مجلة جامعة صحار للأعمال المستدامة
Sohar University Journal of Sustainable Business

Sohar University Journal of Sustainable Business
e-ISSN: 3007-5572 | Volume 1, Issue 1, 57-71
DOI: 10.69983/SUJSB/0101/05, January 2025

INTELLIGENT DATA SYSTEMS FOR SUSTAINABLE PREVENTION AND MANAGEMENT OF G6PD DEFICIENCY: A MARKOV CHAIN MODEL APPROACH

Divya Rani

Medanta Super Speciality Hospital, Jharkhand, India

ABSTRACT

Adopting sustainable healthcare practices is crucial for effectively managing genetic disorders. Despite significant advances in our understanding of the genome over the last decade, there are notable limitations in epidemiological and analytic approaches for investigating the effects of common chronic genetic diseases like Glucose-6-phosphate dehydrogenase (G6PD) deficiency in Oman. This study aims to use pedigree charting and Markov Chain Model analysis to determine the pattern of gene distribution for G6PD deficiency in the family trees of Omani nationals. It also identifies the pattern of gene distribution for G6PD deficiency by synchronizing medical data, including the genetic history of families, and applying Markov's model to find the family generation at which G6PD deficiency becomes normal homozygote and reaches an equilibrium state. Based on the analysis using the Markov Chain Model, it is observed that after a few generations, the population attains a normal homozygote state. However, this equilibrium is not consistent across all families. In some families, the generation becomes normal after 18 stages, while in others it becomes normal after 12 stages. Leveraging intelligent data systems can reduce long-term healthcare costs and improve population health, ultimately fostering economic sustainability in Oman.

Keywords: Markov Chain, G6DP, genetics, Anemia, Oman

Paper Type: Research paper

INTRODUCTION

Sustainable healthcare in Oman focuses on creating a resilient health system that meets current needs without compromising future generations. Emphasizing preventive care and public health initiatives helps reduce the burden on healthcare facilities. Technological advancements, like telemedicine and electronic health records, enhance efficiency and accessibility. Oman's population is distinguished by large family sizes, advanced maternal and paternal ages (Rajab et

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)
©Sohar University Journal of Sustainable Business

al., 2013). As a result, the presence of gene-related defects, various genetic anomalies, and syndromes is extremely high in Oman's general population (Islam, 2012). In a new report, Oman's Ministry of Health revealed that citizens in the country inherit more than 300 diseases from their parents.(Al Mawali et al., 2022). The likelihood of a child developing a genetic disease such as G6PD is 3.5 to 4.7 per 1,000 live births, according to the report, that the Sultanate has a newborn birth defect or any of the genetic diseases at a rate of 5 to 7 percent, compared to a global average of 4.4 percent (Al-Sheryani et al., 2023). According to statistics, congenital disorders account for 37% of newborn deaths and premature births (Rajab et al., 2013). Genetic disorders account for 10% of neonatal mortality and 52% of mortality in older children (Al-Sheryani et al., 2023). The prevalence of these genetic disorders, comparable to rates in other Arab countries, is highly concerning.

Managing genetic diseases can be challenging, as treatments are often time-consuming and costly, and currently, no definitive cure exists. The patient must endure the trauma of taking medicine or receiving blood transmission for the rest of his or her life, or he or she must undergo organ transplantation. According to studies by Center of Arab Genomic studies, 10% of Omani nationals carry genes for sickle cell anemia, 3% carry mutations to genes with beta thalassemia and 45% has genes for alpha thalassemia (Al-Riyami & Ebrahim, 2003). Also, down syndrome is as common among 1 in every 350 live births.

The most common genetic disease being Glucose-6-phosphate dehydrogenase (G6PD) deficiency accounting for 12% in women and 28% among men. In total, a skyscraping percentage of around 60% of Omani nationals are carrying genes for genetic disorders or, suffering from different genetic anomalies (Al-Maawali, 2022)(Rajab et al., 2013). It is a significant challenge for the country because 60% of its 2.5 million-person at least have one type of genetic blood disorder deficiency exists in the population. (Sulaiman et al., 2001). Health officials emphasized because inter-tribal marriage is so common in Oman, premarital tests should be performed, which would result in a significant decrease in blood-related genetic disorders (Bruwer et al., 2022).

Many people have advocated for mandatory premarital tests in order to protect their children from inheriting genetic disorders like G6PD deficiency (Alangari et al., 2023). In general, no public health measures are aimed at avoiding congenital and genetic diseases, which are also constrained by cultural and legal constraints. Some countries have begun low-cost prevention programs, such as premarital carrier screening, for certain common genetic disorders G6PD deficiency. In this region, strategies for genetic disorder prevention should include the community genetics integration into primary care, education, and expansion of existing specialized genetic services. The goal of this research is to aware the Omani nationals about the G6PD (Glucose-6-phosphate dehydrogenase**)** deficiency and how to prevent them by finding the equilibrium state/normal state of their family generations by using Markov's Chain Model and also to highlight the use and simplicity of Markov chain modeling in the study of G6PD disorders. This study focuses on integrating advanced data systems and preventive measures to enhance public health. Leveraging technologies like intelligent data systems and Markov Chain models, Oman aims to identify and mitigate the risks associated with genetic disorders. This approach not only improves patient outcomes by enabling early diagnosis and tailored interventions but also reduces long-term healthcare costs. By promoting preventive care, efficient data

management, and community education, Oman advances toward a more sustainable and economically viable healthcare system, ensuring better health for future generations.

