

Bayesian Networks for Diagnosing Childhood Malaria in Malawi

by

Sanya Bathla Taneja

B.Tech., Indira Gandhi Delhi Technical University for Women, 2018

Submitted to the Graduate Faculty of the
School of Computing and Information in partial fulfillment
of the requirements for the degree of
Master of Science in Intelligent Systems

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

SCHOOL OF COMPUTING AND INFORMATION

This thesis was presented

by

Sanya Bathla Taneja

It was defended on

April 10, 2020

and approved by

Gregory Cooper, Professor, Department of Biomedical Informatics and Intelligent Systems
Program

Gerald Douglas, Assistant Professor, Department of Biomedical Informatics

Thesis Advisor: Shyam Visweswaran, Associate Professor, Department of Biomedical
Informatics and Intelligent Systems Program

Copyright © by Sanya Bathla Taneja

2020

Bayesian Networks for Diagnosing Childhood Malaria in Malawi

Sanya Bathla Taneja, B.Tech.

University of Pittsburgh, 2020

Infectious diseases such as malaria are responsible for the majority of under-five deaths in low- and middle-income countries. Accurate diagnosis and management of illnesses can help in reducing the global burden of childhood morbidity and mortality. While trained healthcare workers deliver treatment for common childhood illnesses in healthcare facilities in Malawi, there is a significant lack of diagnostic support in rural health centers. With recent trends in artificial intelligence in global health, we hypothesize that a data-driven approach to diagnosis of childhood illnesses may address the challenges faced in health centers in low-resource countries. In this study, we aim to utilize Bayesian networks to diagnose cases of childhood malaria in Malawi. We develop two Bayesian network (BN) models for diagnosis of malaria using clinical signs and symptoms. The first model is created manually, while the other combines an Augmented Naïve Bayes approach with model editing by an expert. The models are learnt using a national survey dataset which contains sick child observations including patient information, diagnosis, and symptoms. The target malaria diagnosis is taken as the result of the malaria rapid diagnostic test (mRDT). The performance of the BN models is further compared to traditional machine learning classifiers on the basis of accuracy, area under the receiver operating characteristic curve (AUC), F1 score, sensitivity and specificity. We also present an experimental framework that can be used to model the malaria diagnostic support in the rural health centers. The manually created BN model achieves accuracy of 63.6% and AUC of 0.583. The Augmented Naïve Bayes model considers associations between the variables and achieves an accuracy of 62.7% and AUC of 0.581. BN models provide a powerful, efficient and data-driven tool for diagnosis of childhood illness that can lead to a more evidence-based clinical practice in Malawi. The simplicity and interpretability of BN models offers a unique approach to diagnostic support in low-resource countries. As BN models are representative of the population from which the data has been derived, this approach can be generalized to other childhood illnesses in different regions of the world.

Table of Contents

Preface.....	viii
1.0 Introduction.....	1
2.0 Methods.....	4
2.1 Data Description	4
2.2 Bayesian Network Models	7
2.2.1 Data Preprocessing	7
2.2.2 Model Development and Evaluation	8
2.3 Comparison to Machine Learning Classifiers	9
2.3.1 Data Preprocessing	9
2.3.2 Classification.....	9
2.4 Translation to Clinical Practice	10
3.0 Results	13
3.1 Data Description	13
3.2 Bayesian Network Models	15
3.3 Comparison to Machine Learning Classifiers	18
3.4 Translation to Clinical Practice	19
4.0 Discussion.....	21
Bibliography	25

List of Tables

Table 1: Variables included in the models, their categories and values	6
Table 2: Data characteristics in the SPA dataset (where mRDT result is available)	14
Table 3: Contingency table for classification of malaria on the test data using the manual BN model	17
Table 4: Contingency table for classification of malaria on the test dataset using the hybrid BN model.....	18
Table 5: Performance metrics for all models on the test dataset	19

List of Figures

Figure 1: Experimental decision model to use BN predictions for decision support.....	10
Figure 2: Manually created BN in GeNIe Modeler	16
Figure 3: Hybrid BN using ANB structure learning with expert knowledge.....	16
Figure 4: BN with subset of arcs with high strength of influence values.....	17
Figure 5: Figure 5: ROC curves for classification of malaria in test dataset using (a) manual BN model and (b) hybrid BN model.....	18
Figure 6: Sensitivity analysis of the decision model in Figure 1.	20

Preface

We acknowledge the significant contributions of Dr. Marian Michaels, UPMC Children's Hospital of Pittsburgh and Professor of Pediatrics at the University of Pittsburgh School of Medicine, in this study. Dr. Michaels has been instrumental in developing the Bayesian models through her knowledge and expertise in infectious diseases in children. Dr. Marek Druzdzel provided key inputs for the development of the Bayesian models in GeNIe. We further acknowledge the contributions of Rashid Deula and his help during the field visits in Malawi to health centers and health posts.

The models described in this thesis were created using the GeNIe Modeler, available free of charge for academic research and teaching use from BayesFusion, LLC, <http://www.bayesfusion.com/>. The dataset is publicly available through the DHS website, <https://www.dhsprogram.com/>.

