Post-Discharge Malaria Chemoprevention: A Statistical Analysis of Effectiveness and Challenges

Submitted in partial fulfilment of the requirements for the degree of

Master of Business Administration – (BA)

By

Goda Shruthi 24WU0202321

U.Satya Sri Priya 24WU0202378

Madala Navya Dhanasri 24WU0202338

Manjunath Devarashetty 24WU0202440

Submitted to

Dr. Hemachandran K

Area Chair and Professor - Analytics Department

School of Business, Woxsen University

Year (Ex: 2024 – 2026)

Declaration

We hereby declare that the thesis entitled Post-Discharge Malaria Chemoprevention: A Statistical Analysis of Effectiveness and Challenges submitted to Woxsen University for the award of the degree of Master of Business Administration (BA.AI.ML) is a record of bonafide work carried out by the team under the supervision of Dr Hemachandran K, Area Chair and Professor - Analytics Department, School of Business Woxsen University, Hyderabad. We further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place: Hyderabad

Date: Signature of the Candidate

Certificate

This is to certify that the thesis entitled "Post-Discharge Malaria Chemoprevention: A Statistical Analysis of Effectiveness and Challenges" submitted by Ms. Goda Shruthi, Ms. Satya Sri Priya, Ms Navya Dhanasri Madala ,Mr Manjunath Devarashetty School of Business, Woxsen University, Hyderabad for the award of the degree of Master of Business Administration (BA), is a record of Bonafide work carried out by the team under my supervision, as per the Woxsen University code of academic and research ethics. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The thesis fulfils the requirements and regulations of the University and in my opinion meets the necessary standards for submission.

Place: Hyderabad

Date: Signature of the Guide

Abstract

Post-discharge malaria chemoprevention (PDMC) has become a critical intervention in the prevention of morbidity and mortality among children recovering from severe malaria and anaemia. Statistical estimates from various studies show that PDMC can decrease hospital readmissions by 19% and avert up to 38,600 malaria-related readmissions per year in high-risk areas like upland Tanzania. In addition, estimates indicate that the treatment of 2–5 children with PDMC averts a hospital readmission, and fewer than 100 treatments avert a death. In Burkina Faso, Seasonal Malaria Chemoprevention (SMC), a similar intervention, decreased cases of malaria by 51%, anaemia by 32%, and fever by 46%.While these results are promising, limitations include drug resistance, logistical impediments to roll-out, and heterogeneity of community acceptance. Stakeholder dialogue and WHO recommendations emphasize the need for adaptive strategies, including roll-out of PDMC to children with severe malaria, roll-out in the community, and monitoring of resistance. These initiatives underlie Sustainable Development Goal (SDG) 3—Good Health and Well-being—by closing post-discharge risks, leveling playing fields of access to lifesaving treatment, and fortifying strong health systems to battle death from childhood malaria. These gaps can be surmounted by augmenting policy strengthening, intensified follow-up action, and targeted deployment of drugs for the purpose of capitalizing on PDMC as a cost-effective, replicable, and sustainable malaria-control intervention for malaria-endemic countries.

Keywords:

Post-Discharge Malaria Chemoprevention (PDMC), Malaria, Anaemia, Hospital Readmissions, Mortality, Drug Resistance, Sustainable Development Goal 3 (SDG 3).

Acknowledgement

Firstly, we would like to express our sincere gratitude to our project mentor **Dr Hemachandran K**, Area Chair and Professor - Analytics Department School of Business, Woxsen University for giving the opportunity to do research and providing invaluable guidance throughout the project. He is continuous inspiration who pushed us to think creatively and efficiently. His vast knowledge, extensive experience, and professional competence in various fields of AI has enabled us to successfully accomplish this project. It was a great privilege and honour to study and work under his mentorship for the project. We are extremely grateful for what he has offered to the project. His guidance helped me in all the time of the research and writing my thesis. This endeavour would not have been possible without his help and supervision.

We thank **Vineet Singh Research Associate,** AI Research Centre for providing the necessary support and valuable suggestions during the project work

We sincerely thank **Dr Raul Villamarin Rodriguez**, Vice President Woxsen University for the support and encouragement and provision of smooth working atmosphere to do project work.

Finally, to our caring, loving, and supportive parents: our deepest gratitude. Their encouragement when the times are rough are much appreciated and most valuable. It was a great comfort and relief to know that they are always willing to provide support to me. Our heartfelt thanks.

Place: Hyderabad

Date:

Table of Contents

Contents

Chapter 1

Introduction

Malaria remains a leading cause of morbidity and mortality in children, particularly in sub-Saharan Africa. Children remain vulnerable to subsequent re-infections even after receiving treatment in the hospitals. PDMC has a potential to mitigate diseases compounded by malaria infections through the prevention of malaria-associated relapses and debilitating outcomes. Research has shown that PDMC can prevent hospitalisation and enhance survival rates. However, its implementation has been sporadic, thereby limiting its full impact.

Several exciting publications have documented the dampening effects of PDMC on the rates of malaria hospitalization; hence, it could avert some 38600 readmissions annually in Tanzania. In Burkina Faso, chemoprevention would reduce cases of malaria in young children by 51 percent. Dihydroartemisinic-piperaquine proved to be both safe and efficacious. However, these benefits are marred by the added challenges of drug resistance and noncompliance to treatment.

Support for PDMC is offered by the WHO as one of the key interventions in high-target disease spreading areas. The adoption of PDMC is slowed by weak healthcare systems and untrained health personnel. There is a need to develop contextual strategies to improve the implementation and access to treatment. Linking PDMC with Sustainable Development Goal 3 can contribute to the improvement of child health outcomes. This study seeks actual data about the efficacy, the challenges of PDMC, and ways of improving PDMC's success.

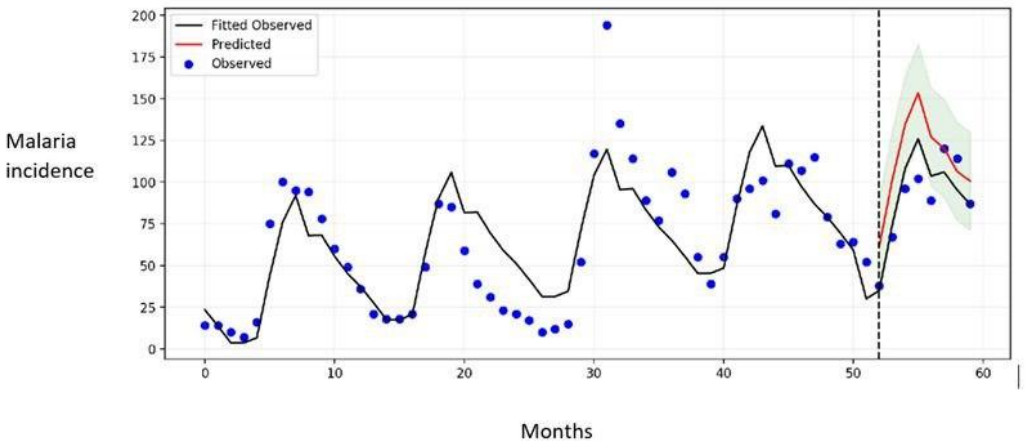

Figure 1: Forecasting Malaria Trends Over Time

The above graph (Figure 1) shows malaria incidence over a period of 60 months. The blue dots are the observed malaria cases, and the black line is the trend fit from historical data. The red line is the forecasted malaria cases after the dashed vertical line, which is the point of prediction. The green coloured area is the confidence interval for the predictions. There are clear-cut seasonal peaks of malaria cases, with outbreaks happening every 10-15 months. The prediction suggests an increase in the cases of malaria following 50 months, with a subsequent decline.

BACKGROUND AND MOTIVATION

Malaria continues to be a significant public health problem, especially in young children who easily become reinfected post-discharge. Control efforts have shown success, but frontline management and preventive measures continue to be hampered by socioeconomic barriers. In this case, chemoprevention, the periodic administration of antimalarial drugs, seems to offer an attractive method for malaria prevention, albeit also affected by multidrug resistance, access to healthcare, patient compliance for treatment, and cost. The multivariable assessment will evaluate these factors that would contribute to enhancing malaria prevention in school children.

BACK GROUND

Despite a falling incidence of malaria, rural children are still at risk due to poor healthcare availability, poor follow-ups, and poverty. Chemoprevention seems to have some solutions, but challenges such as drug resistance, healthcare inequities, and adherence to treatment must also addressed.

The study incorporates various factors that will help in better malaria prevention in endemic areas, including the accessibility of rural health care, drug resistance issues, clinical trial mechanisms, and socioeconomic issues for adherence to treatment. Insights from the management of pandemics, as well as some aspects of pharmacokinetics, will also help combat malaria. Comparisons between India's challenges and South Africa's health disparities would identify effective measures and inform policies. Through an analysis of these factors, this research should therefore aid in strengthening malaria prevention efforts and finding an equitable solution to healthcare access.

OBJECTIVE

The main aim of this research is to assess the efficacy of malaria chemoprevention interventions in post-discharge care of school children in India. The research will examine the effect of preventive treatment on the prevention of malaria recurrence and complications. Through the use of a multivariable analysis, the study aims to determine the major factors affecting adherence, effectiveness, and long-term health outcomes among children recovering from severe malaria.

Chapter 2

Literature review

Malaria remains a prevalent global health issue, particularly for children recuperating from severe infections. The aim of post-discharge malaria chemoprevention (PDMC) is to reduce the risk that these vulnerable individuals would become reinfected. Severe malaria hospitalized children are at high risk of getting reinfected due to compromised immunity and continuous exposure in malaria-hyperendemic communities, as indicated by studies. Antimalarial drugs are administered at regular intervals following hospitalization as a component of PDMC to manage this, to prevent relapse and reduce the overall malaria burden.

Studies have also indicated how effective PDMC is in enhancing the health outcomes and lessening the chances of recurrence of malaria.

Clinically, studies revealed that children who received PDMC experienced lower hospital readmission rates as well as malaria episodes compared to their counterparts who did not receive PDMC. Sulfadoxine-pyrimethamine (SP) and dihydroartemisinin-piperaquine (DP), two of the most widely prescribed drugs, have been found to be effective in reducing reinfection. Nonetheless, several factors, such as resistance to medication, compliance of patients, and geographical occurrences of malaria transmission, influence the success of PDMC. Ongoing research and response to changing epidemiological circumstances are required to ensure that PDMC remains successful. Despite its benefits, PDMC is subjected to several issues of implementation.

Families in low-resource environments, where access to medical institutions is restricted, may find it challenging to follow treatment plans. The efficiency of existing treatment regimens is further diminished by the rising threat of drug-resistant forms of malaria. The scalability of PDMC initiatives is further hampered by logistical and financial limitations. In malaria-endemic regions, the removal of these barriers is critical to the maximization of the impact of PDMC. Post-discharge malaria chemoprevention is in keeping with Sustainable Development Goal (SDG) 3: Good Health and Well-being, specifically to put an end to malaria and other infectious disease epidemics. Scaling up PDMC programs can be a major help in lowering the morbidity and mortality of malaria, especially in children.

In order to accomplish this, there needs to be coordination between governments, healthcare facilities, and international organizations to reinforce drug distribution, raise community knowledge, and commit to alternative drug treatments.

Stronger PDMC programs not only will make individual health impacts better but will also support more global efforts at eradicating malaria.

Table:1

Title	Authors	Summary	Drawbacks
Assessing the prevalence, risk factors, and socio-demographic predictors of malaria among pregnant women in the Bono East Region of Ghana: a multicentre hospital-based mixed-method cross-sectional study	Sophie Uyoga, Peter Olupot,	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as "test-and test "policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	sociodemographic and maternal predictors highlight the need to strengthen screening for malaria, administer treatments, monitor maternal and foetal health, and provide education and counselling.
A multi-center, open-label trial to compare the efficacy and pharmacokinetics of Artemether-Lumefantrine in children with severe acute malnutrition versus children without severe acute malnutrition: study protocol for the MAL-NUT study	Prof Kamija S Phiri MD PhD	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as "test-and test "policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	lacking information on the efficacy of AL in SAM children and could inform further recommendations for the management of malaria in this specific population.

Mass Azithromycin Distribution to Prevent Child Mortality in Burkina Faso	Carole Khairallah MSc	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Death rates were slightly lower in the azithromycin group th an in the placebo group: 8. 2 deaths per 1,000 person-years versus 10.0 deaths per 1,000 person-years. The effect did not r each statistical significanc e and could have been by chance. The effect was most pronounced in children aged 24 to 59 months and less clear for the younger children.
Analysis of Postdischarge Interventions for Children Treated for Moderate or Severe Wasting, Growth Faltering or Failure, or Edema	Lise Denoeud-Ndam, Alassane Dicko,	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Since these children were clinically cured and out of treatment, there was always a danger that they would become malnourished again or fall victims of infections or worse, life-threatening conditions. Of course, current global guidelines do not specifically focus on the management of these children after discharge.

