

Research Article

Markov Chains to Predict Malaria Incidence and Death in Gazira State, Sudan From 2001 to 2021

Badawi Osman Mohammed Fadlelkarim 1,2*

¹Department of MIS, Faculty of Business, Albaha University, Kingdom of Saudi Arabia ²Department of Statistics, Faculty of Economics and Political Sciences, Omdurman Islamic University, Sudan

Abstract

Background: Malaria is considered the most deadly and difficult parasitic disease in the world. This study aims to use Markov chains to predict the probability patterns of stability or change in malaria incidence and deaths.

Methods: Markov chains were used to analyze the data on malaria incidence and deaths through the Windows Quantitative Systems for Business (WINQSB) program. Data was obtained from the Ministry of Health, Gazira State, Health Information Centre, Sudan. The data is a time series, from 2001 to 2021 per year, according to three cases of decrease, stability, and increase. A transitional matrix is built for the three cases.

Results: The results revealed that the probability that malaria incidence and deaths will reach a stable state in one year and in the long run; the probability of transitioning to an increased state was 0.66 of malaria incidence; and the probability of moving to a decreased state was 0.52 of malaria deaths.

Conclusion: The results show that the malaria incidence will increase and malaria deaths will decrease in the short and long run from 2022 to 2030 in Gazira State. It is necessary to reinforce means and resources for case management and to investigate the determinants of the situation. Thus, strategies are urgently needed to arrest the unacceptably high incidence and death rates.

Keywords: Markov Chain, predicting, malaria incidence, malaria death, Gazira State

Corresponding Author: Badawi Osman Mohammed Fadlelkarim; email: bfadelkarim@bu.edu.sa

Received: 4 January 2024 Accepted: 18 March 2024 Published: 28 June 2024

Production and Hosting by KnE Publishing

© Badawi Osman Mohammed Fadlelkarim. This article is distributed under the terms of the Creative Commons
Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Editor-in-Chief: Prof. Nazik Elmalaika Obaid Seid Ahmed Husain, MD, M.Sc, MHPE, PhD.

○ OPEN ACCESS

1. Introduction

In 1907, A. Markov first introduced the fundamental concepts of Markov Chains. Soon after, prominent mathematicians like A. Kolmogorov and W. Feller further developed this theory. However, it wasn't until the 1960s that the true significance of Markov Chains for natural, social, and other applied sciences was fully recognized. This powerful mathematical system is known for its perplexity and burstiness, transitioning between a finite or countable number of possible states. It is a memoryless random process, where the next state depends solely on the current state and is not influenced by previous events. Markov chains have multiple practical applications as statistical models for real-world processes [1].

Many studies have been conducted in the form of time series to predict malaria incidence and deaths, but very few have used Markov chains for the prediction. Moreover, no study has analyzed the effectiveness of Markov chains in predicting malaria incidence and deaths.

Malaria has long been one of the world's major health challenges, particularly in tropical and subtropical countries. It is a leading cause of death in Africa. Malaria remains a serious obstacle to Africa's socioeconomic development and is estimated to be the cause of approximately 90% of deaths in Africa [2]. According to the WHO [3], an estimated 247 million cases of malaria and 619,000 malaria deaths were reported in 2021. Overall, in 2021, Africa was home to 95% of malaria cases and 96% of malaria deaths. Children under five years of age accounted for 80% of all malaria deaths in the region.

Therefore, this study aimed to predict the malaria incidence and deaths and examine three cases of the exchange rate (decrease, stability, and

increase) using Markov chains. The methodology used is analytical using Markov chains.

Malaria is considered the most deadly and difficult parasitic disease in the world [4]. Although malaria mortality is down worldwide, the incidence rate is increasing in several nations. Due to a variety of physical and social environmental conditions, groups within a country experience malaria in varied ways in terms of both time and space. A nation's malaria patterns are not country-wide; rather, they differ among communities for a variety of physical and socioenvironmental reasons [3].