1.0 Introduction

Infectious diseases such as malaria are responsible for the majority of under-five deaths in low- and middle-income countries (LMIC). Accurate diagnosis and management of illnesses can help reduce the global burden of childhood morbidity and mortality. As per the Malaria Indicator Survey in 2017, the prevalence of malaria in children under five in Malawi is 24%, with this percentage being higher in rural areas (48%). The *P. falciparum* parasite accounts for 98% of malaria infections and all severe cases and malaria deaths (1). In Malawi, trained community-based healthcare workers known as Health Surveillance Assistants (HSAs) deliver treatment for common childhood illnesses including malaria at under-five clinics or health posts. Rural health centers, staffed with a medical assistant or clinical officer, offer alternate treatment locations for children in rural areas. For the majority of the population, these sites serve as the primary points of care (1). This study aims to utilize Bayesian network (BN) models and machine learning classifiers to develop a data driven approach to diagnosis of childhood malaria that can be implemented in the health centers and health posts.

The current standard for the management of childhood malaria is defined in a set of clinical guidelines developed by the World Health Organization (WHO) (2). Based on these guidelines, a child presenting with suspected malaria at a healthcare site should have a confirmatory diagnosis based on microscopy or the malaria rapid diagnostic test (mRDT). Historically, presumptive treatment of febrile illness with anti-malarial drugs has been a common approach. Following the national policy in 2010, use of mRDTs for suspected malaria is standard practice before administration of artemisinin combination therapy (ACT) drugs. The scale up of ACT coverage, distribution of mRDTs, and significant efforts to provide community-based care for childhood malaria have led to a decline in the disease burden (3). However, there are still a number of challenges that hinder effective management of malaria in health facilities in the rural communities.

Rural health posts are often characterized by limited availability of resources, unavailability of diagnostic testing facilities, and lack of expert clinicians (4). In a study conducted in 2017 in Malawi, Klootwijk et. al. reported lack of microscopy facilities in all the health centers surveyed (5). mRDTs and HIV tests are often the only diagnostic tests available at the health care

facilities (6). Even so, stock outs of mRDTs and first line ACT drugs are common, especially during the malaria season and in rural areas. The Malawi Service Provision Assessment (SPA) survey reported that mRDTs are only available in 85% of the facilities, with hospitals having the highest proportion available (95%) and health posts the lowest (19%) (6). Common reasons for stock outs include late and inaccurate reporting of supplies, drug pilferage and overprescribing of anti-malarial and antibiotic drugs (4,7). As HSAs are encouraged to focus on adherence to the guidelines, the unavailability of mRDTs leads to three common scenarios at the health posts. The child may be referred to a secondary health center or tertiary hospital, the HSA treats every child presenting with fever presumptively (if ACT drugs are in stock), or in the worst case, the health post remains closed while mRDTs are out of stock. More often than not, the guardians of the child cannot arrange transportation to the referral or alternate site and the child remains untreated (5). While diagnostic tests and essential drugs are provided free of cost to patients at all public health facilities in Malawi, data on availability and affordability of drugs shows that a single course of treatment is unaffordable for a major part of the population (8). This can be a problem if the guardians are advised to purchase drugs unavailable at the healthcare facilities. Given the high-volume patient settings and increasing evidence of non-adherence to paper-based traditional case management guidelines (9), it is imperative to provide diagnostic and treatment support to the healthcare workers.

Technological advances may help in tackling some of the challenges above. The promise of artificial intelligence (AI) and data-centric models for healthcare in LMIC has recently begun to see light (10). While decision support systems are common in many high-income countries, migration of the same technologies to LMIC is impossible due to the unique challenges in resource-constrained countries. The distinct needs, diseases, demographics, and standards of care call for a different approach for personalized medicine in LMIC by adopting tools specifically designed for use in these areas (11). Prior attempts to develop clinical decision support in Malawi have focused on implementing electronic versions of existing guidelines, rather than personalized evidence-based algorithms and data (12,13). There is a significant lack of diagnostic support for the healthcare workers in these applications. We propose to implement a diagnostic model for childhood malaria in Malawi to aid in treatment in cases where mRDT is unavailable or is in limited supply, with a goal toward sustainable use of both mRDTs and ACT drugs.

A recent review of electronic clinical decision algorithms (eCDA) in LMICs identifies the lack of effective, integrated diagnostic tools as a contributing factor to childhood morbidity and mortality (13). In addition to better diagnosis of diseases and support for rational use of drugs, the review identifies components of an eCDA that will be crucial in closing gaps in the primary care management systems in low-resource countries. These include algorithms for specific regions, openly available evidence-base and content, automated data collection for monitoring and evaluation, and syndromic-based surveillance systems (13). BN models can be aptly used for prediction and diagnosis of diseases using country or region-specific data. BNs are probabilistic graphical models that can use data to learn associations between variables such as a disease diagnosis and symptoms. Each variable or node has a conditional probability table that contains the parameters of the network, and this network is representative of the data and, as such, of the population from which the parameters have been obtained. BN models can be used for classification of diseases where they compute the posterior probability of the target (disease) node given the observed signs and symptoms (14).