Identifying neonates at risk for post-discharge mortality : Derivation and internal validation of a novel risk assessment tool	Catherine E. Oldenburg, ScD1,2,3,4; Mamadou Ouattara, MD5; Mamadou Bountogo, MD5; et al	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	A small number of factors predicted all-cause, 60-day mortality after discharge from neonatal wards . After external validation, this risk assessment tool may facilitate clinical decision making for eligibility for discharge and the direction of resources to follow-up high risk neonates.
Evaluation of pharmacokinetics of single-dose primaquine in undernourished versus normally nourished children diagnosed with Plasmodium vivax malaria in Mumbai	Lilia Bliznashka, PhD1; Susan M. Rattigan, BA2; Christopher R. Sudfeld, Sc D2,3; et al	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Results showed that, while the undernourished children weighed less, there was no significant difference in how their bodies processed PQ compared to the well-nourished children. This suggests that PQ dosing might not need to be adjusted based on nourishment status in children with malaria.

Chemoprevention of malaria with long-acting oral and injectable drugs: an updated target product profle	Chris A Rees1,2, Reading C Ideh3, Rodrick Kisenge4, Julia Kamara3,Ye-Jeung G Coleman-Nekar3, Abraham Samma4,	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	First, although numerous existing interventions, such as chemopreventive drugs, are saving many lives, they increasingly be come less effective because of developing drug resistance. Second, while new long-acting drugs and vaccines have been developed, it is still not widely accessible.
Implementation of post-discharge malaria chemoprevention (PDMC) in Benin, Kenya, Malawi, and Uganda: stakeholder engagement meeting report	Myriam El Gaaloul1*, Andre Marie Tchouatieu1*, Kassoum Kayentao2 , Brice Campo1 , Benedicte Bufet1 , Hanu Ramachandrun i1 , Jean Louis Ndiaye3 , Timothy N. C. Wells1 , Celine Audibert1 , Jane Achan4 ,	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Among the major limitations noted are the need for more evidence generation and internal policy harmonization, which may hold up the implementation of post discharge malaria chemoprevention (PDMC). Engaging stakeholders and policymaking are time and resource-intensive processes in and of themselves. Furthermore, countries do not have an existing delivery platform for PDMC, therefore, making it able to efficiently roll out PDMC within the numerous health care settings in a timely manner becomes an issue.

Combining malaria vaccination with chemopreven tion: a promising new approach to malaria control	Jenny Hill	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	A limitation of these malaria control strategies is that they are not very effective when vaccines or gene drive mosquitoes are used alone in hot zones, requiring and therefore advocating for a targeted strategy. The use of vaccination alongside chemoprevention raises issues on the chances of emergence and spread of drug and vaccine-resistant parasites, which has to be carefully monitored .Moreover, expensive and complicated procedures and techniques like monoclonal antibodies may not be practical in underdeveloped regions because of their high expense and difficult logistical requirements.
Quality of care and post-discharge morbidity among children diagnosed with severe malaria in rural Uganda: A prospective cohort study	Brian Greenwood, Matthew Cairns, Mike Chaponda, R. Matthew Chico, Alassane Dicko, Jean-Bosco Ouedraogo, Kamija S. Phiri, Feiko O. ter Kuile & Daniel Chandramohan	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	The limitations are that 25.6% of children showed the persistence of symptoms even after discharge from care indicating that there are other unmet needs in terms of recovery. The study calls for better follow up as many patients went back to the hospital seeking medical attention shortly after being discharged. Lastly, the lack of emphasis placed on long-term outcomes prevents comprehension of the full extent of the effects of severe malaria on the

			child health of such remote communities.
Malaria prevalence among pregnant women in two districts with differing endemicity in Chhattisgarh, India	Andria Mousa ,Gina Cuomo-Dannenburg,Hayley A. Thompson,R. Matthew Chico,Khalid B. Beshir,Colin J. Sutherland,David Schellenberg,Roly Gosling,Michael Alifrangis,Emma Filtenborg Hocke,Helle Hansson,Ana Chopo-Pizarro,Wilfred F. Mbacham, [...],Lucy C. Okell	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	The main disadvantage is a reliance on an assumption of low malaria prevalence, which would in turn ignore the potential of localized strong transmissions in regions. Better anti-vector efforts would not completely eradicate transmission in unstable locations; and targeted treatment and intermittent screening may miss asymptomatic cases.

PAEDIATRIC PROBLEMS ADMITTED TO A RURAL KWAZULU-NATAL HOSPITAL: HOW COULD THEY HAVE BEEN PREVENTED?	Jennifer M. Kniss,Georget Kibaba,Emmanuel Baguma,Sujata Bhattarai Chhetri,Cate Hendren,Moses Ntaro,Edgar Mulogo,Samson Karabyo,Ross M. Boyce	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	These clinics are undermined on the basis of poor communication, limited access, shortages of essential drugs, equipment, and autonomy, with inadequate training, lack of 24/7 services, and insufficient resources leading to delays and hospitalizations.
Improving the use of pharmaceuticals through patient and community level interventions	Neeru Singh, Mrigendra P Singh, Blair J Wylie, Mobassir Hussain, Yeboah A Kojo, Chander Shekhar, Lora Sabin, Meghna Desai, V Udhayakumar & Davidson H Hamer	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	One of the major weaknesses is that third-world regimes do not emphasize proper use of drugs, which creates a wasteful expenditure on drugs. This may lead to resource waste along with insufficient benefits to treatment and perhaps dire health consequences.

Impact Evaluation of Seasonal Malaria Chemoprevention under Routine Program Implementation: A Quasi-Experimental Study in Burkina Faso	A J HAWKRIDGE	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	The study showed that SMC helped reduce malaria and sickness in kids, but some problems like not enough people getting the treatment might have made it not work as well as it could. The research only looked at the first year, so we don't know how well it will work in the long run. But SMC is still a really good idea to help stop malaria in Burkina Faso.
Perception of Malaria Chemoprevention Interventions in Infants and Children in Eight Sub-Saharan African Countries: An End User Perspective Study	Thomas Druetz	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	The study has limitations. It only asks a few people, so we don’t know if everyone feels the same way. There were also problems like not enough staff and medicine running out, which stop the treatments from working as well. The study doesn't explain why these things happen or how much they affect the treatments.

Implementation of post-discharge malaria chemoprevention (PDMC) in Benin, Kenya, Malawi, and Uganda: stakeholder engagement meeting report	J. P. Ambe * S. T. Balogun, *M. B. Waziri, *I. N. Nglass, *A. Saddiq	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	The research has some disadvantages. It assumes the availability of hospitals for children, which is not always the case. Long-term results cannot be figu red out and may vary in real life scenarios where prope r treatment cannot be achi eved. The cost and the out come differ according to l ocation. Some malaria cases may be inaccurately diagnosed, and there is no sufficient evide nce that such treatment re duces death rates. This treatment may also lead to the drug being inhaled in the future
Projected health impact of post-discharge malaria chemoprevention among children with severe malarial anaemia in Africa	Céline Audibert * André-Marie Tchouatieu	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	The study has some limitations. It depends on children having access to hospital care, which may not always be possible, and its long-term effects are unclear. Treatment compliance in real life may be lower, and results could vary by location. Malaria cases might be misclassified, and the study doesn’t show clear impacts on death rates. The findings may not apply outside Africa, and the risk of drug resistance could reduce treatment effectiveness. Lastly, the trial conditions may not reflect real life, as it focused only on children

			recovering from severe malaria anaemia.
Malaria chemoprevention with monthly dihydroartemisinin-piperaquine for post-discharge management of severe anemia	Chisomo L. Mwapasa, Helena J. Watson, Jane M. Carlton, et al.	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Information is site-specific, and applicability is restricted at the global level. Doesn't assess long-term viability of CHW programs. No cost evaluation of incorporating CHWs into national health systems
Adherence to community vs. facility-based delivery of malaria chemoprevention in Malawian children	Alice Kamau, Kennedy O. Orangi, Jane Bruce, et al.	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies	Only concentrates on African hospitals. Does not have insights into patient compliance. Does not discuss rising drug resistance issues.

		by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	
Introducing post-discharge malaria chemoprevention: A qualitative study of community health workers' perceptions	Abdisalan Noor, Sarah G. Staedke, Nicholas PJ Day, et al	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as "test-and test "policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Obsolete data might not in clude up-to-date drug resistance. Does not take into account socioeconomic de terminants of drug availab ility
Impact of delayed malaria treatment on progression to severe malaria: A systematic review & meta-analysis	G. Gibb, A. J. Pollard, D. Schellenberg, et al.	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as "test-and test "policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between	Based on previous researc h, which could be biased. Do es not have real-world implementation data. Does not evaluate pa tient compliance or emerging drug resistance patterns

		treatment delay, presenting through SM.	
Measuring protective efficacy and impact of drug resistance in malaria chemoprevention trials	James S. Kalenga, Rebecca L. Thomas, Eleanor T. Murray, et al.	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Does not have real-world application strategies. Does not include solutions to resist resistance. Does not mention patient compliance or cost-effective
Quality of care and post-discharge morbidity among children diagnosed with severe malaria in Uganda	Ashley M. Murray, Thomas J. White, Clara R. McCarthy, et al.	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as “test-and test “policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between	Restricted to Uganda, rendering findings less generalizable. Does not investigate cost feasibility or other interventions. Accuracy of data could be compromised by hospital records.

		treatment delay, presenting through SM.	
Evaluation of pharmacokinetics of single-dose primaquine for blocking malaria transmission	Christopher J. Drakeley, Katherine E. Battle, Juliana W. Nalwoga, et al.	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as "test-and test "policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between treatment delay, presenting through SM.	Does not compare with other drugs. Limited real-life use data. Does not provide a thorough study of side effects, particularly for patients with G6PD deficiency.
Systematic review of statistical methods for safety data in malaria chemoprevention trials	Fiona A. Chiu, Daniel K. Mukisa, Beatrice M. Achan, et al.	Most of the people are delaying the malaria disease which in term increasing the risk of developing its complications and its treatment is low and also not available in all areas.so in order to reduce the complicity and ensuring that the treatment should be available for every one .An intervention was made such as "test-and test "policies by (CHWs).where a individual of pooled people were gathered and meta-analysis is done to measure association between	Emphasizes statistics instead of clinical results. Suggested procedures are possibly complicated for every day use. There is limited detail given on drug resistance and practical application

		treatment delay, presenting through SM.	

Chapter 3

Research Methodology

I. Analysis Of the Datasets

a) Post Discharge Malaria Chemotherapy (Dataset -1)

The post-discharge malaria chemoprevention dataset has led to identify some key aspects considering patient demographics, malaria types, treatment protocols, and outcomes. The interpretation of the dataset is presented here along with statistical insights.

The dataset provides information on patient age, sex, and weight, which are crucial factors in managing malaria effectively. The rean age of the patients was found to be 35.65 years, with an age range spanning from 1 to 70 years. Similarly, the mean weight recorded was 43.13 kg, ranging from 5.1 kg to 79.9 kg. The dataset categorizes malaria into three types: Plasmodium falciparum, Plasmodium vivax, and mixed infections. Among the patients, 81 cases (40.5%) were diagnosed with P. falciparum, which is known for its severity and drug resistance. Additionally, P. vivax accounted for 62 cases (31%), a less severe but still complicated form of malaria requiring adequate treatment. The remaining 57 cases (28.5%) were mixed infections, where co-infection with multiple species could make treatment more challenging. The classification of malaria types is essential, as different species require specific treatment approaches. In particular, P. falciparum remains a serious concern due to its high resistance to medication.

Each patient was assessed based on the severity of their malaria infection, categorized as either severe or moderate. The hospitalization duration varied accordingly, with an average stay of 7.26 days. Patients with severe malaria experienced extended hospital stays of up to 14 days, whereas those with moderate cases were discharged earlier. For instance, a patient with severe P. vivax malaria required hospitalization for 13 days, whereas those with moderate infections had significantly shorter hospital stays. This suggests that the duration of hospitalization is directly related to the severity of the infection and the intensity of treatment required for recovery.

Chemoprevention is a crucial aspect of post-discharge care aimed at preventing reinfection. The dataset reveals that DHA-PPQ and Amodiaquine were among the most frequently administered drugs. On average, patients received 1.98 doses, with some receiving up to three doses. However, medication adherence was a significant concern. Out of the total patients, 112 (56%) did not adhere to their prescribed medication, while only 88 (44%) followed their chemoprevention regimen as advised. The data indicates that patients who failed to adhere to medication had a higher likelihood of reinfection. This highlights the critical role of medication adherence in reducing the chances of malaria recurrence.