BACKGROUND OF THE STUDY

Oman was historically the main business port of the Arabian Gulf region. Because of its extensive trade activities, the Omani people have a diverse social and ethnic background. According to the survey, which was designed, carried out, and funded by the Sultanate of Oman's Ministry of Health, the statistics are very concerning.(Al-Riyami & Ebrahim, 2003). According to the Oman Hereditary Disorder Association (OHDA), 10% of the Omani population is genetically predisposed to at least one dangerous blood disorder.

According to a recent study conducted in the Sultanate of Oman, G6PD (Glucose-6-phosphate dehydrogenase**)** is very common in society. It is reported to occur 17.9% more frequently in children under the age of five (25 per cent with males and 10 per cent with females) (Rajab et al., 2013). Males were more than twice as likely as females to have G6PD deficiency in all regions of the Sultanate. Based on the rates listed above, it is possible to calculate that in the Sultanate as a whole, the statistics are as in table 1:

Table 1: Genotype traits in the Sultanate of Oman

#	Genetic Trait	Population Statistics
1	G6PD deficiency	Under 5 years of age = 44,733 children
2	Sickle cell trait	14,306
3	Sickle cell disease	474
4	Beta-thalassaemia trait	5393
5	Bets-thalassaemia major	175
6	Mean Hb level	Children under the age of five had 10.9 g/dl, males had 10.89 g/dl, and females had 10.99 g/dl.
7	Anemic	Under 5 years of age = ½ of the children
8	Mild anemia was predominant	46% (sixty-five percent of children aged 0 to or equal to one year)
9	Moderate	4%
10	Severe anemia	0.2%

Source: *https://ghdx.healthdata.org*

With these frightening statistics spreading throughout society, the Ministry of Health (https://www.moh.gov.om) is attempting to identify and eliminate the causes. Pre-marital screening appeared to be the prudent course of action in order to avoid having more blood-disease patients. Premarital screenings are recommended for couples to protect their future children from genetic health problems.

G6PD DEFICIENCY

G6PD deficiency is genetically transmitted from one or both parents to a child. G6PD (Glucose-6-phosphate dehydrogenase) deficiency is a hereditary condition that is inherited in an X-linked recessive manner marked by low G6PD enzyme levels. Arab countries have a high prevalence of genetic diseases, with a higher prevalence in Saudi Arabia (39.8 percent), Syria (30 percent), and Oman (29 percent) than in other Arab countries(Natarajan & Joseph, 2021). Only a few studies have been conducted in Arab countries to investigate the prevalence of G6PD mutations and their functional role in illness

LITERATURE REVIEW

Comparing various genetic models has been an active area of research due to the multitude of models in use. Since the late 1980s, scholars have developed Markov chain models of genetic algorithms to enhance the understanding of these genetic models (Munkhammar & Widén, 2018). Most genetic algorithms can be represented as Markov chains over populations. This study reviews a well-established Markov chain model, examines the G6PD research report, and constructs new Markov chain models. A notable application of the Markov Chain Model was to study gene distribution in albinism, proposing that one parent must be a typical homozygote with selective pressures applied (Sun et al., 2023). It was concluded that after six generations, all individuals would be normal homozygotes, leading to the extinction of the albino gene in the population. Similarly, a discrete-time Markov chain model was employed to tackle the issue of infectious diseases, observing that hepatitis B became more infectious over time compared to tuberculosis and HIV despite the low initial infection probability.

Digital microfluidics (DMF) has also been utilized to quantitatively and kinetically evaluate G6PD activity. In a study involving 86 samples (Paxinou et al., 2021), DMF results were compared with two reference methods, achieving a reproducibility of 3.8% over 5 days across two operators and instruments (Cornel et al., 2024). Hidden Markov Models (HMMs) have been developed for various computational biology challenges, including multiple sequence alignment, homology detection, protein sequence categorization, and genome annotation (Sun et al., 2014). The principles, techniques, practices, and policies influencing European population genetic screening programs were examined from professional and scientific perspectives, aiming to raise awareness among healthcare professionals and policymakers (King et al., 2022). Methods included analyzing current professional regulations, regulatory frameworks, and relevant documents. HMMs have also been used to identify patterns in genomic sequences, assessing model performance by varying the number of states in the Markov model and using several performance indicators, utilizing publicly available genetic data (Grzymski et al., 2020).

A weighted Markov chain was introduced for predicting future incidence states, using standardized self-coefficients as weights based on infectious disease incidence properties, with the Markov chain Monte Carlo approach examining these characteristics to maximize long-term benefits. The applications of Markov chains have been highlighted, demonstrating their usefulness in applied mathematics. The main causes of G6PD deficiency have been identified as

the absence of a screening program, uncertainty regarding jaundice onset and progression, rapid increases in bilirubin levels, and early hospital discharge without proper monitoring of high-risk infants (Lacaze et al., 2022), advocating for screening programs and public health initiatives including education, surveillance, and training.