BNs have been developed for efficient diagnosis and risk prediction of many diseases (15–17). In this study, we developed BN models for diagnosis of malaria in children under 5 years of age in Malawi. We evaluated two variations of BNs for this task, both utilizing a publicly available dataset from Malawi. We further compared the results of the BNs models to other machine learning classifiers. Finally, we demonstrated the integration of BN models with a decision-analytic framework that can be implemented at a typical rural health post in Malawi.

2.0 Methods

We first give a description of the Malawi Service Provision Assessment (SPA) (6) data used in this study to develop and evaluate the models. We developed two BN models, one created manually with parameters learnt using the dataset, and the second with an automated learning approach known as Augmented Naïve Bayes to create the structure of the model and learn the parameters. The performance of these models was then compared to three machine learning methods for binary classification.

2.1 Data Description

We used the SPA data in this study to train and evaluate the models. The SPA survey was conducted between July 2013 and February 2014 by the Ministry of Health, with support from the Demographic and Health Survey (DHS) program, to assess the status of health facilities and quality of healthcare in Malawi. Data were collected from 977 facilities including 119 hospitals, 489 health centers, 55 dispensaries, 369 clinics, and 28 health posts in the three major regions in the country, and are representative at the national level by facility type and managing authority (6). These data have been used previously in studies to assess the quality of care and treatment for pneumonia in Malawi (18). The data file is publicly available through the DHS program (19).

The survey data contains observations of 3,441 children aged 2 to 59 months presenting with an illness complaint in outpatient facilities. This includes observations about the examination, guardian exit interview, and an expert healthcare provider's re-examination recorded by an observer during the outpatient procedure for each child. The dataset contains information about the sick child, including demographic details (age, date of birth, gender), symptoms of disease (fever, diarrhea, anemia, etc.), test results (if any), and the provider's diagnosis.

We took the result of the mRDT for the children in the dataset as the gold standard malaria diagnosis for developing the models in this study. The test is known to have high sensitivity (0.997) and specificity (0.995) for diagnosis of malaria (20) and is the standard practice to obtain

parasitological confirmation of the disease as per both the WHO and national malaria treatment guidelines (21). Thus, it is an objective and appropriate indicator of the presence of the disease. To this effect, if the mRDT result for a child was positive, we considered malaria present, and absent if the test result was negative. The variable is also referred to as ‘malaria’ or ‘malaria diagnosis’ in the following sections.

While it would be ideal to have the mRDT result for each child in the dataset, this is not the case. Only 1,139 children in the dataset had a reported mRDT result. Other cases either included a microscopy-based diagnosis or the provider’s diagnosis without any further details. We, thus, excluded all cases from our analysis without the mRDT result and used the observations from this subset to create and evaluate the models. Table 1 shows the variables from the dataset that were used in the models, along with their categories and values. These variables were chosen due to their presence in the childhood illness management guidelines (2) as well as expert domain knowledge.

Table 1: Variables included in the models, their categories and values

Category	Variable	Values
Diagnosis	Malaria	Positive Negative
Patient Information	Age	Less than 2 months 2-12 months 13-24 months 25-60 months Over 60 months Unknown
	Duration of Illness	Less than or equal to 2 days 3-15 days 16-30 days Over 30 days Unknown
Sign or Symptom	Conscious	Yes No Unknown
	Anemia	Present Absent Unknown
	Convulsions	
	Cough or Difficulty Breathing (CDB)	
	Diarrhea	
	History of Fever	
	Fever (temperature > 37.5 C)	
	Lethargic	
	Malnutrition	
	Unable to feed	
	Vomiting	

The dataset with all variables in Table 1 was split into 80% training set and 20% test set, stratified on the target variable i.e. the malaria diagnosis. The models were developed using the training set and we evaluated the performance of each model using the test dataset.

2.2 Bayesian Network Models

A BN is a probabilistic graphical model that consists of nodes representing variables and arcs depicting associations between the variables. Each node in the network has a conditional probability table, also referred to as the parameters of the node. BNs can function as classifiers where the classifier gives the posterior probability distribution of the target node (such as a disease diagnosis) given the values of all other nodes in the network (22). When learning a BN from data, the nodes represent the dataset variables. We can learn the parameters or both the structure and the parameters of the network using the dataset. This study used the Augmented Naïve Bayes method to learn the structure of the BN for malaria diagnosis and compared it to a manually constructed BN model. For each node X in the model, the parameters were computed by estimating the conditional probability of the node given its parents $pa(X)$ from the dataset (14).

2.2.1 Data Preprocessing

We developed the models using the variables in Table 1. Continuous variables such as age and duration of illness were discretized. Age was divided by months (2-12, 13-24, 25-60). The intervals are based on the varying epidemiology of the disease in children of different ages. The duration of illness was discretized by number of days as shown in Table 1. All the signs and symptoms in the model can be ‘Present’, ‘Absent’ or ‘Unknown’. All variables, except the malaria diagnosis, can have an ‘Unknown’ value due to the missing observations of variables in the dataset. As mentioned previously, the malaria variable contains the mRDT result for each child.