1.1. Related works

- The study of Akter Akhi et al. [5], which deals with a continuous-time Markov chain and stochastic differential equations approach for modeling malaria propagation, discusses a transient numerical simulation and their outcomes strongly correlate with the actual scenario.
- 2. The study of Osman Mohammad [6] deals with Markov chains for forecasting the probabilities of the exchange rate of the Sudanese pound against the US Dollar for 1999–2015 and which concludes that the probability of moving to the stable state in the annual exchange rate matrix will arrive after one year if the current conditions continue. The probability of moving to decrease, stability, and increase states in the exchange rate are 0.29, 0.00, and 0.71 respectively for the Sudanese pound against the US dollar.
- 3. Kadium Hamza et al. [7] conducted a Markov chain-based exchange rate forecasting of the Iraqi Dinar against the US Dollar. The outcome showed that the likelihood of an increase is

0.0698, that of stability is 0.9302, and that of a decline is zero.

4. The study of Hussien et al. [8], which deals with the Statistical Methods for Predicting Malaria Incidences using data from Sudan, revealed that malaria is the leading cause of illness and death in Sudan. Malaria epidemics pose a threat to the entire society, placing a tremendous cost on both the populace and the government. The objective of this study was to create useful and comprehensible time series models for forecasting levels of incidence in the future. The findings indicate that in Gadaref, Gazira, North Kordofan, and Nord, the transform approach outperforms other methods by a substantial margin, whereas in Khartoum, the moving average model outperforms other models by a significant margin. Of all the previous studies, this study is the only one, to the best of the authors' knowledge, in which Markov chains were used to predict infection and deaths from malaria.

2. Materials and Methods

This study uses Markov chains to predict the probabilities of malaria incidence and deaths because it is suitable for the nature of the data as it is discrete and a new method for prediction.

If the future state of a stochastic process is dependent on the current state unaffected by previous processes, it is referred to be a Markov process. The kind of state space and the kind of time parameters are used to categorize Markov processes. A Markov process can have a discrete state or a continuous state in state space. Consequently, there are four fundamental kinds of Markov processes: [6] Markov chain in

discrete time, Markov chain in continuous time, Markov process discrete, and Markov process in continuous time [9].

2.1. The structure of Markov processes

A stochastic process that alternates between discrete states at predetermined or arbitrary times is called a jump process. The system enters a state during such a procedure and remains in it for a predetermined period, known as the holding time. The process' sample path remains constant between T_k and T_k+1 if the number of hops is $T_0 =$ $0 < T_1 < T_2 < ...$ A jump chain is created if the jump times are discrete. The stochastic processes can be Markov processes if the status in the future does not depend on past situations, but only on the present status [10]. If the following three conditions are met, the random process $\{X_n: n : \in T\}$ is called a Markov chain requirement [11]: The case space for this process is separated, the parameter space for this process is separated, if the process depends only on current circumstances and is independent of past circumstances. The transition probability matrix (TPM) is as follows:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & p_{n2} & p_{nn} \end{pmatrix} \to (1).$$

- 1. The TPM conditions are [12]: Square matrix, nonnegative elements as $1 \geq p_{ij} \geq 0$, total row equal one $\sum_{j \in T} p_{ij} = 1$
- 1. One-step (TP) [11]: Transition probability from case to case by one step as:

$$p_{ij}^{(n, n+1)} = P(X_{n+1} = j X_n = i),$$

$$i, j = 0, 1, 2, ..., \rightarrow (2).$$

1. Stability transition probabilities: Transition probability from situation ito situation j in step is stable if it does not depend on n, and it is represented by p_{ij} in this case:

$$p_{ij} = P(X_{n+1} = j | X_n = i), \quad \forall n \to (3).$$

- 2. Homogeneous Markov chain: Markov chain is homogeneous if it has stable transition probabilities.
- 1. The *n*-step transition probability [11]: It is represented by $p_{ii}^{(n)}$ as follows:

$$p_{ij}^{(n)} = P\left(X_{n+m} = j X_m = i\right),\,$$

$$n \ge 0$$
, $i, j = 0, 1, 2, ..., \rightarrow (4)$.