Studies on G6PD deficiency polymorphisms in neonatal hyperbilirubinemia among Egyptian neonates provided insights into the Mediterranean variant's prevalence in Egyptian infants (Luzzatto et al., 2020). Genetic/genomic test information characteristics, such as predictability, immutability, and uniqueness, are crucial for developing information security policies and addressing "genetic exceptionalism" concerns, aiding in the integration of genetic/genomic information into healthcare (Ong et al., 2017). Combining genotype and deep phenotype data from Electronic Health Records (EHRs) has been shown to accelerate genetic diagnoses, with EHR narratives automating phenotype-driven clinical exome or genome analysis, promoting broader adoption of genomic medicine. A recursive hierarchical extension of hidden Markov models, called Hierarchical Hidden Markov Models (HHMM), has been introduced, efficiently estimating model parameters from unlabeled data using an effective estimation method (Mesa et al., 2016). Mechanisms in electronic health records have been developed for requesting a pharmacogenomics panel either preemptively or in response to a specific drug indication. These studies underscore the diverse applications and advancements in using Markov chain models and related methodologies to study genetic patterns, disease incidences, and the integration of genomic data into healthcare.

MANAGEMENT AND PREVENTIONS OF G6PD DEFICIENCY

Many factors, such as genes, environment, and lifestyle, are shared by families. These factors, when considered together, can provide ideas to medical problems that may run in families. Healthcare personals can identify whether a person, another family member, or future generations are at a higher risk of developing a specific condition by observing disorder patterns in relatives. Children, siblings and sisters, parents, aunts and uncles, nieces and nephews, grandparents, and cousins are all included in a detailed record. A family medical history is a document that contains information about a person's and his or her close relatives' health (Khalifa et al., 2021; Yoon et al., 2002). A person's family history can also help predict the risk of rarer conditions caused by single gene mutations, such as G6PD deficiency (Hassan et al., 2014). A person who has no family history of an illness, on the other hand, may be predisposed to developing that ailment. Knowing one's family medical history enables them to take precautionary measures to lower their risk (Beutler, 1996). According to healthcare providers, people who have a medical condition that runs in their family may benefit from regular checkups or testing.

Proposed method

The method adopted in this study involves several crucial steps. Initially, health data is stored through digital devices, ensuring accurate and efficient data collection. Following this, the digital

data undergoes processing to prepare it for further analysis. The processed data is then analyzed to derive meaningful results from the digital information. Once the analysis is complete, the pedigree of the suspect is developed, which involves mapping out the genetic relationships and potential inheritance patterns within the family. Finally, Markov's model is applied to the suspected family data to identify and predict genetic patterns and potential disease risks, thus providing a comprehensive approach to understanding genetic influences within the family.

A person's future health cannot be predicted based on his or her family's medical history; it only provides risk information. At least three generations should be included in the history. Gather the following information in an electronic sheet for each person: gender, year of birth, ethnicity, medical health condition, conditions affecting one's mental health, including a history of substance abuse, pregnancy complications such as miscarriage, stillbirth, birth defects, and infertility, age at diagnosis for each condition, and lifestyle factors including diet, exercise, and tobacco use. For deceased relatives, include the age at the time of death as well as the cause of death.

Intelligent data integration and decision support throughout the patient journey

The quality and availability of the data heavily influences decision-making in today's digitalized healthcare environment. Smart data integration may significantly improve decision-making quality, particularly in healthcare, where clinical decision-makers confront various impediments and challenges throughout the patient journey. Gathering enough information from the patient's family history which is used to screen for disease risk and to guide treatment because making decisions is a time-consuming procedure that is difficult to complete just via dialogue between patients and their health care professionals.

The patient e-card stores personal and patient information. The card stores previous diseases that were chronic and/or significant with diagnosis dates, long-term medications with dosages, allergies with dates of diagnosis, vaccines with dates, and surgical procedures with dates of operation, clinic name, and summary information (Bianco et al., 2013). The card also contains information about the patient's most recent examination and prescriptions.

On the basis of the available intelligent data, digital Genetic health information system process the data and generate valid information regarding the G6PD suspect as per the flow chart in Figure 1. After analyzing the information by the health professionals, a pedigree of a suspect drawn to derive a final conclusion related to G6PD deficiency.

Pedigrees

Based on a Specific family's electronic medical history, a "pedigree" of a family tree will be generated, assisting medical professionals in determining the risk of G6PD disease. A pedigree (or family tree) is used to present family history in a clear manner, making it easier to identify abnormal inheritance patterns. A pedigree chart shows the phenotypes of a specific gene in family members and ancestors. It uses standardized symbols to represent family members and

relationships as in figure 2. A pedigree can be created much faster than information can be recorded about a patient's family history in writing, and it enables disease patterns to emerge as the pedigree is drawn (Lacy, 2012). Because family histories change all the time, the pedigree can be easily updated on future visits. Most of the time, three generations are enough to produce reliable results. Inquiries about names, birthdays, and clinically relevant data. Progeny, miscarriages, or unintended childlessness. Based on the availability of random electronic health records, a pedigree genetic analysis (for 5 generations) is done.