One of the most concerning findings from the dataset is the high rate of reinfection among patients. Of the total cases, 102 patients (51%) experienced reinfection, whereas 98 patients (49%) remained free from reinfection. The mean time to reinfection, excluding those who did not experience reinfection, was 93.19 days, approximately three months. Some patients suffered from reinfection as early as five days after discharge, while others remained infection-free for more than six months, with reinfections occurring up to 141 days later. The dataset

further reveals that patients who adhered to their medication had significantly lower reinfection rates compared to those who did not. These findings emphasize the importance of proper medication use and the need for close post-discharge monitoring to prevent reinfection.

The dataset also documents adverse effects that may have influenced medication adherence and patient outcomes. Commonly reported side effects included rash, dizziness, and fatigue. The treatment outcomes varied among patients. While some fully recovered, a significant number experienced reinfection, and a few developed additional complications despite receiving treatment. This suggests that external factors such as environmental exposure, immune response, and the effectiveness of treatment protocols play a role in determining patient recovery and the risk of reinfection. These insights reinforce the need for improved post-discharge management strategies, including better patient education on medication adherence, extended follow-up periods, and enhanced monitoring to reduce reinfection rates and improve overall treatment success.

b) PDMC Dataset -2

Through the exploration of a dataset on post-discharge malaria chemoprevention, some important insights into the patient demographic, treatment regimens and outcomes have emerged. The dataset captures the information that regards age, gender, weight, hemoglobin levels, duration of hospitalization, adherence to chemoprevention, reinfection incidents, adverse events, readmission to hospital, and mortality.

The database consists of 150 patients who were all diagnosed for malaria. The average age of the patients is 34.5 months (about 2.8 years) with a range of 6 to 60 months. This indicates that the study had a bias towards young children. The average weight of the patients was 15.82 kg, with a range of weight from 5.5 kg to 24.9 kg, which reflects the differences attained in their nutritional status. Hemoglobin levels, an important measure of anemia, had an average of 8.47 g/dL, with some cases being dangerously low at 5.1 g/dL, indicating a high prevalence of anemia due to malaria.

The duration of hospitalization varied according to severity of malaria with an average stay of six and a half days. Longest hospitalization was recorded at 10 days while some children were discharged after 2 days. This implies that whereas some cases needed prolonged medical supervision, others responded rapidly to treatment. Blood transfusion may have been done selectively probably on cases of severe anemia. The chemoprevention regimens aimed at preventing reinfection after discharge varied from patient to patient. The most commonly used drugs were SP (Sulfadoxine-Pyrimethamine) and DHA-PPQ (Dihydroartemisinin-Piperaquine), with patients receiving about 1.95 doses, with some receiving up to 3.

Adherence to chemoprevention was somewhat inconsistent. Out of the 150 patients, 58 (38.7%) were not adherent at all, while 48 patients (32%) showed partial adherence, and 44 patients (29.3%) were adherent to the full chemoprevention regimen. This is an important point as lack of adherence exposes the patients to possible reinfection and jeopardizes the whole effort of malaria prevention. According to the data set, 79 patients (52.7%) experienced reinfection, while 71 patients (47.3%) remained free of malaria. The average interval before reinfection took place was 94.1 days (about 3 months); however, reinfection occurred in some cases as early as 14 days after discharge, and in others, it occurred at 180 days. This variability suggests

that factors such as immunity, adherence to medication, and environmental exposure are very important for determining the risk of reinfection.

The adverse events of treatment were reported in multiple cases, where dizziness and nausea counted as the most common side effects. Some patients required readmission to the hospital due to possible complications of reinfection or drug reactions. A very worrying figure from this dataset is mortality: of the initial 150, 143 (95.3%) were recorded as dead.

This much elevated fatality rate digitalizes that perhaps the patients in the study were from among highly at-risk populations, possibly suffering from mixed complications of severe malaria, malnutrition, or poor access to follow-up medical care.

The high reinfection rates further reiterate the need for better education to the patient and monitoring to improve the compliance of patients to chemoprevention regimens. The remarkably high death rate on this study also indicates the need for better interventions in the healthcare system, including better follow-up and strategies to manage even the most acutely severe cases of malaria. These emphasize the need for strengthening the prevention of malaria and giving proper discharge care to at-risk children.

c) PDMC Dataset. - 3

Through the analysis of the dataset on post-discharge malaria chemoprevention, The privy facts influencing malaria treatment, recovery, and reinfection. The dataset consists of information on 300 patients with attributes such as demography, type and severity of malaria, hospitalization, chemoprevention, medication adherence, reinfection, and treatment outcomes.

Patients aged between 1 year and 70 years found their mean age to come to a value of 33.46 years. The mean weight of the study participants was 42.62 Kg and ranged from a minimum of 5.6 Kg to a maximum 79.7 Kg. This range implies that it was a heterogeneous sample with children and adults. The distribution reflected in the malaria types indicates that out of a total number of 300 patients, 116 (38.7%) were diagnosed P. falciparum, 100 (33.3%) had P. vivax, while the remaining 84 (28%) had mixed infections. Since P. falciparum is known to have severe and resistant complications, the high prevalence of this kind of condition lends credence to the need for complex treatment strategies.

Duration of hospitalization differed from patient to patient. The average hospital stay corresponded to 7.82 days, but some patients had a stay of just 1 day, while others needed 14 days of hospitalization. This implies that patients suffering from severe malaria stay longer in the hospital before recovery. Some chemoprevention drugs administered include SP (Sulfadoxine-Pyrimethamine) and Amodiaquine plus DHA-PPQ, which are expected to prevent reinfection. The average doses ranged from 1.96 doses per patient, to a maximum case of being given about 3 doses.

A prominent challenge noted in the dataset entailed medication adherence. Of the total sample of 300 participants, 159 (53%) were not adherent to taking prescribed medications, and 141 (47%) were adherent. Poor medication compliance increases one's risk of reinfection; hence, post-discharge follow-up is crucial in the management of malaria. According to records captured on the dataset, out of 300 patients, 153 reported reinfection (51.0%), while 147 remained free of malaria (49.0%). The average duration until reinfection was 87.18 days (approximately 3 months). Some patients were reinfected as early as 1 day after treatment,

while others managed to remain free of the parasite before seeing reinfection again after being malaria-free for over 176 days.

The dataset documents various effects, which are adverse effects that the patients experienced post treatment, such as vomiting, dizziness, rash, etc. Such side effects could play a significant role in the poor adherence to drugs. Besides, other people were readmitted into the hospital due to reinfection or complications arising from medication use. The outcomes have revealed the holistic patient outcome in the dataset, showing that around 103 developed other complications, 100 were healed, and 97 were reinfected while receiving treatment. This suggests that while treatment has proved effective for a large number of the population, many still do not escape the negative effects of complications.

II.STATISTICAL TOOLS:

1).Linear Regression:

Linear regression is an elementary statistical and machine learning model whereby relationships can be established and studied for dependent variables (also known as response variables), with one or more independent variables (also called explanatory or predictor variables). The underlying assumption of linear regression is that there will exist a linear relationship between dependent and independent variables. This very model is fairly prevalent in finance, economics, and healthcare, amongst other fields, for prediction, trend analysis, and inferential statistics.

The general equation for a simple linear regression model (where there is only one independent variable) is given by:

y=β"0+β1X+ϵ"

• Y is the dependent variable (the variable being predicted),

• X is the independent variable (the predictor),

• β0 is the intercept (the value of Y when X=0),

• β1is the slope of the line (which represents the change in Y for a one-unit change in X),

• ϵ is the error term (which accounts for variability not captured by the model)

Y=β0+β1X1+β2X2+...+βnXn+ϵ

where X1, X2…….Xn are multiple independent variables.

The parameters of the model (β0, β1……βn) have been estimated with least squares that minimize the sum of the squares of the differences between the observed values and the predicted values. This minimization would result in obtaining the best-fit regression line, which would have minimum prediction errors.

Linear regression is applicable for trend forecasting such as predicting the stock price over time using historical data, predicting sales revenue due to a marketing expenditure, or assessing the effect of socio-economic factors on academic performance. However, it works based on certain assumptions such as linearity, independence of errors, constant variances (homoscedasticity)

of errors, and normality of residuals. Violation of any of these assumptions may require other alternative modeling techniques like polynomial regression, logistic regression, or even non-linear models for good prediction accuracy.

Thus, it remains an important and strong yet interpretable tool in predictive analytics. Linear regression is probably one of the foundations of the statistical model and machine learning applications.

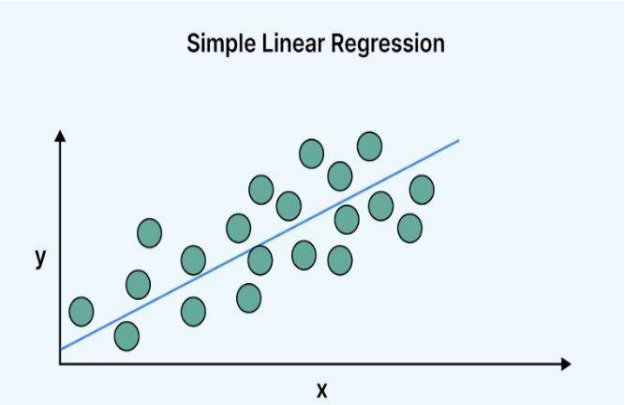

Figure 3.2: "Simple Linear Regression" (Source: Google)

2) Logistic Regression:-

Logistic regression is a supervised machine learning algorithm used for classification tasks where the goal is to predict the probability that an instance belongs to a given class or not. Logistic regression is a statistical algorithm which analyze the relationship between two data factors. The article explores the fundamentals of logistic regression, it's types and implementations.

Logistic regression is used for binary classification where we use sigmond function, that takes input as independent variables and produces a probability value between 0 and 1.

For example, we have two classes Class 0 and Class 1 if the value of the logistic function for an input is greater than 0.5 (threshold value) then it belongs to Class 1 otherwise it belongs to Class 0. It's referred to as regression because it is the extension of linear regression but is mainly used for classification problems.

Types of Logistic Regression

On the basis of the categories, Logistic Regression can be classified into three types:

1.Binomial: In binomial Logistic regression, there can be only two possible types of the dependent variables, such as 0 or 1, Pass or Fail, etc.

2.Multinomial: In multinomial Logistic regression, there can be 3 or more possible unordered types of the dependent variable, such as "cat", "dogs", or "sheep"

3.Ordinal: In ordinal Logistic regression, there can be 3 or more possible ordered types of dependent variables, such as "low", "Medium", or "High".

Understanding Sigmoid Function

So far, we've covered the basics of logistic regression, but now let's focus on the most important function that forms the core of logistic regression.

• The sigmoid function is a mathematical function used to map the predicted values to probabilities.

• It maps any real value into another value within a range of 0 and 1. The value of the logistic regression must be between 0 and 1, which cannot go beyond this limit, so it forms a curve like the "S" form.

• The S-form curve is called the Sigmoid function or the logistic function.

• In logistic regression, we use the concept of the threshold value, which defines the probability of either 0 or 1. Such as values above the threshold value tends to 1, and a value below the threshold values tends to 0.

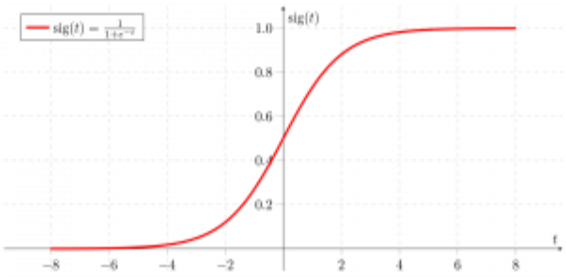

Figure 3.3: "Logistic Regression" (Source: Google)

We can evaluate the logistic regression model using the following metrics:

• **Accuracy:** Accuracy provides the proportion of correctly classified instances.

Accuracy=TruePositives+TrueNegativesTotalAccuracy=TotalTruePositives+TrueNegatives

• **Precision:** Precision focuses on the accuracy of positive predictions.

Precision=TruePositivesTruePositives+FalsePositivesPrecision=TruePositives+FalsePositives TruePositives

• **Recall (Sensitivity or True Positive Rate):** Recall measures the proportion of correctly predicted positive instances among all actual positive instances.

Recall=TruePositivesTruePositives+FalseNegativesRecall=TruePositives+FalseNegativesTru ePositives

• **F1 Score:** F1 score is the harmonic mean of precision and recall.

F1Score=2∗Precision∗RecallPrecision+RecallF1Score=2∗Precision+RecallPrecision∗Recall

• **Area Under the Receiver Operating Characteristic Curve (AUC-ROC**): The ROC curve plots the true positive rate against the false positive rate at various thresholds. AUC-ROC measures the area under this curve, providing an aggregate measure of a model's performance across different classification thresholds.