It can be written in a matrix as:

$$P^{(n)} = \begin{pmatrix} p_{00}^{(n)} & p_{01}^{(n)} & p_{02}^{(n)} & \cdots \\ p_{10}^{(n)} & p_{11}^{(n)} & p_{12}^{(n)} & \cdots \\ p_{20}^{(n)} & p_{21}^{(n)} & p_{22}^{(n)} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \rightarrow (5).$$

1. Chapman–Kolmogorov equation [11]: Determines transition probabilities from one state to another after *n* events.

$$p_{ij}^{(n)} = p_{ik}^{(m)} p_{kj}^{(n-m)} 1 \le m \le n-1 \to (6).$$

2.2. Stationary and steady state

The term "stationary" means that the statistical properties of a random process do not change over time. A steady state occurs when the stochastic

process continues until the transition number ratio stabilizes at some value in each case, in which case we call it steady-state probability [13]. When P is the TPM of a stochastic process with m finite cases, then:

$$\lim_{m \to \infty} P^m = V = \begin{bmatrix} v \\ v \\ \vdots \\ \vdots \\ v \end{bmatrix} \to (7).$$

Since the only probability vector is: $v^0 = (v_1, v_2, ..., v_n)$, where: $\sum_{i=1}^m v_i = 1$ and $0 \le v_i \le 1$.

The Stationary distribution for the coming term is computed by: $v = v.P \rightarrow (8)$.

When m approximates to ∞ , the transition probabilities $p_{ij}^{(m)}$ with m steps will depend on the final step and not depend on the primary step, this means that after the large number of trials, the chain will be stationary [14].

3. Results

This study is based on yearly and annual data of malaria incidence, obtained from the Ministry of Health, Gazira State - Health Information Centre, Sudan. The data is a time series from 2001 to 2021. The data were analyzed using the Markov chains theory to make a forecast and other related analyses through the WINQSB program. The study examines three cases of the exchange rate (decrease, stability, and increase), and the data represented in Appendix (1).

3.1. The analysis of malaria incidence by years from 2001 to 2021

Table 1 presents malaria incidence as a series of Markov chains by sample space.

Based on the data in Table 2, we can estimate the transition probabilities to move from case i to case j using the following relationship:

$$p_{ij} = \frac{N_{ij}}{\sum_{k \in S} N_{ik}} \ \forall i, j \in S.$$

Thus, the estimated TPM becomes:

$$P = \begin{bmatrix} \frac{2}{7} & \frac{5}{7} \\ \frac{5}{14} & \frac{9}{14} \end{bmatrix} = \begin{bmatrix} 0.29 & 0.71 \\ 0.3 & 0.64 \end{bmatrix}.$$

The initial vector v^0 can get it out of rows sum in Table 2 as follows:

$$v^0 = \left(\frac{7}{21} \frac{14}{21}\right) = (0.330.67).$$

The prediction of any situations under study for the next period (next year) comes by multiplying the initial vector v^0 by TPM (P).

The case of stability and stationery: The case can reach the state of stability and stationery by using the formula (8) through the WINQSB program for accurate results and reduce the time as follows:

$$v^n = v^{n-1}P$$
.

The predictable situations under study for the next period (one year) are as follows:

$$v^1 = v^0 P$$
.

where

$$v^0 = (0.330.67)$$

and

$$P = \left[\begin{array}{cc} 0.29 & 0.71 \\ 0.3 & 0.64 \end{array} \right]$$

. Then,
$$v^1 = v^0 P = \begin{pmatrix} 0.34 & 0.66 \end{pmatrix}$$
.

The predictable situations under this study for the next period (two years) are as follows:

$$v^2 = v^1 P = \left(\begin{array}{c} 0.34 & 0.66 \end{array} \right).$$

The predictable situations under this study for the next period (30 years) are as follows:

$$v^{30} = v^{29}P = \left(0.34 \ 0.66 \right).$$

It is clear from the above vector that a case of stability and stationery can be accessed after one year and the possibilities may continue indefinitely (if current conditions persist).

3.2. The Analysis of malaria deaths by years from 2001 to 2021

Malaria death represents a series of Markov chains by sample space, as in the following table.

We can estimate the transition probabilities to move from case *i* to case *j* using the data given in Table 4 by using the following relationship:

$$p_{ij} = \frac{N_{ij}}{\sum_{k \in S} N_{ik}} \quad \forall i, j \in S.$$

Thus, the estimated TPM becomes:

$$P = \begin{pmatrix} \frac{4}{11} & \frac{1}{11} & \frac{6}{11} \\ \frac{0}{0} & \frac{0}{0} & \frac{1}{1} \\ \frac{7}{9} & \frac{0}{14} & \frac{2}{9} \end{pmatrix} = \begin{pmatrix} 0.36 & 0.09 & 0.55 \\ 0.0 & 0.0 & 1.00 \\ 0.78 & 0.0 & 0.22 \end{pmatrix}.$$

The initial vector v^0 can get it out of rows sum in Table 2 as follows:

$$v^0 = \left(\begin{array}{cc} \frac{11}{21} & \frac{1}{21} & \frac{9}{21} \end{array}\right) = \left(\begin{array}{cc} 0.52 & 0.05 & 0.43 \end{array}\right),$$

where the prediction of any situations under study for the next period (next year) comes by multiplying the initial vector v^0 by TPM (P).