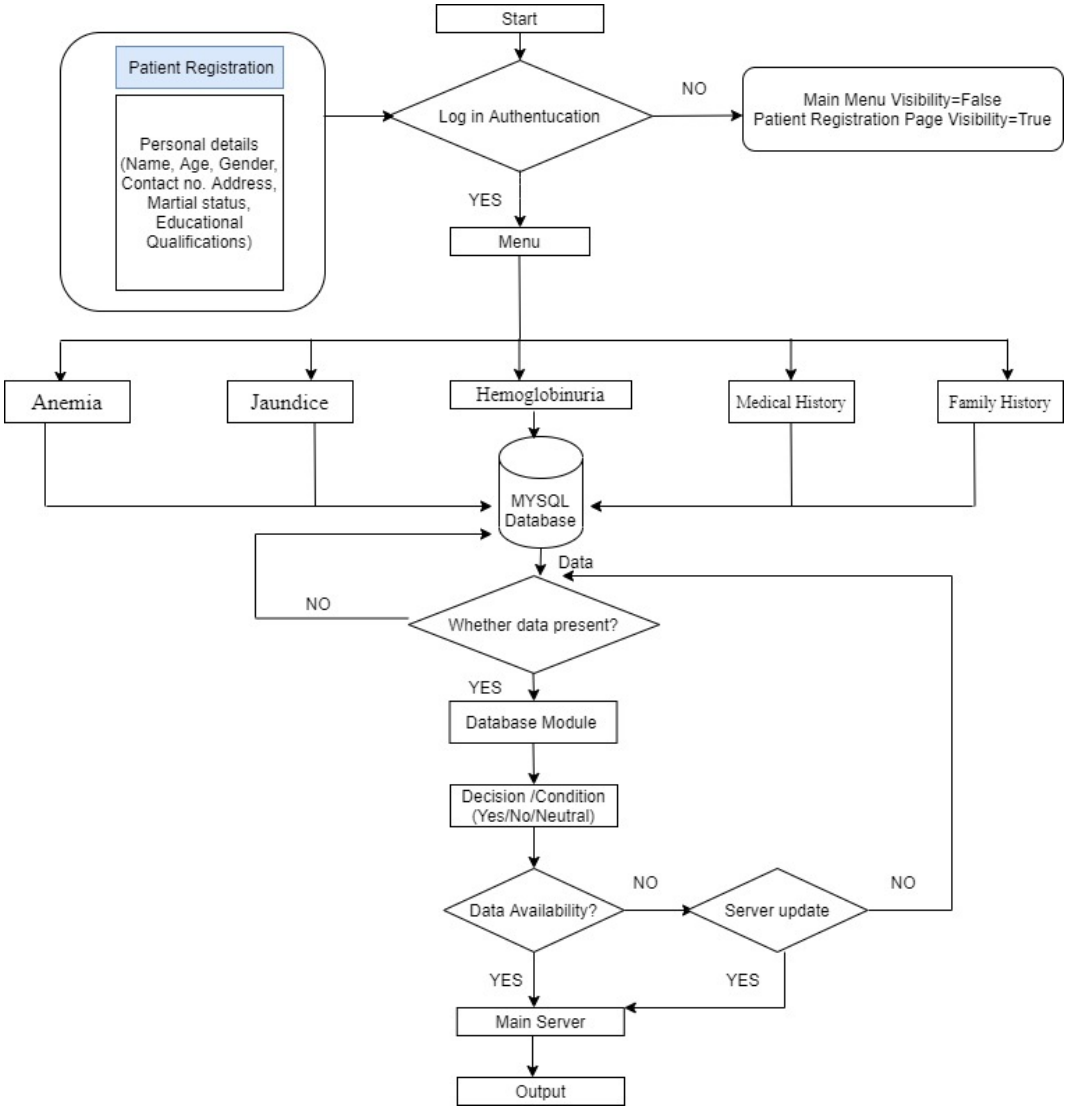

Figure 1: Process of digital information of G6PD deficiency suspects

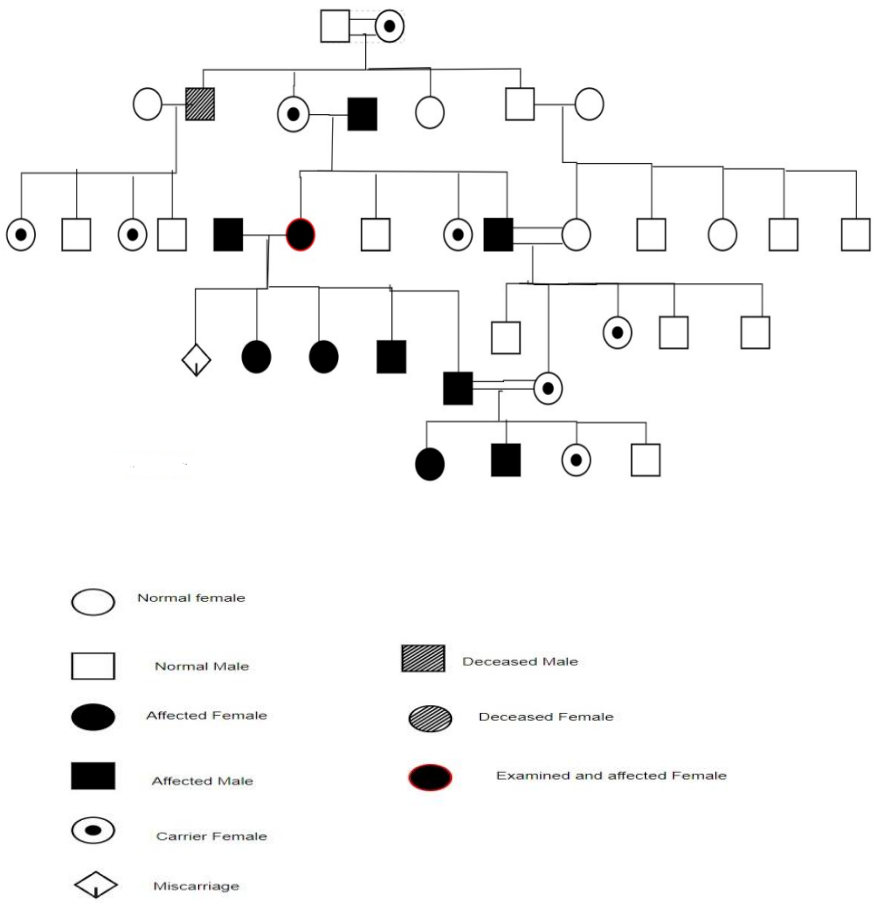

Figure 2: Family pedigree to determine inheritance pattern of the G6PD deficiency.

Pedigree chart helps doctors to analyze and assess the risk of G6PD deficiency. So as per the pedigree chart of the particular sample, analysis about G6PD deficiency are as:

The randomly examined female has consanguineous marriages in family with history of miscarriages, chronic anemia, jaundice and sclera confirmed diagnosis of G6PD deficiency in almost every generation.

The generational analysis of pedigree in consanguineous marriages reveals distinct patterns across five generations. In the first generation, there is a consanguineous marriage, which sets the stage for the genetic analysis of subsequent generations. The second generation includes one deceased male and one carrier female. In the third generation, there are carrier females and an affected female who has been examined. Moving to the fourth generation, both males and females are affected. Finally, in the fifth generation, as a result of continued consanguineous marriages, there is a probability of affected males and females as well as carrier females, as indicated by the pedigree chart.25% of son=healthy.

50% of children=may get affected

25% of daughter=carrier.

By doing the test during the engagement period has been reported as one of the significant reasons for marriage continuation despite abnormal premarital screening and counseling results. As a result, more effort is required to test at an early stage.

Markov Chain Model

The pattern of gene distribution on G6PD is determined using a six-state Markov Chain Model. A Markov chain, also known as a Markov process, is a stochastic model that describes a series of potential events in which the probability of each occurrence is determined only by the state obtained in the preceding event. In simpler terms, it is a process for which future outcomes can be predicted only based on its current state, and, more significantly(Khalifa et al., 2021), such predictions are just as good as those that could be made knowing the process's whole history. A Markov chain model is one of the most important tools in stochastic processes (Pes & Dore, 2022). A "Transition Matrix" and an "Initial Distribution" are required for the formation of a Markov chain.

At time n, the transition matrix is the matrix $P(n) = (Pi\ j(n))$, which means that the $(i\ j)^{th}$ element of $P(n)$ equals $Pi\ j(n)$. The transition matrix satisfies:

(i) $Pi\ j(n) \geq 0\ \forall i, j$ (the entries are non-negative)
(ii) $\sum j\ Pi\ j(n) = 1\ \forall i$ (the rows sum to 1)