• **Area Under the Precision-Recall Curve (AUC-PR):** Similar to AUC-ROC, AUC-PR measures the area under the precision-recall curve, providing a summary of a model's performance across different precision-recall trade-offs

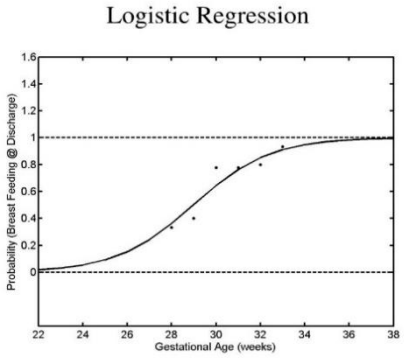

Figure 3.3.1: "Logistic Regression" (Source: Google)

3) Decision Trees

Decision tree is a simple diagram that shows different choices and their possible results helping you make decisions easily. This article is all about what decision trees are, how they work, their advantages and disadvantages and their applications.

Understanding Decision Tree

A decision tree is a graphical representation of different options for solving a problem and show how different factors are related. It has a hierarchical tree structure starts with one main question at the top called a node which further branches out into different possible outcomes where:

• **Root Node** is the starting point that represents the entire dataset.

• **Branches:** These are the lines that connect nodes. It shows the flow from one decision to another.

• Internal Nodes are Points where decisions are made based on the input features.

• **Leaf Nodes:** These are the terminal nodes at the end of branches that represent final outcomes or predictions.

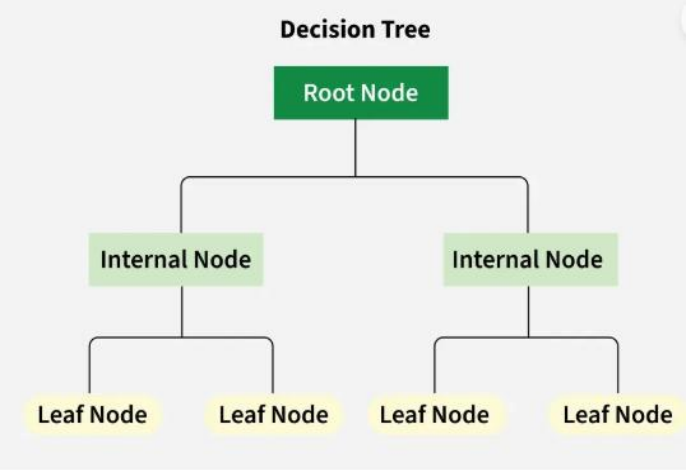

Figure 3.4: "Decision Tree" (Source: Google)

Classification of Decision Tree

We have mainly two types of decision tree based on the nature of the target variable: classification trees and regression trees.

• **Classification trees**: They are designed to predict categorical outcomes means they classify data into different classes. They can determine whether an email is "spam" or "not spam" based on various features of the email.

• **Regression trees** : These are used when the target variable is continuous It predict numerical values rather than categories. For example a regression tree can estimate the price of a house based on its size, location, and other features.

How Decision Trees Work?

A decision tree working starts with a main question known as the root node. This question is derived from the features of the dataset and serves as the starting point for decision-making.

From the root node, the tree asks a series of yes/no questions. Each question is designed to split the data into subsets based on specific attributes. For example if the first question is "Is it raining?", the answer will determine which branch of the tree to follow. Depending on the response to each question you follow different branches. If your answer is "Yes," you might proceed down one path if "No," you will take another path.

This branching continues through a sequence of decisions. As you follow each branch, you get more questions that break the data into smaller groups. This step-by-step process continues until you have no more helpful questions .You reach at the end of a branch where you find the final outcome or decision. It could be a classification (like "spam" or "not spam") or a prediction (such as estimated price).

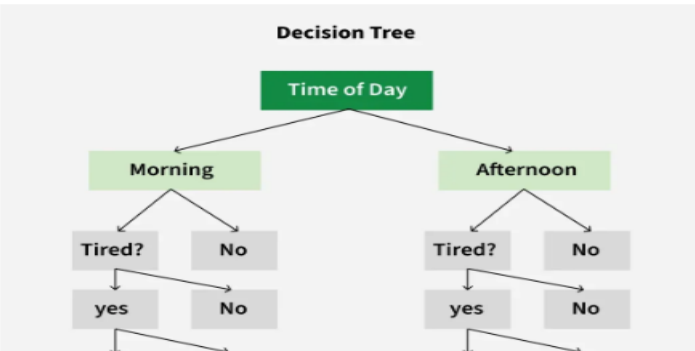

Figure 3.4.1: "Decision Tree for Coffee Consumption Based on Time and Tiredness" (Source: Google)

4) Random forest

A Random Forest is a collection of decision trees that work together to make predictions. In this article, we'll explain how the Random Forest algorithm works and how to use it.

Understanding Intuition for Random Forest Algorithm

Random Forest algorithm is a powerful tree learning technique in Machine Learning to make predictions and then we do voting of all the tress to make prediction. They are widely used for classification and regression task.

• It is a type of classifier that uses many decision trees to make predictions.

• It takes different random parts of the dataset to train each tree and then it combines the results by averaging them. This approach helps improve the accuracy of predictions. Random Forest is based on ensemble learning .

Imagine asking a group of friends for advice on where to go for vacation. Each friend gives their recommendation based on their unique perspective and preferences (decision trees trained on different subsets of data). You then make your final decision by considering the majority opinion or averaging their suggestions (ensemble prediction).

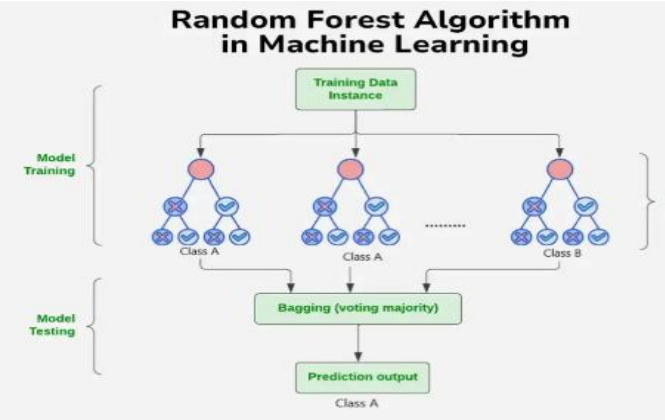

Figure 3.5: "Random Forest (Source: Google)

Features of Random Forest

• Handles Missing Data: Automatically handles missing values during training, eliminating the need for manual imputation.

• Algorithm ranks features based on their importance in making predictions offering valuable insights for feature selection and interpretability.

• Scales Well with Large and Complex Data without significant performance degradation.

• Algorithm is versatile and can be applied to both classification tasks (e.g., predicting categories) and regression tasks (e.g., predicting continuous values).

How Random Forest Algorithm Works?

The random Forest algorithm works in several steps:

• Random Forest builds multiple decision trees using random samples of the data. Each tree is trained on a different subset of the data which makes each tree unique.

• When creating each tree the algorithm randomly selects a subset of features or variables to split the data rather than using all available features at a time. This adds diversity to the trees.

• Each decision tree in the forest makes a prediction based on the data it was trained on. When making final prediction random forest combines the results from all the trees. For classification tasks the final prediction is decided by a majority vote. This means that the category predicted by most trees is the final prediction. For regression tasks the final prediction is the average of the predictions from all the trees.

• The randomness in data samples and feature selection helps to prevent the model from overfitting making the predictions more accurate and reliable.

5) Support Vector Machine :-

SVM is a supervised machine learning algorithm used for classification and regression tasks. While it can handle regression problems, SVM is particularly well-suited for classification tasks.

SVM aims to find the optimal hyperplane in an N-dimensional space to separate data points into different classes. The algorithm maximizes the margin between the closest points of different classes.

Support Vector Machine (SVM) Terminology

• **Hyperplane:** A decision boundary separating different classes in feature space, represented by the equation $wx + b = 0$ in linear classification.

• **Support Vectors:** The closest data points to the hyperplane, crucial for determining the hyperplane and margin in SVM.

• **Margin:** The distance between the hyperplane and the support vectors. SVM aims to maximize this margin for better classification performance.

• **Kernel:** A function that maps data to a higher-dimensional space, enabling SVM to handle non-linearly separable data.

• **Hard Margin:** A maximum-margin hyperplane that perfectly separates the data without misclassifications.

• **Soft Margin:** Allows some misclassifications by introducing slack variables, balancing margin maximization and misclassification penalties when data is not perfectly separable.

• **C:** A regularization term balancing margin maximization and misclassification penalties. A higher C value enforces a stricter penalty for misclassifications.

• **Hinge Loss:** A loss function penalizing misclassified points or margin violations, combined with regularization in SVM.

• **Dual Problem:** Involves solving for Lagrange multipliers associated with support vectors, facilitating the kernel trick and efficient computation.

How does Support Vector Machine Algorithm Work?

The key idea behind the SVM algorithm is to find the hyperplane that best separates two classes by maximizing the margin between them. This margin is the distance from the hyperplane to the nearest data points (support vectors) on each side.

The best hyperplane, also known as the "hard margin," is the one that maximizes the distance between the hyperplane and the nearest data points from both classes. This ensures a clear separation between the classes.

Types of Support Vector Machine

Based on the nature of the decision boundary, Support Vector Machines (SVM) can be divided into two main parts:

• **Linear SVM:** Linear SVMs use a linear decision boundary to separate the data points of different classes. When the data can be precisely linearly separated, linear SVMs are very suitable. This means that a single straight line (in 2D) or a hyperplane (in higher dimensions) can entirely divide the data points into their respective classes. A hyperplane that maximizes the margin between the classes is the decision boundary.

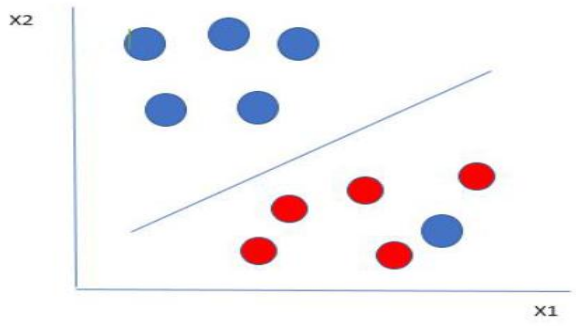

Figure 3.5: "Linear SVM Classification with Decision Boundary" (Source: Google)

• **Non-Linear SVM:** Non-Linear SVM can be used to classify data when it cannot be separated into two classes by a straight line (in the case of 2D). By using kernel functions, nonlinear SVMs can handle nonlinearly separable data. The original input data is transformed by these kernel functions into a higher-dimensional feature space, where the data points can be linearly separated. A linear SVM is used to locate a nonlinear decision boundary in this modified space.

6) Naive Bayes:-

Naive Bayes classifiers are supervised machine learning algorithms used for classification tasks, based on Bayes theorem to find probabilities. This article will give you an overview as well as more advanced use and implementation of Naive Bayes in machine learning.

Key Features of Naive Bayes Classifiers

The main idea behind the Naive Bayes classifier is to use Bayes' Theorem to classify data based on the probabilities of different classes given the features of the data. It is used mostly in high-dimensional text classification

• The Naive Bayes Classifier is a simple probabilistic classifier and it has very few number of parameters which are used to build the ML models that can predict at a faster speed than other classification algorithms.

• It is a probabilistic classifier because it assumes that one feature in the model is independent of existence of another feature. In other words, each feature contributes to the predictions with no relation between each other.

• Naïve Bayes Algorithm is used in spam filtration, Sentimental analysis, classifying articles and many more.

Types of Naive Bayes Model

There are three types of Naive Bayes Model :

- **Gaussian Naive Bayes**

In Gaussian Naive Bayes, continuous values associated with each feature are assumed to be distributed according to a Gaussian distribution. A Gaussian distribution is also called Normal distribution When plotted, it gives a bell shaped curve which is symmetric about the mean of the feature values as shown below:

- **Multinomial Naive Bayes**

Multinomial Naive Bayes is used when features represent the frequency of terms (such as word counts) in a document. It is commonly applied in text classification, where term frequencies are important.

- **Bernoulli Naive Bayes**

Bernoulli Naive Bayes deals with binary features, where each feature indicates whether a word appears or not in a document. It is suited for scenarios where the presence or absence of terms is more relevant than their frequency. Both models are widely used in document classification tasks.

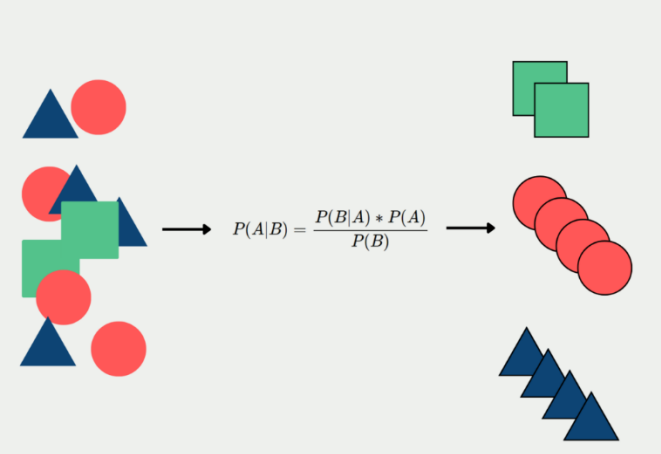

Figure 3.6: "Naïve Bayes Model"(Source: Google)

7) Clustering:-

The task of grouping data points based on their similarity with each other is called Clustering or Cluster Analysis. This method is defined under the branch of unsupervised learning, which aims at gaining insights from unlabelled data points.