2.2.2 Model Development and Evaluation

We used the GeNIe modeler to create the two BNs. The variables in Table 1 formed the nodes of the network and both networks consisted of the malaria variable as the target node. An arc drawn from malaria to convulsions, with malaria as the parent node and convulsions as a child node, implies that malaria is associated with convulsions.

For the first model, we created the nodes and linked the nodes in the network using domain knowledge of malaria from experts and literature. In this variation, all signs or symptoms were assumed to be independent of each other. We used the training data to learn the parameters of the network by populating the conditional probability tables of the nodes using conditional frequencies from the dataset. In the dataset, the marginal probability of malaria being “present” is 0.36. This corresponds to the dataset wherein 36% children have a positive malaria diagnosis.

While the model created manually is simple and interpretable, the assumption of independence between the nodes often does not hold in the real world. It is important to model the relationships between the various symptoms to get an accurate diagnosis. Thus, for the second BN model, we developed a hybrid network by first learning an Augmented Naïve Bayes structure and then improving it with expert knowledge. As the algorithm may learn associations that are biased due to the dataset used, each arc in the Augmented Naïve Bayes structure is evaluated by an expert to assess its relevance and fit in the structure clinically.

An Augmented Naïve Bayes (ANB) model (14) is an extension of the Naïve Bayes model, where the target variable forms the root of the network and all other nodes have an incoming arc from the root node. However, unlike the Naïve Bayes model, an ANB model allows associations between the nodes (other than the root node) that are learnt from the dataset and represented with arcs between them. In this structure, an arc from diarrhea to convulsions implies that the probability of convulsions depends both on the state of malaria and of diarrhea in a patient. This allows the model to account for the statistical associations between the variables present in the dataset. Learning an ANB model is akin to finding the best BN structure with the target variable as the root. Most studies involving ABNs focus on a special variant of the model known as Tree-Augmented Naïve Bayes (TANB) (14) due to the efficiency with which it can be learned. However, the GeNIe modeler allows efficient learning of general associations between variables in an ANB model, and thus, we use it.

Once the BN models were created, we used them to classify the malaria diagnosis on the test dataset. We computed the accuracy, area under the Receiver Operating Characteristic curve (AUC), precision, F1 score, sensitivity, and specificity for each model. The AUC value indicates the diagnostic discrimination performance of the model, where perfect performance has an AUC of 1.

2.3 Comparison to Machine Learning Classifiers

We trained binary machine learning classifiers with the SPA dataset and compared their performance to that of the BN models. We used the same train and test split of the data after preprocessing.

2.3.1 Data Preprocessing

The variables described in Table 1 were used for training and evaluating the machine learning classifiers. Instead of discretizing the continuous variables, we used a standard scaler to scale the ‘age’ and ‘duration of illness’ to unit variance. We imputed the mean for both age and duration of illness for all observations when these values were missing. All the categorical variables were used in the same way as in the BN models, i.e. with 3 possible values – present, absent and unknown.

2.3.2 Classification

We trained three different binary classifiers, namely, logistic regression, random forest and Bernoulli Naïve Bayes to predict the malaria diagnosis. The Naïve Bayes model relies on the assumption of independence between all the variables (as in the first BN model above) except the target variable, while logistic regression and random forest do not work under this assumption. We trained the models using the scikit-learn library (23) in Python. We used the training data to obtain the best combination of parameters for each classifier over five folds of the training data using the

grid search method. We evaluated the models using the test dataset and derived the accuracy, AUC, precision, F1 score, sensitivity and specificity on a threshold equal to 0.5.

2.4 Translation to Clinical Practice

The BNs give the posterior probability of malaria after the nodes of the network have been observed. We present an experimental decision model in Figure 1 to use the predictions from the BN models for clinical decision support. A healthcare worker can use the decision model to decide whether or not to use the mRDT as well as prescribe ACT drugs to a sick child presenting with symptoms at a healthcare facility.