A stochastic matrix is any matrix that meets (i), (ii) above. As a result, the transition matrix is stochastic.

Let $(Xn)n \in N_0$ be a random variable sequence with values in S. In this case, n represents the time at which the state Xn happens. So, to finish the Markov chain, designate Ds as the collection of discrete distributions on S such that $Ds = (P = (Pi)_{j \in s:}\ P_i \geq 0, \sum_{j \in S} Pi = 1)$

In this case, the distribution as row vectors. So $P_0 = (P_{0i})_{i \in S} \in Ds$ the initial distribution of the chain $(Xn)_{n \in N0}$ *if* $P[X_0 = 1] = P_{0i}$, $\forall$ *states* $i \in s$.

A discrete time the square transition matrix *P. Pij* fully describes the Markov process, indicates that the likelihood of the system existing in state 1 at time $t = 0$ is implied by the probability of the system transitioning from state i to state *Pi (0) j*.

A probability state vector is a set of state probabilities that add up to one. A state is a mutually exclusive and exhaustive situation that a process can be in at any one time. If $Pi\ (0) = 1$, then $Pi\ (0) = 0$ for $i\ j = 1,2,3$ and for i, j. The transition matrix is a set of conditional probabilities for changing states.

The following is the Markov's chain model for six transitions:

XX = Normal Female
XX_1 = Carrier Female
X_1X_1 = affected Female
XY = Normal Male
X_1Y = affected Male

Table 2: Genetic scenarios

Male Genes	Female Genes	XX	XX1	X1X1	XY	X1Y
X1	X1	X1X	X1X		X1Y	X1Y
X	X1	XX	X1X		XY	X1Y
X	X	XX	XX		XY	XY
Y	X	XY	XY		XY	XY
X1	XX1	XX1	X1X1		XY	X1Y
X1	X1X1	X1X1	X1X1		X1Y	X1Y
XY	XX1	0	0.5		0	0
XY	X1X1	0	0	0.25	0.25	0
XY	XX	0	0	1	1	0
X1Y	XX1	0	0.5	0	0	0.5
X1Y	X1X1	0.25	0.25	0	0	0.25
X1Y	XX	0.5	0	0	0	0.5

The table 2 illustrates the various genotypes of offspring based on different combinations of male and female genes, showing the possible outcomes when specific genes from parents combine. When the male genes are X1 and the female genes are X1, the offspring will have the genotypes X1X (affected female) and X1Y (affected male). For the combination of male X and female X1 genes, the offspring will have the genotypes XX (normal female), X1X (carrier female), XY (normal male), and X1Y (affected male). When both the male and female genes are X, the offspring will be XX (normal female) and XY (normal male) with a probability of 1. If the male genes are Y and the female genes are X, the offspring will be XY (normal male) with a probability of 1. When the male genes are X1 and the female genes are XX1, the offspring will be XX1 (carrier female), X1X1 (affected female), XY (normal male), and X1Y (affected male). If the male genes are X1 and the female genes are X1X1, the offspring will be X1X1 (affected female) and X1Y (affected male).