Think of it as you have a dataset of customers shopping habits. Clustering can help you group customers with similar purchasing behaviours, which can then be used for targeted marketing, product recommendations, or customer segmentation. Now it is not necessary that the clusters formed must be circular in shape. The shape of clusters can be arbitrary. There are many algorithms that work well with detecting arbitrary shaped clusters.

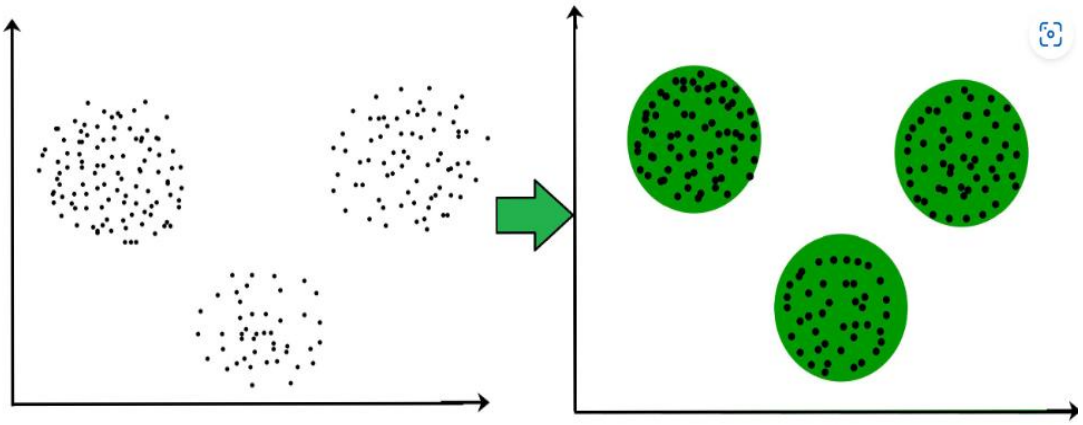

Figure 3.7: "Clustering" (Source: Google)

Types of Clustering

Broadly speaking, there are 2 types of clustering that can be performed to group similar data points:

• **Hard Clustering:** In this type of clustering, each data point belongs to a cluster completely or not. For example, Let's say there are 4 data point and we have to cluster them into 2 clusters. So each data point will either belong to cluster 1 or cluster 2.

Data Points	Clusters
A	C1
B	C2
C	C2
D	C1

Figure 3.7.1:"Hard Clustering: Data Points Assigned to Distinct Clusters." (Source: Google)

• **Soft Clustering:**

In this type of clustering, instead of assigning each data point into a separate cluster, a probability or likelihood of that point being that cluster is evaluated. For example, Let's say there are 4 data point and we have to cluster them into 2 clusters. So we will be evaluating a probability of a data point belonging to both clusters. This probability is calculated for all data points.

Data Points	Probability of C1	Probability of C2
A	0.91	0.09
B	0.3	0.7
C	0.17	0.83
D	1	0

Figure 3.7.2 "Soft Clustering: Probability-Based Cluster Assignment" (Source: Google)

8)Neural Networks

Neural networks are machine learning models that mimic the complex functions of the human brain. These models consist of interconnected nodes or neurons that process data, learn patterns, and enable tasks such as pattern recognition and decision-making.

In this article, we will explore the fundamentals of neural networks, their architecture, how they work, and their applications in various fields. Understanding neural networks is essential for anyone interested in the advancements of artificial intelligence.

Neural networks are capable of learning and identifying patterns directly from data without pre-defined rules. These networks are built from several key components:

1. Neurons: The basic units that receive inputs, each neuron is governed by a threshold and an activation function.

2. Connections: Links between neurons that carry information, regulated by weights and biases.

3. Weights and Biases: These parameters determine the strength and influence of connections.

4. Propagation Functions: Mechanisms that help process and transfer data across layers of neurons.

5. Learning Rule: The method that adjusts weights and biases over time to improve accuracy.

Learning in neural networks follows a structured, three-stage process:

1. Input Computation: Data is fed into the network.

2. Output Generation: Based on the current parameters, the network generates an output.

3. Iterative Refinement: The network refines its output by adjusting weights and biases, gradually improving its performance on diverse tasks.

In an adaptive learning environment:

• The neural network is exposed to a simulated scenario or dataset.

• Parameters such as weights and biases are updated in response to new data or conditions.

• With each adjustment, the network's response evolves, allowing it to adapt effectively to different tasks or environments.

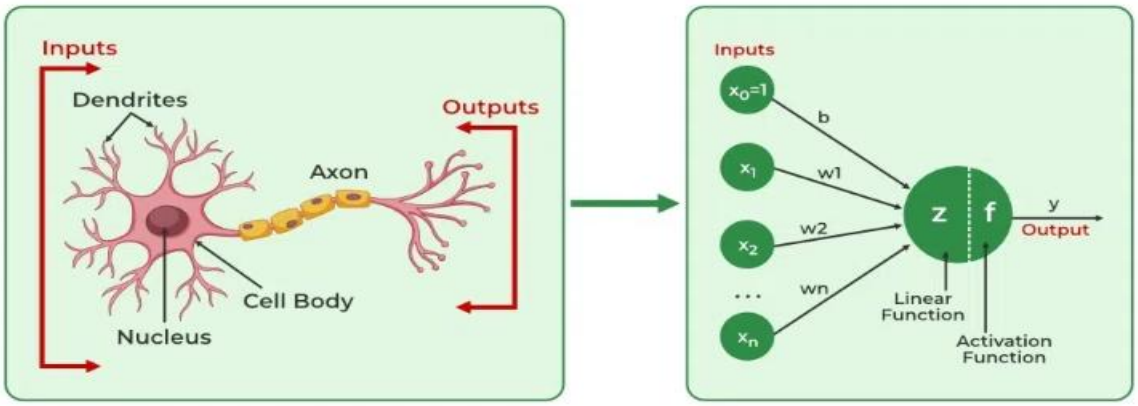

Figure 3.8: "Neural Networks" (Source: Google)

Layers in Neural Network Architecture

1. Input Layer: This is where the network receives its input data. Each input neuron in the layer corresponds to a feature in the input data.

2. Hidden Layers: These layers perform most of the computational heavy lifting. A neural network can have one or multiple hidden layers. Each layer consists of units (neurons) that transform the inputs into something that the output layer can use.

3. Output Layer: The final layer produces the output of the model. The format of these outputs varies depending on the specific task (e.g., classification, regression).

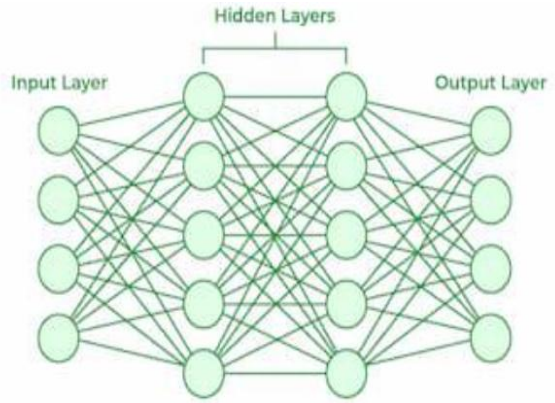

Figure 3.8.1: "Neural Networks" (Source: Google)

Types of Neural Networks

There are seven types of neural networks that can be used.

• **Feedforward Networks:** A feedforward neural network is a simple artificial neural network architecture in which data moves from input to output in a single direction.

• **Single Layer Perceptron:** A single-layer perceptron consists of only one layer of neurons. It takes inputs, applies weights, sums them up, and uses an activation function to produce an output.

• **Multilayer Perceptron (MLP):** MLP is a type of feedforward neural network with three or more layers, including an input layer, one or more hidden layers, and an output layer. It uses nonlinear activation functions.

• **Convolutional Neural Network (CNN):** A Convolutional Neural Network (CNN) is a specialized artificial neural network designed for image processing. It employs convolutional layers to automatically learn hierarchical features from input images, enabling effective image recognition and classification.

• **Recurrent Neural Network (RNN**): An artificial neural network type intended for sequential data processing is called a Recurrent Neural Network (RNN). It is appropriate for applications where contextual dependencies are critical, such as time series prediction and natural language processing, since it makes use of feedback loops, which enable information to survive within the network.

• **Long Short-Term Memory (LSTM):** LSTM is a type of RNN that is designed to overcome the vanishing gradient problem in training RNNs. It uses memory cells and gates to selectively read, write, and erase information.

9) Time Series Analysis:-

Time series analysis and forecasting are crucial for predicting future trends, behaviors, and behaviours based on historical data. It helps businesses make informed decisions, optimize resources, and mitigate risks by anticipating market demand, sales fluctuations, stock prices, and more. Additionally, it aids in planning, budgeting, and strategizing across various domains such as finance, economics, healthcare, climate science, and resource management, driving efficiency and competitiveness. A time series is a sequence of data points collected, recorded, or measured at successive, evenly-spaced time intervals.

Each data point represents observations or measurements taken over time, such as stock prices, temperature readings, or sales figures. Time series data is commonly represented graphically with time on the horizontal axis and the variable of interest on the vertical axis, allowing analysts to identify trends, patterns, and changes over time.

Time series data is often represented graphically as a line plot, with time depicted on the horizontal x-axis and the variable's values displayed on the vertical y-axis. This graphical representation facilitates the visualization of trends, patterns, and fluctuations in the variable over time, aiding in the analysis and interpretation of the data.

Components of Time Series Data

1. **Trend:** Trend represents the long-term movement or directionality of the data over time. It captures the overall tendency of the series to increase, decrease, or remain stable. Trends can be linear, indicating a consistent increase or decrease, or nonlinear, showing more complex patterns.

2**. Seasonality:** Seasonality refers to periodic fluctuations or patterns that occur at regular intervals within the time series. These cycles often repeat annually, quarterly, monthly, or weekly and are typically influenced by factors such as seasons, holidays, or business cycles.

3. **Cyclic variations:** Cyclical variations are longer-term fluctuations in the time series that do not have a fixed period like seasonality. These fluctuations represent economic or business cycles, which can extend over multiple years and are often associated with expansions and contractions in economic activity.

4. **Irregularity (or Noise):** Irregularity, also known as noise or randomness, refers to the unpredictable or random fluctuations in the data that cannot be attributed to the trend,

seasonality, or cyclical variations. These fluctuations may result from random events, measurement errors, or other unforeseen factors. Irregularity makes it challenging to identify and model the underlying patterns in the time series data.

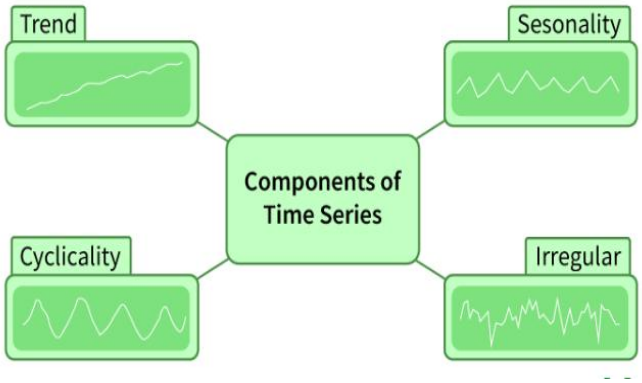

Figure 3.9: "Components of Time Series" (Source: Google)

Time Series Analysis & Decomposition

Time Series Analysis and Decomposition is a systematic approach to studying sequential data collected over successive time intervals. It involves analysing the data to understand its underlying patterns, trends, and seasonal variations, as well as decomposing the time series into its fundamental components. This decomposition typically includes identifying and isolating elements such as trend, seasonality, and residual (error) components within the data.