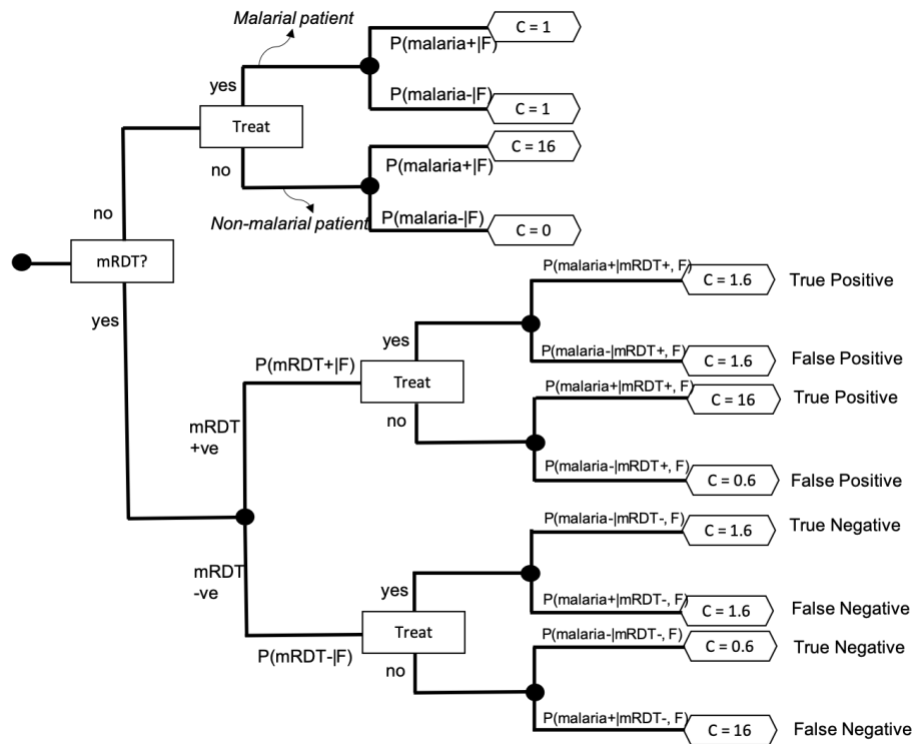


Figure 1: Experimental decision model to use BN predictions for decision support.

The decision is driven by the expected cost values of the ‘mRDT?’ and ‘Treat’ nodes in Figure 1. This is a common approach to create decision models and influence diagrams (24). The expected cost value of the [mRDT=no] branch was calculated using the probability of malaria from the Bayesian network and costs associated with each decision as-

$$\begin{aligned} \text{Expected cost of [mRDT?=no]} = \\ \min(1.0 * P(\text{malaria+|F}) + 1.0 * P(\text{malaria-|F}), 16 * P(\text{malaria+|F}) + 0 * P(\text{malaria-|F})) \end{aligned}$$

In Figure 1, P(malaria+|F) is the probability from the BN model that the malaria diagnosis is positive given the findings of the patient. P(malaria-|F) is the probability from the Bayesian network that the malaria diagnosis is negative given the findings. The cost values (in hexagons) in the decision model reflect the perspective of the payer in healthcare system. These costs depend on the resources used, namely, ACT for treatment and mRDT for diagnosis from the payer’s point of view. The cost of an mRDT kit in the model is US \$0.60 (8) and drug cost per child dose for first line treatment of uncomplicated malaria is US \$1.00 (25). There is a high cost (US \$16.00) in the model associated with the decision to not treat a child with malaria, with the assumption that the cost may increase to 10 times the original cost of treatment and diagnosis for cases with unresolved malaria as the disease becomes severe. While treatment and tests are provided free of cost to patients at public health facilities in Malawi, sustainable resource utilization is driven by minimizing the costs of these resources from a payer’s perspective.

The expected cost of the [mRDT=yes] branch was computed as -

$$\begin{aligned} \text{Expected cost of [mRDT?=yes]} = \\ P(\text{mRDT+|F}) * (\text{Expected Cost of [Treat] when mRDT+}) + P(\text{mRDT-|F}) * (\text{Expected Cost of [Treat] when mRDT-}) \end{aligned}$$

where

$$\begin{aligned} \text{Expected cost of [Treat] when mRDT+} = \\ \min(1.6 * P(\text{malaria+|mRDT+, F}) + 1.6 * P(\text{malaria-|mRDT+, F}), 16 * P(\text{malaria+|mRDT+, F}) + 0.6 * P(\text{malaria-|mRDT+, F})) \end{aligned}$$

$$\begin{aligned} \text{Expected cost of [Treat] when mRDT-} = \\ \min(1.6 * P(\text{malaria-|mRDT-, F}) + 1.6 * P(\text{malaria+|mRDT-, F}), 0.6 * P(\text{malaria-|mRDT-, F}) + 0.6 * P(\text{malaria+|mRDT-, F})) \end{aligned}$$

In the above equations, $P(\text{malaria+}|\text{mRDT+}, F)$ is the probability that a child has malaria given that the mRDT result is positive and the findings of the patient, and $P(\text{malaria-}|\text{mRDT-}, F)$ is the probability that a child does not have malaria given that the mRDT result is negative and the findings of the patient. $P(\text{mRDT+}|F)$ and $P(\text{mRDT-}|F)$ represent the probabilities of the mRDT being positive or negative respectively, given the findings of the patient. The probabilities in the above equations were calculated with Bayes' theorem using sensitivity and specificity of the mRDT (0.997 and 0.995 respectively). The model in Figure 1 only utilized the cost of the test and drugs. These costs can be increased when the mRDT is in limited or no supply. We can then calculate the expected cost values of each branch and take a decision for testing and treatment based on which branch has a lower expected cost.

While the decision model presented above are not complete or perfect, this process allowed us to perform a sensitivity analysis on the posterior probabilities computed by the BN models. To perform sensitivity analysis, we calculated the expected costs of the branches in Figure 1 by varying the value of the posterior probability of the malaria diagnosis from 0 to 1. We were thus able to find the thresholds for this decision model based on the test values.