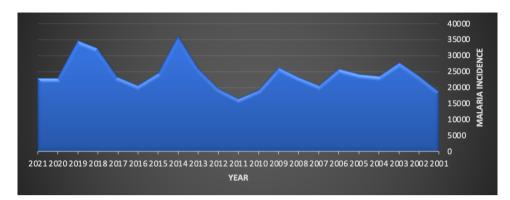


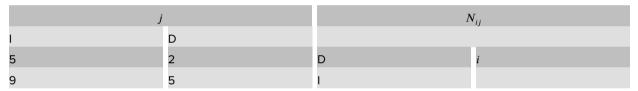
Figure 1: Malaria incidence by years.

TABLE 1: Malaria incidence as a series of Markov chains by sample space.

2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
I	I	I	D	l	l	D	I	l	D	D
2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	
I	I	I	D	D	I	I	I	D	I	

D: decrease; S: stability; I: increase.

TABLE 2: The construction of the Transitional Probability Matrix by finding the sum of cases in Table 1.



 N_{ij} : number of times to move from casei i o case j.

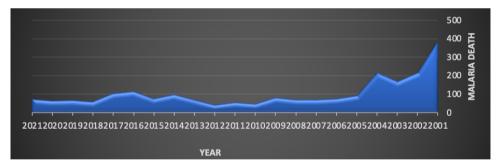


Figure 2: Malaria deaths by years.

TABLE 3: Malaria deaths as a series of Markov chains by sample space by years.

2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
I	D	D	I	D	D	D	S	I	D	I
2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	
D	I		D	I	D	D	1	D	l	

D: decrease; S: stability; I: increase.

The case of stability and stationery: The case can reach the state of stability and stationery by

using the formula (8) through the WINQSB program for accurate results and reduce the time as follows:

$$v^n = v^{n-1}P.$$

TABLE 4: The construction of the Transitional Probability Matrix by finding the sum of cases in Table 1.

 N_{ii} represent several times to move from case *i* to case *j*.

The predictable situations under study for the next period (one year) are as follows:

$$v^1 = v^0 P$$
.

where

$$v^0 = \left(\begin{array}{cc} 0.52 & 0.05 & 0.43 \end{array}\right)$$

and

$$P = \begin{pmatrix} 0.36 & 0.09 & 0.55 \\ 0.0 & 0.0 & 1.00 \\ 0.78 & 0.0 & 0.22 \end{pmatrix}$$

Then,
$$v^1 = v^0 P = \begin{pmatrix} 0.52 & 0.05 & 0.43 \end{pmatrix}$$
.

The predictable situations under this study for the next period (two years) are as follows:

$$v^2 = v^1 P = \left(\begin{array}{cc} 0.52 & 0.05 & 0.43 \end{array} \right).$$

The predictable situations under this study for the next period (30 years) are as follows:

$$v^{30} = v^{29}P = \left(\begin{array}{cc} 0.52 & 0.05 & 0.43 \end{array}\right).$$

It is clear that from the above vector that the case of stability and stationery can be accessed after one year and the possibilities may continue indefinitely (if the current conditions persist).

4. Discussion

The significance of this study lies in its use of Markov chains to predict the likelihood of