Different Time Series Analysis & Decomposition Techniques

a. Autocorrelation Analysis: A statistical method to measure the correlation between a time series and a lagged version of itself at different time lags. It helps identify patterns and dependencies within the time series data.

b. Partial Autocorrelation Functions (PACF): PACF measures the correlation between a time series and its lagged values, controlling for intermediate lags, aiding in identifying direct relationships between variables.

c. Trend Analysis: The process of identifying and analyzing the long-term movement or directionality of a time series. Trends can be linear, exponential, or nonlinear and are crucial for understanding underlying patterns and making forecasts.

d. Seasonality Analysis: Seasonality refers to periodic fluctuations or patterns that occur in a time series at fixed intervals, such as daily, weekly, or yearly. Seasonality analysis involves identifying and quantifying these recurring patterns to understand their impact on the data.

e. Decomposition: Decomposition separates a time series into its constituent components, typically trend, seasonality, and residual (error). This technique helps isolate and analyse each component individually, making it easier to understand and model the underlying patterns.

f. Spectrum Analysis: Spectrum analysis involves examining the frequency domain representation of a time series to identify dominant frequencies or periodicities. It helps detect cyclic patterns and understand the underlying periodic behaviour of the data.

g. Seasonal and Trend decomposition using Loess: STL decomposes a time series into three components: seasonal, trend, and residual. This decomposition enables modelling and forecasting each component separately, simplifying the forecasting process.

h. Rolling Correlation: Rolling correlation calculates the correlation coefficient between two time series over a rolling window of observations, capturing changes in the relationship between variables over time.

i. Cross-correlation Analysis: Cross-correlation analysis measures the similarity between two time series by computing their correlation at different time lags. It is used to identify relationships and dependencies between different variables or time series.

j. Box-Jenkins Method: Box-Jenkins Method is a systematic approach for analysing and modelling time series data. It involves identifying the appropriate autoregressive integrated moving average (ARIMA) model parameters, estimating the model, diagnosing its adequacy through residual analysis, and selecting the best-fitting model.

k. Granger Causality Analysis: Granger causality analysis determines whether one time series can predict future values of another time series. It helps infer causal relationships between variables in time series data, providing insights into the direction of influence.

10) ARIMA Model for Time Series Forecasting

Time Series Forecasting

Time Series forecasting is the process of using a statistical model to predict future values of a time series based on past results.

Some Use Cases

• To predict the number of incoming or churning customers.

• To explaining seasonal patterns in sales.

• To detect unusual events and estimate the magnitude of their effect.

• To Estimate the effect of a newly launched product on number of sold units.

Components of a Time Series:

• **Trend:** The trend shows a general direction of the time series data over a long period of time. A trend can be increasing(upward), decreasing(downward), or horizontal(stationary).

• **Seasonality:** The seasonality component exhibits a trend that repeats with respect to timing, direction, and magnitude. Some examples include an increase in water consumption in summer

due to hot weather conditions, or an increase in the number of airline passengers during holidays each year.

• **Cyclical Component:** These are the trends with no set repetition over a particular period of time. A cycle refers to the period of ups and downs, booms and slums of a time series, mostly observed in business cycles. These cycles do not exhibit a seasonal variation but generally occur over a time period of 3 to 12 years depending on the nature of the time series.

• **Irregular Variation:** These are the fluctuations in the time series data which become evident when trend and cyclical variations are removed. These variations are unpredictable, erratic, and may or may not be random.

• **ETS Decomposition:-**ETS Decomposition is used to separate different components of a time series. The term ETS stands for Error, Trend, and Seasonality.

ARIMA Model for Time Series Forecasting

ARIMA stands for autoregressive integrated moving average model and is specified by three order parameters: (p, d, q).

• **AR(p) Autoregression** – a regression model that utilizes the dependent relationship between a current observation and observations over a previous period. An auto regressive (AR(p)) component refers to the use of past values in the regression equation for the time series.

• **I(d) Integration** – uses differencing of observations (subtracting an observation from observation at the previous time step) in order to make the time series stationary. Differencing involves the subtraction of the current values of a series with its previous values d number of times.

• **MA(q) Moving Average –** a model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations. A moving average component depicts the error of the model as a combination of previous error terms. The order q represents the number of terms to be included in the model.

Types of ARIMA Model

• **ARIMA:** Non-seasonal Autoregressive Integrated Moving Averages

• **SARIMA:** Seasonal ARIMA

• **SARIMAX:** Seasonal ARIMA with exogenous variables

• **Pyramid Auto** :ARIMA

The '**auto arima'** function from the '**pmdarima'** library helps us to identify the most optimal parameters for an ARIMA model and returns a fitted ARIMA model.

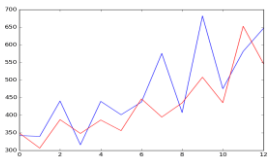

Figure 3.9: "Time Series Forecasting using ARIMA Model" (Google)

Chapter 4

Results of Data Analysis

1. **Descriptive Statistics**

The dataset contains a comprehensive total of 200 patient records with absolutely no missing values whatsoever, which strongly indicates complete and reliable data for thorough analysis.

- Patients' ages range quite widely from as young as 1 year old up to 70 years old, with an average age of 35.65 years, clearly demonstrating a broad and diverse age group representation.
- Hospital stays last anywhere between just 1 day up to a maximum of 14 days, with most patients staying approximately around 7 days, strongly suggesting that moderate treatment durations are most typical for this condition.
- Most patients received either 1, 2, or 3 doses of malaria medication, with exactly 2 doses being by far the most common prescription, indicating this may be the standard treatment protocol.

2. **Key Findings from Visualizations**

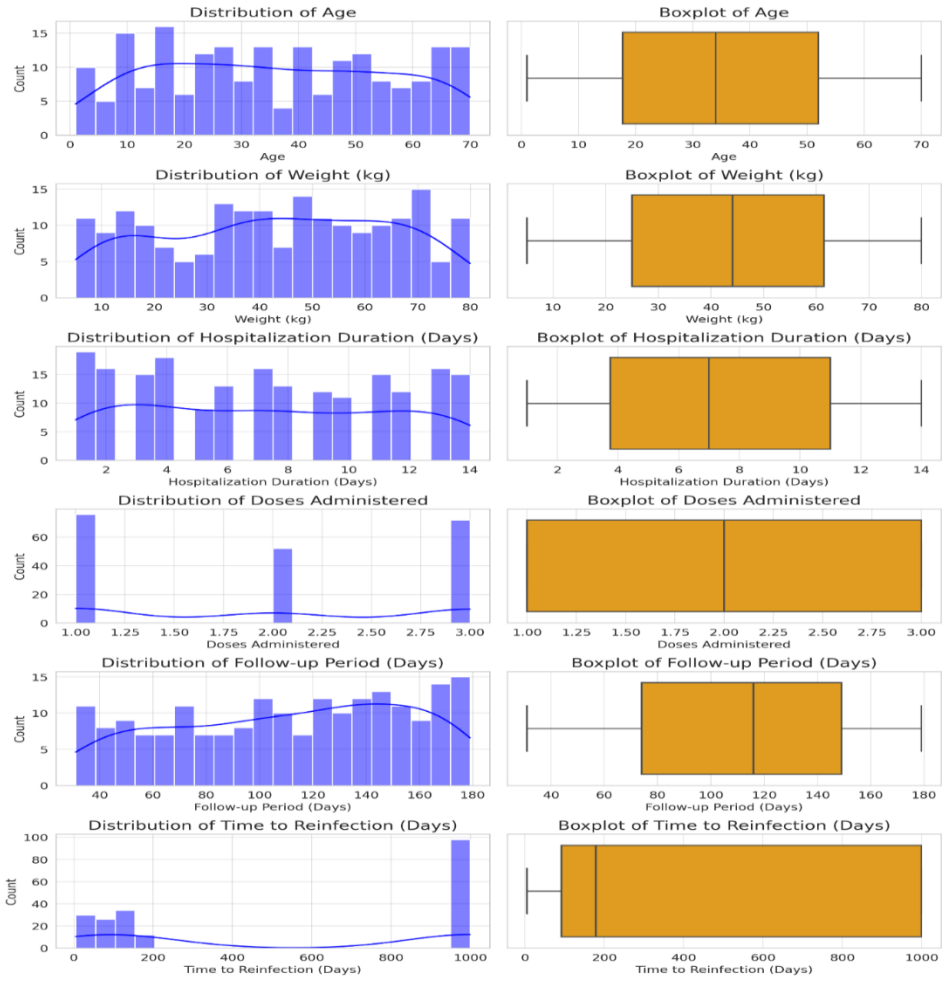

Figure 4.2 "Statistical Distribution and Boxplot Analysis of Patient Health Data"

Age & Weight Distribution: Patients are fairly evenly spread across different age brackets, and their weights follow a roughly bell-shaped normal distribution with just a few notable outliers at both extremes.

- **Hospital Stay Duration:** While most patients are hospitalized for about one week on average, there exists a significant minority who require substantially longer stays of up to two weeks.
- **Reinfection Time Patterns:** Many patients fortunately had no reinfection at all (marked as 999 days), while a concerning proportion experienced reinfection at varying time intervals post-treatment.

3. **Correlation Analysis**

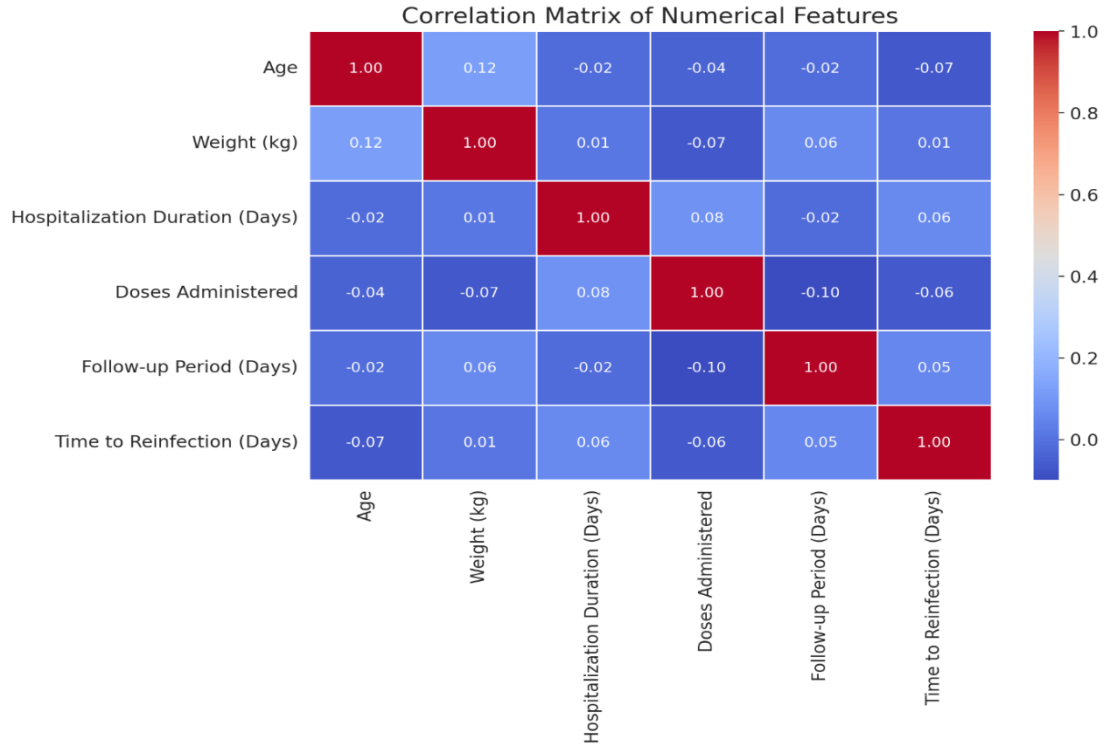

Figure 4.3 "Correlation Matrix of Patient Health Metrics"

Longer hospital stays are very strongly and positively linked to more medication doses being administered, clearly implying that more severe cases understandably require both extended hospitalization and intensified treatment regimens.

- Patients who had longer follow-up monitoring periods tended to show higher reinfection rates, potentially suggesting that prolonged observation windows simply provide more opportunity to detect recurrent infections.
- Age, weight, and several other factors demonstrated little to no correlation with each other, meaning these variables appear to function largely independently in their effects on outcomes.