3.0 Results

3.1 Data Description

The SPA dataset contains 3,441 observations. After excluding cases where the result of the mRDT was unknown, we used 1,139 observations with 14 variables (Table 1) to develop the models. The data characteristics for the variables are included in Table 2. Of note, 36.4% (415) of the total cases had a positive malaria diagnosis. The majority of the children were aged between 24 to 60 months (35.7%). The duration of illness varied from 0 days to over 30 days, although the duration period of 0 to 2 days was the most common (51.7%). A large proportion of children presented with history of fever (69.3%) and CDB (62.8%). These were the most common complaints, followed by vomiting (29.1%) and diarrhea (26.3%).

We found that every field in Table 2, except anemia and malnutrition, sometimes has missing values. These were coded as ‘Unknown’ in the BN models. The proportion of missing values ranged from 0 to 6.4%, with the highest missing values in ‘history of fever’ and the lowest missing values in the anemia and malnutrition variables.

Table 2: Data characteristics in the SPA dataset (where mRDT result is available)

Variable	Number of children with mRDT result n = 1,139	% (n = 1,139)
Diagnosis – Malaria		
mRDT positive	415	36.4
mRDT negative	724	63.6
Patient Information		
Age (in months)		
2-12	379	33.3
12-24	294	25.8
24-60	407	35.7
Other	21	1.8
Duration of Illness (in days)		
<=2	589	51.7
3 to 15	502	44.1
15 to 30	11	1.0
More than 30	37	3.2
Sign or Symptom (Present)		
History of Fever	789	69.3
Fever (temperature)	307	27.0
Diarrhea	299	26.3
Vomiting	331	29.1
Malnutrition	11	1.0
Convulsions	55	4.8
Unable to feed	137	12.0
Lethargic	228	20.0
Unconscious	12	1.1
CDB	715	62.8
Anemia	93	8.2

3.2 Bayesian Network Models

Figure 2 shows the BN created manually with all the variables as nodes and the links between them. The variables duration of illness and age are parents of the malaria node, while the signs and symptoms are children of the malaria node. This is modeled so that the arcs represent the influence of patient characteristics (age in months, duration of illness) on the disease, and these are propagated through the network to influence the observed signs and symptoms. There is an arc from the malaria diagnosis to each sign or symptom in the model. This is a common approach to develop expert-elicited BN models (26). The model created in Figure 1 is simple and interpretable, and assumes conditional independence among signs and symptoms.

Figure 3 shows the hybrid BN created using the ANB learning combined with expert knowledge. As this model learnt associations between the variables, there were several arcs between the nodes in addition to an arc from the malaria diagnosis to each of the other nodes. The model contained 42 arcs with the 14 variables as nodes. To illustrate the influences in the model, we present the model with only a subset of the arcs from the complete model in Figure 4. Figure 4 shows the hybrid model with all nodes and arcs from the malaria node to all other nodes. In addition to these, we included the arcs with high ‘strength of influence’ values. For each arc in the complete model, the strength of influence measures the Euclidean distance between the conditional probability distributions of the nodes linked by that arc. This was calculated over each node conditional on the states of the parent node and implies a higher association between two nodes when the strength of influence is large (27). The red arcs in the model in Figure 4 had high strength of influence values calculated from the network parameters. These included association of consciousness of a child with lethargy, anemia, convulsions, diarrhea, and duration of illness, fever with cough or difficulty breathing (CDB), and the age of child with history of fever.