stability or change in malaria infection patterns and death rates. Markov chains are known for their high accuracy in prediction compared to other statistical forecasting tools such as time series. Between the years 2001 and 2021, Figure 1 shows that the highest incidence occurred in 2014 and the lowest in 2011. Table 1 shows that malaria incidence increased by 67%. Table 2 shows the estimated TPM results as follows: 0.29 probability that malaria incidence changes from decrease to decrease; 0.71 probability that malaria incidence changes from decrease to increase; 0.30 probability that malaria incidence changes from increase to decrease; 0.64 probability that malaria incidence changes from increase to increase. Also, the results of an initial vector (v^0) can get it out of rows sum in Table 2 as follows: 0.33 probability that all malaria incidence will move to a decrease state and 0.67 probability that all malaria incidence will move to an increase state. The predictable situations under this study, that is, malaria incidences from 2022 to 2030, reach the state of stability and stationery by using formula (8) through the WINQSB program, as follows: 0.66 probability of transitioning to a state of increased malaria incidence in the period 2022–2030. Also, Figure 1 shows the highest malaria deaths in 2001 and the lowest in 2012. Table 3 shows that malaria deaths decreased by 52%. Table 4 shows the estimated TPM results by finding the sum of cases in Table 3 as follows: 0.36 probability that malaria deaths change from decrease to decrease; 0.09 probability that malaria deaths change from decrease to stability; 0.55 probability that malaria deaths change from decrease to increase; 0.00 probability that malaria deaths change from stability to decrease; 0.00 probability that malaria deaths change from stability to stability; 1.00 probability that malaria deaths change from stability to increase; 0.78 probability that malaria deaths change from increase to decrease; 0.00 probability that malaria deaths change from increase to stability; and 0.22 probability that malaria deaths change from increase to increase. Also, the results of an initial vector (v^0) can get it out of rows sum in Table 4 as follows: 0.52 probability that all malaria deaths will move to a decrease state. The predictable situations under this study of malaria death from 2022 to 2030 reach the state of stability and stationery by using formula (8) through the WINQSB program, as follows: 0.52 probability of transitioning to a state of decreased malaria deaths from 2022 to 2030. The study agrees with Hussien et al. [8] that malaria is one of the leading causes of disease and deaths in Sudan. In addition, the study agrees with Osman Mohammad [6] in predicting the probability of transition to an increased state, but disagrees with the study of Kadium Hamza and others [7] in predicting the probability of transition to a stable state. The results of this study indicate an urgent need to strengthen malaria control programs in Sudan, as it is a major cause of incidence and death. Health policymakers can use these findings to develop more effective strategies to reduce the spread of malaria and improve healthcare provided to those infected. It also highlights the importance of developing early warning systems to predict the escalation of malaria cases and act quickly to reduce its spread.

5. Conclusion

The results show that the malaria incidence has increased and malaria deaths have decreased in the short and long run from 2022 to 2030 in Gazira State. First, the probability that malaria incidence will reach a steady state after 2021 or in the period 2022-2030 is as follows: 0.34 probability of transitioning to a state of decreased malaria incidence by 34%, and 0.66 probability of transitioning to a state of increased malaria incidence by 66%. Second, the probability that malaria deaths will transition to a steady state after 2021 or in the period 2022-2030 is as follows: 0.52 probability of transitioning to a state of decreased malaria deaths by 52%; 0.05 probability of transitioning to a state of stability in malaria deaths by 5%; and 0.43 probability of transitioning to a state of increased malaria deaths by 43%. This study differed from previous studies in that it dealt with predicting the probabilities of infection and deaths from malaria using Markov chains, while previous studies used other methods, and some of them dealt with predicting Markov chains in economic topics. The results indicate that Markov chain prediction is useful for ongoing intervention strategies by governmental and non-governmental agencies in Gazira State to effectively control the disease. It is necessary to reinforce means and resources for case management and to investigate the determinants of that situation. So, strategies are urgently needed to arrest the unacceptably high incidence and death rates.

Limitations

This study did not take into account all social, economic, and environmental factors that may affect the prevalence of malaria. For future research, more comprehensive studies should be

conducted that include a wider range of data and different factors that influence malaria incidence and mortality. Predictive models using Markov chains that take into account social, economic, and environmental variables should also be developed. The study relies on ready data on malaria infection and deaths for the period 2001–2021, and the probabilities of increase or decrease were predicted without specifying numbers.

Declarations

Acknowledgements

The author is thankful to the Ministry of Health, Gazira State - Health Information Centre, Sudan for providing the study data.

Ethical Considerations

A letter was written to the Ministry of Health, Gezira State - Health Information Centre, Sudan to provide study data.

Competing Interests

The author has no competing interests.

Availability of Data and Material

The manuscript includes all the required data.

Funding

The author did not receive any funding from any organization.