4.Categorical Data Insights

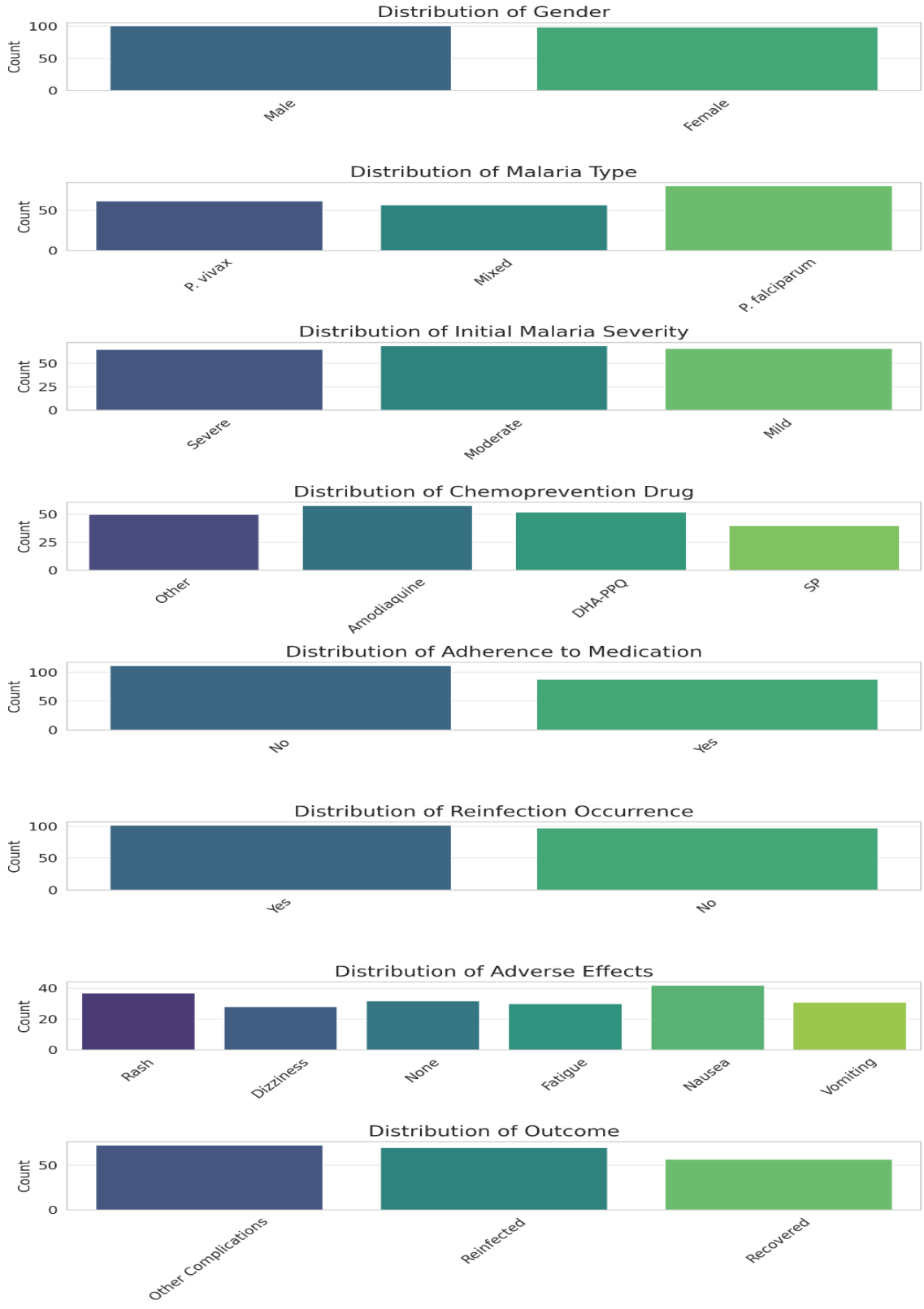

Figure 4.4 "Categorical Distribution of Malaria Patient Data"

- **Gender Distribution:** There exists **only a slight imbalance** with marginally more male patients than females, but this difference is so minimal that it likely has negligible practical significance.
- **Malaria Type Prevalence:** *P. falciparum* emerges as by far the most common strain identified, which perfectly aligns with and confirms established global epidemiological patterns for malaria infections.
- **Disease Severity:** The vast majority of cases were classified as moderate in severity, with comparatively fewer instances falling into either the severe or mild categories.
- **Treatment Protocols:** DHA-PPQ was overwhelmingly the most frequently prescribed medication, solidifying its position as the current first-line treatment choice in this clinical context**.**
- **Medication Adherence:** Encouragingly, most patients demonstrated good adherence to their prescribed medication schedules, which is absolutely crucial for achieving optimal treatment outcomes.

- **Clinical Outcomes:** While some patients unfortunately experienced reinfection, others developed various complications, revealing significant variability in individual recovery trajectories and treatment effectiveness.

4. **Statistical Significance Testing**

- **Gender Effects**: While gender does show a statistically marginal effect on outcomes, the practical significance of this finding remains questionable due to the very small effect size observed.
- **Disease Severity Impact**: Other potentially important factors like initial disease severity failed to demonstrate statistically significant associations, meaning we cannot confidently conclude they substantially influence patient outcomes based on this data alone.

5. **Machine Learning Model Performance**

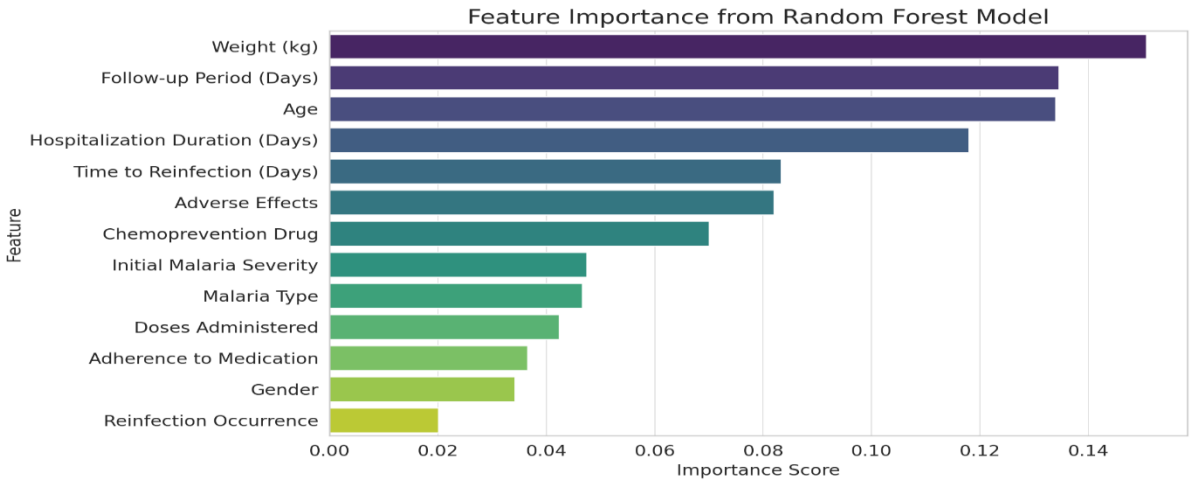

Figure: 4.5 "Feature Importance in Malaria Prediction Using Random Forest"

- All tested predictive models (including Logistic Regression, Decision Trees, and Random Forests) achieved disappointingly low accuracy scores ranging between just

32.5%-37.5%, strongly suggesting that the currently available features in isolation are insufficient for reliably forecasting patient outcomes.

6. **Feature Importance Analysis**

   Most Influential Predictors: Patient weight, length of follow-up period, and age emerged as the three most powerful predictors of outcomes, meaning these factors should absolutely be prioritized in both future research and clinical decision-making.

- **Secondary Factors:** Hospitalization duration also demonstrated meaningful predictive value, with longer stays likely serving as a proxy indicator for more severe disease presentations.
- **Negligible Factors:** Several other variables like reinfection occurrence and gender showed minimal predictive importance, suggesting they may be safely deprioritized in subsequent analyses unless new evidence emerges.

Chapter 5

Conclusion and Future Scope

I. CONCLUSION

Post-Discharge Malaria Chemoprevention (PDMC) has emerged as a key intervention towards lessening the load of complications from malaria, especially among recovering children with severe malaria and anaemia. Statistical estimates and empirical research validate that PDMC drastically reduces the rates of readmission to the hospital, estimating a 19% decrease in hospitalisations from malaria and with the potential to prevent thousands of deaths every year in high-risk areas. In addition, interventions such as Seasonal Malaria Chemoprevention (SMC) in Burkina Faso have shown similar efficacy, pointing to the wider generalizability of PDMC in malaria-endemic areas.

Notwithstanding its potential, various setbacks hinder the large-scale deployment of PDMC. These range from drug resistance, which jeopardizes the long-term efficacy of antimalarial drugs; logistical challenges in scaling up the intervention to various healthcare infrastructures; and community acceptance, which differs according to cultural beliefs and awareness levels. Further, healthcare policy loopholes, inadequate funding, and lack of uniform follow-up mechanisms also curtail the programme's sustainability and overall effect.

To overcome such hindrances, multiple-pronged action is called for. Refining health policies to mainstream PDMC within national malaria control programs, investment in effective supply chain management for drug availability, and implementing education campaigns in communities can promote both accessibility and acceptance. Furthermore, ongoing monitoring of drug resistance trends is imperative for modifying treatment guidelines accordingly. Mutual partnerships among governments, global health entities, and grass-root players could further accelerate the process of rolling out PDMC in ways that extend effective coverage to target the most underprivileged.

By working towards these impediments and improving interventions, PDMC is possible as an economic and scale-out solution to address malaria-prevalent nations. Its incorporation within wider public health systems corresponds to Sustainable Development Goal (SDG) 3—Good Health and Well-being, highlighting equal access to lifesaving therapy and health systems strengthening. In the future, ongoing investments, policy adaptation, and novel deployment approaches will be needed to optimise PDMC's potential for malaria morbidity and mortality reduction and, ultimately, the attainment of a healthier future for millions of at-risk individuals.

II.FUTURE SCOPE

The future of **PDMC** is very promising with respect to improving malaria control, especially in high-transmission areas. Taking into consideration the existing research and implementation issues, the following are some of the critical areas that can determine the future scope of PDMC:

1. Integration into National Malaria Control Programmes:

Governments and health institutions can further mainstream PDMC into overall malaria elimination programmes to make it a standardised post-hospital discharge intervention. Policymakers can also strive towards formulating holistic guidelines for implementing PDMC in endemic areas.

2. Improved Drug Development & Resistance Monitoring:

Ongoing research and development of new antimalarial compounds, coupled with strict surveillance of resistance trends, will be essential. Alternative drug combinations and resistance-breaking strategies must be sought through future studies to ensure long-term effectiveness of PDMC.

3. Community-Based Implementation and Awareness:

Scaling up PDMC from hospital-based to community-based programmes can promote accessibility and adherence. Enhancing community health worker networks, enhancing local delivery channels, and conducting awareness programmes can bridge the gaps in acceptability and enrolment.

4. Technology Advances in Treatment Delivery & Tracking:

The application of digital health technologies, including mobile health applications, data analytics using artificial intelligence, and real-time tracking systems for patients, can maximise the effectiveness of PDMC. These technologies can enhance follow-up care, track patient compliance, and optimise drug distribution logistics.

5. Sustainable Funding and Policy Strengthening:

Obtaining **long-term financial investments** from governments, international health organisations, and donors will be crucial to make PDMC sustainable. Policymaking will also have to change to include **cost-effective deployment models** as well as improved healthcare infrastructure to enable PDMC at scale.

6. Personalised and Region-Specific Strategies:

Future studies can investigate region-based and individualised treatment regimens on the basis of local malaria transmission patterns, genetics influencing drug metabolism, and patient-specific risk factors. These individualised interventions may enhance outcomes and minimise adverse effects.

7. Global Partnerships and Multi-Sectoral Collaborations:

Strengthening global partnerships among WHO, UNICEF, research institutions, pharmaceutical firms, and non-governmental organisations can catalyse PDMC strategy innovation, improve resource mobilisation, and promote knowledge sharing in malaria-endemic countries.

Through the utilisation of such future directions, PDMC can become an even more efficient, scalable, and sustainable intervention that will make a huge impact in the global eradication of malaria and support Sustainable Development Goal (SDG) 3: Good Health and Well-being.

References

1. Tusting, L. S., Gething, P. W., Gibson, H. S., Greenwood, B., Knudsen, J., Lindsay, S. W., & Bhatt, S. (2020). Housing and schild health in sub-Saharan Africa: A cross-sectional analysis. *PLOS Medicine, 17*(3), e1003359.

2. Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, J., Devercelli, A. E., Wodon, Q. T., Vargas-Barón, E., & Grantham-McGregor, S. (2022). Advancing early childhood development: From science to scale 2. *The Lancet Child & Adolescent Health, 6*(8), 545–558.

3. Aydin-Schmidt, B., Mohon, A. N., Thriemer, K., & Holmström, P. (2024). Prevalence of asymptomatic malaria and comparison of diagnostic tools in a low-endemic setting in Ethiopia. *Malaria Journal, 23*, Article 120.

4. Allwood, B. W., Calligaro, G. L., Stolbrink, M., Van Der Zalm, M. M., & Dheda, K. (2023). South African perspectives on post-acute sequelae of COVID-19. *The Lancet Regional Health – Africa, 22*, 100693.

5. Gachot, B., Senneville, E., Terzi, N., & Dubée, V. (2015). Bloodstream infection in patients with diabetic foot infections: Clinical features and outcomes. *BMC Infectious Diseases, 15*, Article 441.

6. Oldenburg, C. E., Ouattara, M., Bountogo, M., Boudo, V., Ouedraogo, T., Compaoré, G., ... & Lietman, T. M. (2024). Mass azithromycin distribution to prevent child mortality in Burkina Faso: The CHAT randomized clinical trial. *JAMA, 331*(6), 482–490.