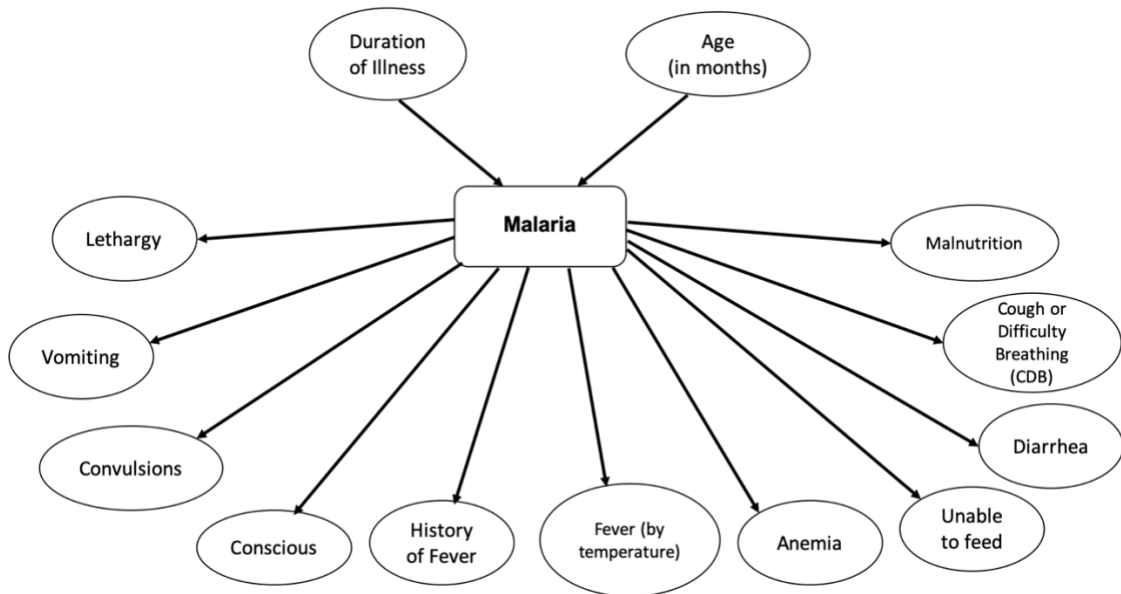


Figure 2: Manually created BN in GeNIe Modeler

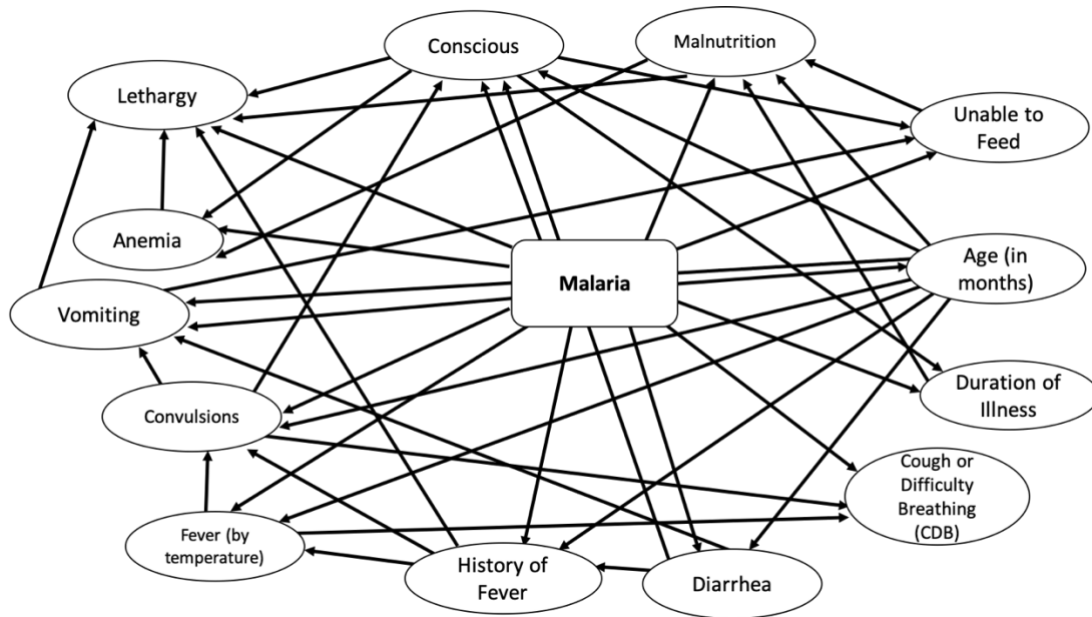


Figure 3: Hybrid BN using ANB structure learning with expert knowledge

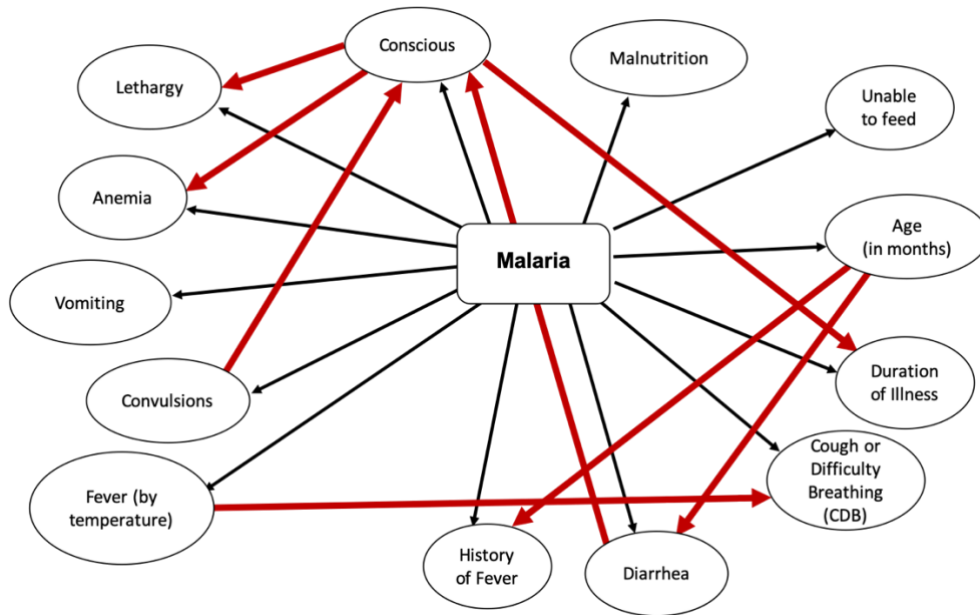


Figure 4: BN with subset of arcs with high strength of influence values

Tables 3 and 4 present the contingency tables for classification of malaria by the BN models on the test dataset. Table 5 shows a set of performance statistics. The model created manually had accuracy of 63.6% and AUC of 0.581. The hybrid model had similar performance in terms of accuracy (62.7%) and AUC (0.583). The manual model achieved a higher F1 score of 0.394 due to better precision (0.5) and recall (0.325) values as compared to the hybrid model. However, the hybrid model had higher specificity (0.848) at the default classification threshold of 0.5 i.e. the ability to classify when malaria is absent.

