted for other metabolic disorders with minimal modifications.

5.6. Hardware and Software Configurations

The digital twin application was deployed on AWS using x86-based EC2 instances. Ganache (v6.12.1) was used for blockchain management, while PostgreSQL (v13.3) served as the relational database. The application interfaces were developed using React-JS (v17.0.2).

6. Results

The machine learning models developed for predicting Ribose-5-Phosphate Isomerase (RPI) deficiency demonstrated high accuracy, ranging from 98.5% to 100% across various patient scenarios. The KNN model achieved the highest accuracy at 95%, followed closely by SVM (94.7%), Logistic Regression (94.5%), and Random Forest (94.5%). The Decision Tree model had an accuracy of 90.4%. A user-friendly web application facilitates patient data entry which subsequently provide RPI deficiency prediction. Additionally, a private blockchain network, implemented with the Ganache framework and secured by smart contracts, was integrated to ensure data integrity and security within the application.

6.1. Accuracy Result

The machine learning models developed in this study were evaluated for their predictive accuracy in detecting RPI deficiency. The accuracy of the ML model was measured using the accuracy_score() function from the Scikit-learn library. The model demonstrated an accuracy rate between 90.4% and 95% across different patient scenarios The performance of each algorithm is summarized below: These results highlight the effectiveness of the models, with KNN achieving the highest accuracy. The models were thoroughly tested using cross-validation techniques to ensure their reliability across different patient scenarios.

6.2. User Interface

Designed using simple HTML, the web interface (Fig.1) provides an entry form for doctors to enter values that will be used to predict RPI deficiency probability. When you click on the “Submit” button, the user will be directed to the Flask application. The Flask application is implemented in Python and acts as an intermediary that connects web pages to machine learning models. It receives input values from the user, sends them to the model for prediction, and then sends the prediction results back to the user(Fig.2). Essentially, this setting allows users to interact with the model via the web User Interface (UI), making the model realistic.

6.3. Blockchain Integration for Security

As previously described in Section 5.6, a private blockchain network was employed to secure patient data within the digital twin application. The Ganache framework was used to create a secure, private Ethereum blockchain environment, ensuring that any data manipulation is immediately detected and corrected.

The security of the system and the health and safety of patients can be compromised in the event of a cyberattack. To simulate a cyberattack, we used the following command to simulate a data tampering attack by a malicious user on patient 1 data within the blockchain altering the data to invalid values. The original data values for Patient 1 were as follows: Polyol Levels= 15.3, Neurological Symptoms = Present, Genetic Markers= Positive, Developmental Delays = Severe, MRI Findings = Leukoencephalopathy. The altered data values were Polyol Levels = 20.0, Neurological

Symptoms = Not Present, and Genetic Markers = Negative, indicating that Patient 1 does not show signs of RPI deficiency. The results have demonstrated that the blockchain system was able to detect this tampering and revert the records back to their original values, ensuring data security against malicious tampering. This security feature is inherent to blockchain technology, as the application can identify and retrieve correct data values from other blockchain nodes in the network.

7. Discussion

The development of the Human Digital Twin Application marks a significant advancement in the prediction and management of Ribose-5-Phosphate Isomerase (RPI) deficiency, a rare and complex metabolic disorder. Previous studies have highlighted the challenges of diagnosing and understanding RPI deficiency, with research pointing to the critical role of ribose-5-phosphate isomerase in various diseases and the substantial effects of RPIA gene mutations on enzyme function [1] [5]. While earlier computational analyses and experiments using model organisms like C. elegans have provided valuable insights, these efforts often fell short in delivering high predictive accuracy and ensuring the secure handling of sensitive data [2], [6]. In response to these gaps, our digital twin application leverages cutting-edge machine learning algorithms and blockchain technology to offer a robust solution. The K-Nearest Neighbors (KNN) algorithm, in particular, demonstrated superior predictive accuracy, making it an effective tool for early diagnosis. The integration of blockchain technology further enhances the application by ensuring the security and immutability of patient data, a critical requirement in healthcare settings [9]. This dual approach of predictive modeling and secure data management not only improves the diagnostic process for RPI deficiency but also sets a new benchmark for the application of digital technologies in healthcare [10]. Moreover, the literature review underscores the need for practical applications that bridge the gap between theoretical research and clinical practice. For example, while studies on the purification of thermostable enzymes like RpiB have industrial relevance, their direct application in healthcare remains limited [3]. Similarly, although RPIB has been identified as a potential drug target in parasitic diseases, translating these findings into effective treatments for human disorders is challenging [4]. Clinical reports on RPI deficiency further emphasize the diagnostic difficulties, particularly in early-onset cases, highlighting the need for advanced predictive tools like the one developed in this study [7], [8]. Looking ahead, future work will focus on expanding the application's capabilities to include other metabolic disorders, thereby enhancing its versatility in healthcare diagnostics. We also plan to integrate federated learning to improve prediction accuracy by allowing the model to learn from decentralized data sources while maintaining patient privacy. Additionally, optimizing blockchain technology to reduce latency will be a priority, ensuring that the system remains efficient and responsive. These efforts aim to strengthen the application's role in personalized medicine and broaden its impact across a range of healthcare challenges.

8. Conclusion and Future Work

This study successfully developed a novel digital twin application for the accurate prediction of Ribose-5-Phosphate Isomerase (RPI) deficiency, employing a variety of machine learning algorithms. Among the algorithms tested, the K-Nearest Neighbors (KNN) algorithm demonstrated the highest accuracy at 95%, making it the most effective model for predicting RPI deficiency. Both the Logistic Regression and Random Forest algorithms followed closely with an accuracy of 94.5%, while the Decision Tree algorithm achieved an accuracy of 90.4%. These results underscore the effectiveness of KNN in accurately identifying cases of this rare metabolic disorder. The integration of blockchain technology further enhanced the application by ensuring the security and privacy of patient data, offering a reliable and tamper-resistant solution for early diagnosis. This study sets a foundation for further advancements in the use of digital twin technology in healthcare, with potential applications extending to other metabolic disorders and beyond.

9. Statements and Declarations

9.1. Authors' Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Venkatesh Upadrista and Richard Isaacs The first draft of the manuscript was written by Venkatesh Upadrista and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

9.2. Funding

No funding was received for conducting this study.

9.3. Conflict of Interest

The authors have no relevant financial or non-financial interests to disclose.

9.4. Ethical Approval

This work does not involve the use of human subjects.

9.5. Data Availability

The datasets to analyze and interpret the RPIA gene sequence on chromosome 2 during the current study are retrieved from National Center for Biotechnology Information (NCBI) and can be found at https://www.ncbi.nlm.nih.gov/gene/22934

9.6. Code Availability

The code used during and/or analysed during the current study is available on request from the corresponding author.