The probability distributions derived from these combinations, when a normal male (XY) and a carrier female (XX1) reproduce, there is a 50% probability of having carrier offspring (XX1, X1Y). When a normal male (XY) and an affected female (X1X1) reproduce, there is a 25% probability for each of the genotypes X1X1, X1Y, and 0% for normal. When both parents are normal (XX and XY), all offspring will be normal with a probability of 1. For an affected male (X1Y) and a carrier female (XX1), there is a 50% probability for each of the genotypes XX1, X1Y. When an affected male (X1Y) and an affected female (X1X1) reproduce, there is a 25% probability for each of the genotypes X1X1 and X1Y. Lastly, when an affected male (X1Y) and a normal female (XX) reproduce, there is a 50% probability of having affected offspring (X1Y) and a 50% probability of carrier offspring (XX1). This comprehensive table combines the genotypic outcomes and their probabilities, offering a clear view of the genetic inheritance patterns across different parental gene combinations.

Table 3: Probability of offspring genotype

Time	X_1X_1	XX_1	XX	XX	XX_1	X_1X_1
1	0.225	0.125	0.275	0.275	0.200	0.125
2	0.1125	0.1875	0.30635	0.30635	0.85625	0.1125
3	0.2703125	0.4234875	0.353225	0.353225	0.4141125	0.2701125
4	0.238584375	0.280140625	0.459096875	0.459096875	0.386012500	0.238584375
5	0.2157953125	0.44534175045	0.52913203125	0.52913203125	1.0264003125	0.2157953125
…..	……..	…….	……..	……	……..	……
…..	……..	…….	……..	……	……..	……
25	0	0.5	0	0	0	0

From the above table 3, it is observed that equilibrium is reached at 25th stage and it can be found as per the given formula $P_{(i)}{}^{T} = P_{(i-1)}{}^{T} P$. The above table shows that the traits of G6PD will no more be in existence after the generations. This is due to the fact that each probability state vector goes to [0,0,0,0,0,0]. This study demonstrates that after a few generations, the population obtained a typical homozygote, which is not fixed for each family. It will vary from family to family, in some family the generation become normal after 12 stages or in some family it will be normal after six stages.

Similarly, for all the six state genotypes or initial state probability vectors equilibrium can be attained by the given formula. As a result, the six state Markov's model predicts that in a few generations, everyone in the family tree will become normal.

DISCUSSION AND FINDINGS

The study aims to adopt sustainable healthcare practices to manage G6PD deficiency in Oman by using pedigree charting and the Markov Chain Model to determine gene distribution patterns in family trees. The methodology synchronizes all medical data, including genetic history, through individual Civil IDs, applying Markov's model to predict the generational equilibrium state of G6PD deficiency. G6PD deficiency, a common genetic disorder in Oman, is genetically transmitted and marked by low G6PD enzyme levels, leading to various health complications. The prevalence of genetic disorders, including G6PD deficiency, is notably high in Oman, with significant implications for public health. This underscores the importance of preventive measures and intelligent data integration in healthcare. The study highlights the use of Markov Chain models, which are stochastic models describing sequences of potential events based on their preceding states. These models help predict future outcomes and are crucial in epidemiological studies. By examining different genetic combinations and their probabilities, the study provides insights into the inheritance patterns of G6PD deficiency. The detailed analysis using the Markov Chain model involves various steps, including storing health data digitally, processing this data, analyzing the results, developing a pedigree chart, and applying the Markov model to the family data. This process helps in understanding how G6PD deficiency propagates through generations and at what point the population reaches a normal homozygote state, indicating the absence of the deficiency. The study also emphasizes the importance of family medical history in predicting health risks. Information such as gender, birth year, ethnicity,

medical conditions, mental health history, pregnancy complications, lifestyle factors, and causes of death in deceased relatives should be documented. This data, integrated into intelligent data systems, can significantly enhance decision-making quality in healthcare. Pedigree charts play a crucial role in this process, as they visually represent family histories and help identify abnormal inheritance patterns. The study presents generational analyses of pedigrees in consanguineous marriages, showing the progression of G6PD deficiency through different generations. This helps in assessing the risk and planning preventive measures. The Markov Chain model's six-state transition analysis provides a detailed understanding of genotype probabilities, demonstrating how different combinations of male and female genes affect offspring genotypes.

The model predicts that after several generations, the population will eventually reach an equilibrium state where G6PD deficiency is no longer present. The findings of this study have significant implications for public health in Oman. By leveraging intelligent data systems and Markov Chain models, the healthcare system can improve early diagnosis, tailor interventions, and reduce long-term healthcare costs. This approach aligns with broader goals of enhancing healthcare sustainability and improving population health. This study demonstrates the effectiveness of using Markov Chain models and intelligent data systems in managing genetic disorders like G6PD deficiency. The integration of advanced technologies and preventive measures can lead to better health outcomes, economic sustainability, and a more resilient healthcare system in Oman. The Markov Chain model proves to be a valuable tool in predicting and managing genetic disorders, offering a feasible approach for future epidemiological studies.