Abbreviations and Symbols

WIN: Windows

QSB: Quantitative Systems for Business

WHO: World Health Organization TPM: Transition Probability Matrix

TP: Transition probability

References

- [1] Voskoglou, M. (2016). Applications of Finite Markov Chain Models to Management. *American Journal of Computational and Applied Mathematics*, 6(1), 7–13. https://doi: 10.5923/j.ajcam.20160601.02
- [2] Nkiruka, O., Prasad, R., & Clement, O. (2021). Prediction of malaria incidence using climate variability and machine learning. *Informatics in Medicine Unlocked*, 22, 100508.
- [3] Akhi, A. A., Kamrujjaman, M., Nipa, K. F., & Khan, T. (2024). A continuous-time Markov chain and stochastic differential equations approach for modeling malaria propagation. *Healthcare Analytics*, 4, 100239, ISSN 2772-4425, https://doi.org/10.1016/j.health.2023.100239
- [4] Lubinda, J., Bi, Y., Hamainza, B., Haque, U., & Moore, A. J. (2021). Modelling of malaria risk, rates, and trends: A spatiotemporal approach for identifying and targeting sub-national areas of high and low burden. *PLoS Computational Biology*. https://doi.org/ 10.1371/journal.pcbi.1008669
- [5] Mohammad Fadlelkarim, B. O. (2022). Markov chains for forecasting of probabilities of exchange rate of Sudanese pound against the US Dollar for the period 1999 to 2015. *Journal of the Faculty of Economics*, 21.
- [6] Hamza, S. K., Ahmed, A. D., & Hussein, S. A. (2020). Forecasting the exchange rate of the Iraqi dinar against the US dollar using Markov chains. *Periodicals* of Engineering and Natural Sciences, 8(2), 626–631. http://pen.ius.edu.ba/index.php/pen/article/view/1280
- [7] Hussien, H. H., Eissa, F. H., & Awadalla, K. E. (2017). Statistical methods for predicting malaria incidences using data from Sudan. *Malaria Research* and *Treatment*, 2017, 4205957. https://doi.org/10. 1155/2017/4205957
- [8] Oliver, C. (2009). Markov processes for stochastic modeling (1st ed.). Elsevier Academic Press.

- http://www.sciencedirect.com/science/book/ 9780124077959
- [9] Rosenblatt, M. (1971). Markov processes. Structure and asymptotic behavior. Springer Berlin. https://doi.org/ 10.1007/978-3-642-65238-7
- [10] Taylor, H. (1998). An Introduction to stochastic modeling (3rd ed.). University Stanford, California. https://appliedmath.arizona.edu/sites/default/files/ 0f04d86a836182cbf608dfc86c7a70f5e5f6_0.pdf
- [11] Mourad, M., & Harb, A. (2013). A predictive model for the daily exchange rate of the EUR/USD using Markov chain and co-integration techniques. Lebanese Science Journal, 14(2), 950–970. https:

- //applications.emro.who.int/imemrf/Lebanese_Sci_J/Lebanese_Sci_J_2013_14_2_93_113.pdf
- [12] Madhava Rao, K. S., & Ramachandran, A. (2016). A Markov approach to exchange rate sentiment analysis of major global currencies. *Open Jour*nal of Statistics, 6. https://www.scirp.org/journal/ paperinformation?paperid=73204
- [13] Serfozo, R. (2009). Basics of applied stochastic processes, probability and its applications. Springer-Verlag Berlin. https://books.google.com.sa/books?id= JBBRiuxTN0QC&printsec=frontcover&hl=ar&source= gbs_ge_summary_r&cad=0#v=onepage&q&f= falsehttps://doi.org/10.1007/978-3-540-89332-5

Appendix 1

TABLE 5: Malaria incidence by years for 2001–2021.

Year	Incidence	Death	Year	Incidence	Death	Year	Incidence	Death
2001	18695	387	2008	23369	69	2015	24472	75
2002	23582	215	2009	26311	81	2016	20666	116
2003	27815	168	2010	19310	49	2017	23349	105
2004	23582	215	2011	16584	55	2018	32295	59
2005	24236	92	2012	19856	42	2019	34667	68
2006	25899	75	2013	26222	71	2020	22995	64
2007	20691	69	2014	36224	97	2021	28326	74

The table was prepared using the data obtained from the Ministry of Health, Gazira State - Health Information Centre, Sudan.