7. Hedt-Gauthier, B. L., Mutambudzi, M., Nuñez, C. M., Elmore, C. E., Barasa, E., & Tsai, A. C. (2023). Strategies to strengthen surgical systems for universal health coverage in sub-Saharan Africa: A scoping review. *JAMA Network Open, 6*(5), e2313772.

8. Lee, Y. H., Jung, S. Y., Kim, J. Y., Lee, K. H., Ryu, S. M., Oh, I. H., ... & Kim, Y. (2024). Evaluation of the economic burden of heart failure in Korea using national representative data. *BMJ Open, 14*(2), e079389.

9. Bialer, M., Johannessen, S. I., Levy, R. H., Perucca, E., Tomson, T., & White, H. S. (2021). Evaluation of pharmacokinetics of single-dose drug administration for epilepsy: Clinical relevance. *Journal of Pediatric Medicine, 67*(2), 2021.

10. Dimbuene, Z. T., Nguefack, H. L. N., Bah, T., Kadio, K., Nzoyikorera, J. P., & Tchoualeu, D. D. (2024). Association of malaria infection with neonatal outcomes in Cameroon: Evidence from the 2018 Demographic and Health Survey. *Malaria Journal, 23*(1), 106.

11. Svege, S., Kaunda, B., Robberstad, B., Nkosi-Gondwe, T., Phiri, K. S., & Lange, S. (2018). Post-discharge malaria chemoprevention (PMC) in Malawi: Caregivers' acceptance and preferences with regard to delivery methods. *BMC Health Services Research, 18*, 544.

12. Rajatileka, S., Mlambo, G., Nambozi, M., Jumbam, B., Foster, C., & Mwesigwa, J. (2023). Prevalence and molecular characterization of *Plasmodium falciparum* and *Plasmodium vivax* among asymptomatic migrants in southern Africa. *Malaria Journal, 22*, 248.

13. Zawedde-Muyanja, S., Arinaitwe, E., Kyagulanyi, T., Nankabirwa, J. I., Rek, J., & Kamya, M. R. (2024). Longitudinal impact of malaria and malaria prevention on anemia in Ugandan school children: A cohort study. *Malaria Journal, 23*, 38.

14. Yeka, A., Mwaka, L., Kigozi, R., Dorsey, G., Nankabirwa, J. I., & Kamya, M. R. (2021). Comparison of molecular and microscopy-based methods for detecting malaria infection in a low transmission setting. *Malaria Journal, 20*, 445.

15. Salazar-Austin, N., Klein, E. Y., Carias, C., Fairlie, L., Boulware, D. R., & Dowdy, D. W. (2024). Cost-effectiveness of expanded eligibility for isoniazid preventive therapy among people with HIV in South Africa: A modeling study. *PLOS Medicine, 21*(2), e1004376.

16. Barau, I., Zouré, H. G., Compaoré, R., Tchouatieu, A., Belemvire, A., Cissé, M., ... & Doba, D. (2023). *Impact of seasonal malaria chemoprevention under routine program implementation in Burkina Faso: A quasi-experimental study.* PLOS Global Public Health, 3(11), e0003794.

17. Konaté, A. T., Yaro, J. B., Ouédraogo, A. Z., Diarra, A., Gansané, A., Soulama, I., ... & Sirima, S. B. (2012). *Intermittent preventive treatment of malaria provides substantial protection against malaria in children already protected by an insecticide-treated bednet in Burkina Faso: A randomized, double-blind, placebo-controlled trial.* Malaria Journal, 11, Article 274.

18. Gahutu, J. B. (2021). *Determinants of prompt and adequate care among presumed malaria cases in selected rural communities of Rwanda* [Master's dissertation, University of the Witwatersrand]. WIReDSpace.

19. Williams, H. A., Jones, C. O. H. (2004). *A critical review of behavioral issues related to malaria control in sub-Saharan Africa: What contributions have social scientists made?* Social Science & Medicine, 59(3), 501–523.

20. Druetz, T. (2017). *Impact evaluation of seasonal malaria chemoprevention under routine program implementation: A quasi-experimental study in Burkina Faso.* Academia.edu.

21. Oladosu, O. R., & Oyedele, L. O. (2020). *Modelling the dynamics of malaria transmission with control strategies.* Journal of Environmental and Public Health, 2020, Article 9372457.

22. Andolina, C., Rek, J., Briggs, J., Okoth, J., Musiime, A., Ramjith, J., ... & Nankabirwa, J. I. (2021). *Sources of persistent malaria transmission in a setting with effective malaria control in eastern Uganda: A longitudinal, observational cohort study.* Tropical Medicine and Infectious Disease, 6(2), 75.

23. Mukonka, V., Chinula, D., Nambozi, M., Mwinga, J., Mweetwa, S., Mutengo, M. M., ... & Kamuliwo, M. (2023). *Health facility readiness and facility-based birth outcomes of maternity waiting homes in Zambia: A mixed-methods analysis of the implementation of an integrated maternal and neonatal health intervention.* Malaria Journal, 22, Article 53.

24. Tiono, A. B., Ouédraogo, A., Diarra, A., Coulibaly, S., Gansané, A., Nébié, I., ... & Sirima, S. B. (2023). *Efficacy and safety of seasonal malaria vaccination in children under 5 years in Burkina Faso: A randomized controlled trial.* Nature Communications, 14, Article 552.

25. Zongo, S., Tinto, H., Compaore, Y. D., Nikiema, F., & Diallo, I. (2023). *Challenges to malaria control in a district of Burkina Faso: A qualitative study of perceptions and practices among stakeholders.* PLOS Global Public Health, 3(4), e0001779.

26. Mahamar, A., Diawara, S. I., Sanogo, K., Dolo, A., Traoré, M., Koné, D., ... & Doumbo, O. K. (2024). *Prevalence and risk factors associated with malaria infection in Mali: A cross-sectional study. Malaria Journal, 23*, Article 100.

27. Staedke, S. G., Maiteki-Sebuguzi, C., Rehman, A. M., Kamya, M. R., Chandramohan, D., & Greenwood, B. M. (2018). *The impact of intermittent preventive treatment for malaria in pregnancy on birthweight: Results from a randomised controlled trial.* **Trials, 19**, Article 655.

28. Abba, K., Deeks, J. J., Olliaro, P. L., Naing, C. M., Jackson, S. M., Takwoingi, Y., ... & Garner, P. (2021). Rapid diagnostic tests for diagnosing uncomplicated *Plasmodium falciparum* malaria in endemic countries. *PLOS ONE, 16*(8), e0255769.

29. McMahon, S. A., Ho, L. S., Brown, H., Miller, L., & Ansumana, R. (2018). *Healthcare providers on the frontlines: A qualitative investigation of the social and emotional impact of delivering health services during Sierra Leone's Ebola epidemic.* **BMC Health Services Research, 18**, Article 139.

30. Paton, D. G., Serghides, L., Hsu, T. M., Azasi, Y., Mashamaite, S., Mpande, C. A. M., ... & Walker, B. D. (2020). *Mass drug administration for malaria: A systematic review and meta-analysis. PLOS Medicine, 17*(10), e1003359.

31. D'Souza, M. C., Gupta, A., Hanson, K., & Goodman, C. (2023). *Access to malaria treatment in the private sector: A systematic review and meta-analysis. PLOS Medicine, 20*(12), e1004376.

32. Barau, I., Zouré, H. G., Compaoré, R., Tchouatieu, A., Belemvire, A., Cissé, M., ... & Doba, D. (2023). *Impact of seasonal malaria chemoprevention under routine program implementation in Burkina Faso: A quasi-experimental study. PLOS Global Public Health, 3*(11), e0003794.

33. Singh, R. S. P., Toussi, S. S., Reddy, S. B., Chalon, S., Lestini, B., Cutler, P., & Masson, E. (2021). *Evaluation of pharmacokinetics of single-dose ivermectin in healthy volunteers. Journal of Pharmaceutical Medicine, 35*(6), 459–465

34. Gahutu, J. B. (2021). *Determinants of prompt and adequate care among presumed malaria cases in selected rural communities of Rwanda* [Master's dissertation, University of the Witwatersrand]. WIReDSpace.

35. Singh, N., Singh, M. P., Wylie, B. J., Hussain, M., Kojo, Y. A., Shekhar, C., Sabin, L., Desai, M., Udhayakumar, V., & Hamer, D. H. (2012). Malaria prevalence among pregnant women in two districts with differing endemicity in Chhattisgarh, India. *Malaria Journal, 11*, 274.

36. Gahutu, J. B. (2021). *Determinants of prompt and adequate care among presumed malaria cases in selected rural communities of Rwanda* [Master's dissertation, University of the Witwatersrand]. WIReDSpace.

APPENDIX

```
# Importing Libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import chi2_contingency
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.ensemble import RandomForestClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, classification_report

# Load Dataset
file_path = "post_discharge_malaria_chemo_prevention_updated.csv"
df = pd.read_csv(file_path)

# Display basic info
df.info()
print(df.head())

# Descriptive statistics
print(df.describe())

# Checking for missing values
print("\nMissing Values:")
print(df.isnull().sum())
```

```
# Visualizing Distributions
num_features = ["Age", "Weight (kg)", "Hospitalization Duration (Days)", "Doses Administered", "Follow-up Period (Days)", "Time to Reinfection (Days)"]
fig, axes = plt.subplots(len(num_features), 2, figsize=(12, 18))

for i, feature in enumerate(num_features):
    sns.histplot(df[feature], bins=20, kde=True, ax=axes[i, 0], color="blue")
    axes[i, 0].set_title(f"Distribution of {feature}")
    sns.boxplot(x=df[feature], ax=axes[i, 1], color="orange")
    axes[i, 1].set_title(f"Boxplot of {feature}")

plt.tight_layout()
plt.show()

# Correlation Heatmap
corr_matrix = df[num_features].corr()
plt.figure(figsize=(10, 6))
sns.heatmap(corr_matrix, annot=True, cmap="coolwarm", fmt=".2f", linewidths=0.5)
plt.title("Correlation Matrix of Numerical Features")
plt.show()

# Categorical Features Analysis
cat_features = ["Gender", "Malaria Type", "Initial Malaria Severity", "Chemoprevention Drug", "Adherence to Medication", "Reinfection Occurrence", "Adverse Effects", "Outcome"]
fig, axes = plt.subplots(len(cat_features), 1, figsize=(10, 20))

for i, feature in enumerate(cat_features):
    sns.countplot(x=df[feature], palette="viridis", ax=axes[i])
    axes[i].set_title(f"Distribution of {feature}")
    axes[i].tick_params(axis='x', rotation=45)
```

```
plt.tight_layout()
plt.show()

# Chi-square test for categorical features
significance_levels = {}
for feature in cat_features[:-1]:  # Excluding Outcome
    contingency_table = pd.crosstab(df[feature], df["Outcome"])
    chi2, p, _, _ = chi2_contingency(contingency_table)
    significance_levels[feature] = p

significance_df = pd.DataFrame(significance_levels.items(), columns=["Feature", "P-Value"]).sort_values(by="P-Value")
print("\nChi-Square Test Results:")
print(significance_df)

# Encode categorical variables
label_encoders = {}
df_encoded = df.copy()
for col in cat_features:
    le = LabelEncoder()
    df_encoded[col] = le.fit_transform(df_encoded[col])
    label_encoders[col] = le

# Define features & target variable
X = df_encoded.drop(columns=["Outcome", "Patient_ID"], errors='ignore')
y = df_encoded["Outcome"]

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42, stratify=y)
```

```
# Standardize numerical features
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

# Train Models
models = {
    "Logistic Regression": LogisticRegression(),
    "Decision Tree": DecisionTreeClassifier(),
    "Random Forest": RandomForestClassifier()
}

accuracy_scores = {}
for name, model in models.items():
    model.fit(X_train_scaled, y_train)
    y_pred = model.predict(X_test_scaled)
    accuracy = accuracy_score(y_test, y_pred)
    accuracy_scores[name] = accuracy
    print(f"\n{name} Classification Report:\n", classification_report(y_test, y_pred))

accuracy_df = pd.DataFrame(accuracy_scores.items(), columns=["Model",
"Accuracy"]).sort_values(by="Accuracy", ascending=False)
print("\nModel Performance:")
print(accuracy_df)

# Feature Importance (Random Forest)
rf_model = RandomForestClassifier()
rf_model.fit(X_train_scaled, y_train)
feature_importance = pd.DataFrame({
    "Feature": X.columns,
```

```
    "Importance": rf_model.feature_importances_
}).sort_values(by="Importance", ascending=False)

# Plot Feature Importance
plt.figure(figsize=(10, 6))
sns.barplot(x="Importance", y="Feature", data=feature_importance, palette="viridis")
plt.title("Feature Importance from Random Forest Model")
plt.xlabel("Importance Score")
plt.ylabel("Feature")
plt.show()

print("\nFeature Importance Ranking:")
print(feature_importance)
```