CONCLUSIONS

As the Human Genome research advances, and more genetic testing is made available, family electronic health history is becoming increasingly important to know about the generations. In the near future, the role of genetic and genomic data in public health will become increasingly important. It has the potential to reveal information (such as disease onset at a young age, close relationships between patients, multiple affected family members, a suspicious genetic history, or diseases with a known genetic basis) that can be used to identify patients who may benefit from real-time consultation with medical genetics or disease-specific medical or surgical specialists. This work might serve as an introduction to the rich literature on Markov chains and its possible uses in epidemiological modeling. While there are many alternatives, more complex approaches for describing illness dynamics, the Markov approach's simplicity is appropriate for determining preliminary quantities of interest in the study of disease dynamics. Leveraging Markov Chain models within intelligent data systems presents a viable strategy for enhancing the sustainable management of G6PD deficiency in Oman. This approach not only supports the immediate needs of public health but also contributes to long-term health sustainability, aligning with broader goals of reducing the burden of genetic disorders and improving overall population health. As a result, the Markov chain model is proposed as a feasible approach for calculating other key epidemiological variables of infectious illnesses and generalizing transition probabilities for future forecasts.

REFERENCES

Al-Maawali, A. (2022). Oman Genome Project is the Future of Using Genomics as the Determinant of Health and Disease in the Society. In *Oman medical journal* (Vol. 37, Issue 6, p. e440). https://doi.org/10.5001/omj.2022.107

Al-Riyami, A., & Ebrahim, G. J. (2003). Genetic Blood Disorders Survey in the Sultanate of Oman. *Journal of Tropical Pediatrics*, *49 Suppl 1*, i1—20. http://europepmc.org/abstract/MED/12934793

Al-Sheryani, A., Al-Gheithi, H., Al Moosawi, M., Al-Zadjali, S., Wali, Y., & Al-Khabori, M. (2023). Molecular Characterization of Glucose-6-phosphate Dehydrogenase Deficiency in Oman. *Oman Medical Journal*, *38*(5), e552. https://doi.org/10.5001/omj.2023.107

Al Mawali, A. H. N., Al Qasmi, A. M., Al Sabahi, S. M. S., Idikula, J., Elaty, M. A. A., Morsi, M., & Al Hinai, A. T. (2017). Oman Vision 2050 for Health Research: A Strategic Plan for the Future Based on the Past and Present Experience. *Oman Medical Journal*, *32*(2), 86–96. https://doi.org/10.5001/omj.2017.18

Alangari, A. S., El-Metwally, A. A., Alanazi, A., Al Khateeb, B. F., Al Kadri, H. M., Alshdoukhi, I. F., Aldubikhi, A. I., Alruwaili, M., & Alshahrani, A. (2023). Epidemiology of Glucose-6-Phosphate Dehydrogenase Deficiency in Arab Countries: Insights from a Systematic Review. In *Journal of Clinical Medicine* (Vol. 12, Issue 20). https://doi.org/10.3390/jcm12206648

Beutler, E. (1996). G6PD: Population genetics and clinical manifestations. *Blood Reviews*, *10*(1), 45–52. https://doi.org/10.1016/S0268-960X(96)90019-3

Bianco, A. M., Marcuzzi, A., Zanin, V., Girardelli, M., Vuch, J., & Crovella, S. (2013). Database tools in genetic diseases research. *Genomics*, *101*(2), 75–85. https://doi.org/10.1016/J.YGENO.2012.11.001

Bruwer, Z., Al Ubaidani, S., Al Kharusi, K., Al Murshedi, F., Al-Maawali, A., Al Sayegh, A., Al Kindy, A., Al Riyami, N., Al Dughaishi, T., Al Salmani, M., Al Hashmi, N., Al Shehhi, M., Al Fahdi, B., Al Amri, S., & Al-Thihli, K. (2022). Uptake of prenatal genetic diagnosis and termination of pregnancy by Omani Muslim families at risk of genetic disorders: experience over a 9-year period. *Journal of Community Genetics*, *13*(3), 303–311. https://doi.org/10.1007/s12687-022-00584-1

Cornel, M. C., van der Meij, K. R. M., van El, C. G., Rigter, T., & Henneman, L. (2024). Genetic Screening—Emerging Issues. In *Genes* (Vol. 15, Issue 5). https://doi.org/10.3390/genes15050581

Grzymski, J., Elhanan, G., Morales-Rosado, J., Smith, E., Schlauch, K., Read, R., Rowan, C., Slotnick, N., Dabe, S., Metcalf, W., Lipp, B., Reed, H., Sharma, L., Levin, E., Kao, J., Rashkin, M., Bowes, J., Dunaway, K., Slonim, A., & Lu, J. (2020). Population genetic screening efficiently identifies carriers of autosomal dominant diseases. *Nature Medicine*, *26*, 1–5. https://doi.org/10.1038/s41591-020-0982-5

Hassan, K. S., Al-Riyami, A. Z., Al-Huneini, M., Al-Farsi, K., & Al-Khabori, M. (2014). Methemoglobinemia in an elderly patient with glucose-6-phosphate dehydrogenase deficiency: a case report. In *Oman medical journal* (Vol. 29, Issue 2, pp. 135–137). https://doi.org/10.5001/omj.2014.33

ISLAM, M. M. (2012). THE PRACTICE OF CONSANGUINEOUS MARRIAGE IN OMAN: PREVALENCE, TRENDS AND DETERMINANTS. *Journal of Biosocial Science*, *44*(5), 571–594.

https://doi.org/DOI: 10.1017/S0021932012000016

Khalifa, Y., Mandic, D., & Sejdić, E. (2021). A review of Hidden Markov models and Recurrent Neural Networks for event detection and localization in biomedical signals. *Information Fusion*, *69*, 52–72. https://doi.org/https://doi.org/10.1016/j.inffus.2020.11.008

King, K., Kaganovsky, J., & Shirts, B. (2022). eP493: Population genetic screening study participants intend to share test results with family members. *Genetics in Medicine*, *24*(3, Supplement), S313–S314. https://doi.org/https://doi.org/10.1016/j.gim.2022.01.525

Lacaze, P. A., Tiller, J., Winship, I., & Group, F. the D. N. A. S. I. (2022). Population DNA screening for medically actionable disease risk in adults. *Medical Journal of Australia*, *216*(6), 278–280. https://doi.org/https://doi.org/10.5694/mja2.51454

Lacy, R. C. (2012). Extending Pedigree Analysis for Uncertain Parentage and Diverse Breeding Systems. *Journal of Heredity*, *2012*(2), 197–205. https://doi.org/10.1093/jhered/esr135

Luzzatto, L., Ally, M., & Notaro, R. (2020). Glucose-6-phosphate dehydrogenase deficiency. *Blood*, *136*(11), 1225–1240. https://doi.org/https://doi.org/10.1182/blood.2019000944

Mesa, A., Basterrech, S., Guerberoff, G., & Alvarez-Valin, F. (2016). Hidden Markov models for gene sequence classification. *Pattern Analysis and Applications*, *19*(3), 793–805. https://doi.org/10.1007/s10044-015-0508-9

Munkhammar, J., & Widén, J. (2018). A Markov-chain probability distribution mixture approach to the clear-sky index. *Solar Energy*, *170*. https://doi.org/10.1016/j.solener.2018.05.055

Natarajan, J., & Joseph, M. A. (2021). Premarital screening for genetic blood disorders — an integrated review on the knowledge and attitudes of Middle Eastern university students. *Middle East Fertility Society Journal*, *26*(1), 19. https://doi.org/10.1186/s43043-021-00065-4

Ong, K. I. C., Kosugi, H., Thoeun, S., Araki, H., Thandar, M. M., Iwagami, M., Hongvanthong, B., Brey, P. T., Kano, S., & Jimba, M. (2017). Systematic review of the clinical manifestations of glucose-6-phosphate dehydrogenase deficiency in the Greater Mekong Subregion: implications for malaria elimination and beyond. *BMJ Global Health*, *2*(3), e000415. https://doi.org/10.1136/bmjgh-2017-000415

Paxinou, E., Kalles, D., Panagiotakopoulos, C. T., & Verykios, V. S. (2021). Analyzing Sequence Data with Markov Chain Models in Scientific Experiments. *SN Computer Science*, *2*(5), 385. https://doi.org/10.1007/s42979-021-00768-5

Pes, G. M., & Dore, M. P. (2022). Acquired Glucose-6-Phosphate Dehydrogenase Deficiency. *Journal of Clinical Medicine*, *11*(22). https://doi.org/10.3390/jcm11226689

Rajab, A., Al Rashdi, I., & Al Salmi, Q. (2013). Genetic services and testing in the Sultanate of Oman. Sultanate of Oman steps into modern genetics. *Journal of Community Genetics*, *4*(3), 391–397. https://doi.org/10.1007/s12687-013-0153-1

Sulaiman, A. J., Al-Riyami, A., Farid, S., & Ebrahim, G. (2001). Oman Family Health Survey 1995. *Journal of Tropical Pediatrics*, *47 Suppl 1*, 1–33.

Sun, J., Lindsay, B., & Rhodes, G. (2023). Markov chain composite likelihood and its application in genetic recombination model. *Journal of Statistical Computation and Simulation*, *94*, 1–27. https://doi.org/10.1080/00949655.2023.2288167

Sun, J., Palade, V., Wu, X., & Fang, W. (2014). Multiple Sequence Alignment with Hidden Markov Models Learned by Random Drift Particle Swarm Optimization. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *11*(1), 243–257.

https://doi.org/10.1109/TCBB.2013.148

Yoon, P. W., Scheuner, M. T., Peterson-Oehlke, K. L., Gwinn, M., Faucett, A., & Khoury, M. J. (2002). Can family history be used as a tool for public health and preventive medicine? *Genetics in Medicine*, *4*(4), 304–310. https://doi.org/10.1097/00125817-200207000-00009

***CORRESPONDING AUTHOR**

Divya Rani can be contacted at: *divyaraniryazan@gmail.com*

CITATION

Rani, D. (2025). Intelligent Data Systems for Sustainable Prevention and Management of G6PD Deficiency in Oman: A Markov Chain Model Approach. *Sohar University Journal of Sustainable Business, 1(1*), 1-15.