Table 3: Contingency table for classification of malaria on the test data using the manual BN model

	Malaria Present	Malaria Absent
Predicted Present	27	27
Predicted Absent	56	118
Total	83	145

Table 4: Contingency table for classification of malaria on the test dataset using the hybrid BN model

	Malaria Present	Malaria Absent
Predicted Present	20	22
Predicted Absent	63	123
Total	83	145

Figures 5 (a) and (b) show the ROC curves for classification of malaria using the manual and hybrid BN models respectively.

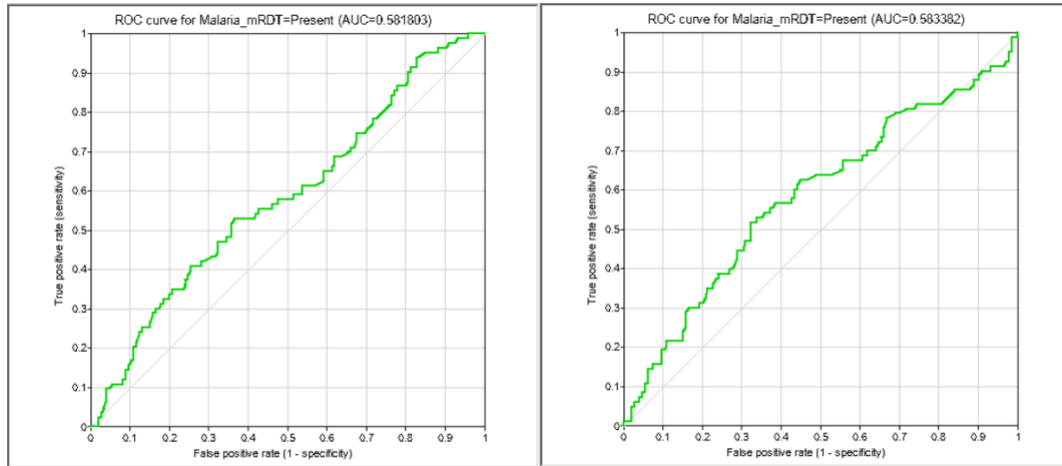


Figure 5: Figure 5: ROC curves for classification of malaria in test dataset using (a) manual BN model and (b) hybrid BN model

3.3 Comparison to Machine Learning Classifiers

Table 5 displays the performance of all models on the test dataset. The accuracy, precision, sensitivity and specificity were derived using a threshold equal to 0.5. While the AUC, accuracy and precision scores were similar for all the models, there was a substantial difference in the sensitivity and specificity values. The machine learning classifiers performed poorly on sensitivity (0.096 and 0.036 for logistic regression and random forest respectively), with low F1 (0.163 and 0.067 respectively) scores as a result. Although the sensitivity was much lower than specificity for all models, the BN models (and the Naïve Bayes classifier) greatly outperformed the logistic

regression and random forest models in their ability to correctly classify cases of positive malaria diagnosis.

Table 5: Performance metrics for all models on the test dataset

Classifier	AUC	Accuracy	Precision	Sensitivity	F1	Specificity
Manual BN Model	0.581	0.636	0.500	0.325	0.394	0.814
Hybrid BN Model	0.583	0.627	0.476	0.241	0.320	0.848
Logistic Regression	0.600	0.640	0.533	0.096	0.163	0.952
Random Forest	0.593	0.636	0.500	0.036	0.067	0.979
Naïve Bayes	0.600	0.600	0.443	0.373	0.405	0.731

3.4 Translation to Clinical Practice

Figure 6 presents the sensitivity analysis of the decision model for implementing clinical decision support (Figure 1). We plotted the posterior probability of malaria given the patient findings from the BN (x-axis) against the expected cost from the decision model (y-axis). The blue and black lines represent the expected cost of the two branches, [mRDT?=yes] and [mRDT?=no], in the decision model respectively. While the expected cost of using the BN model (and not using mDRT) increased at first, it became stable at US \$1.00 when the probability reached 0.1 and above. The expected cost of using the mRDT increased as the probability of the malaria diagnosis, $P(\text{malaria}+|F)$, increased. When $P(\text{malaria}+|F)$ exceeded 0.38, the expected cost of the [mRDT?=yes] branch surpassed that of the [mRDT?=no] branch in the model. The sensitivity analysis suggests that when the probability of malaria diagnosis is between 0 to 0.04 and between 0.38 to 1.0, it is optimal to use the BN model for diagnosis based on the decision model in Figure 1. This can be modified to include additional costs or even utilities for each decision depending on the clinical environment. The decision model and sensitivity analysis provide an effective way to determine the thresholds for decision-making for use of the mRDT and ACT drugs.

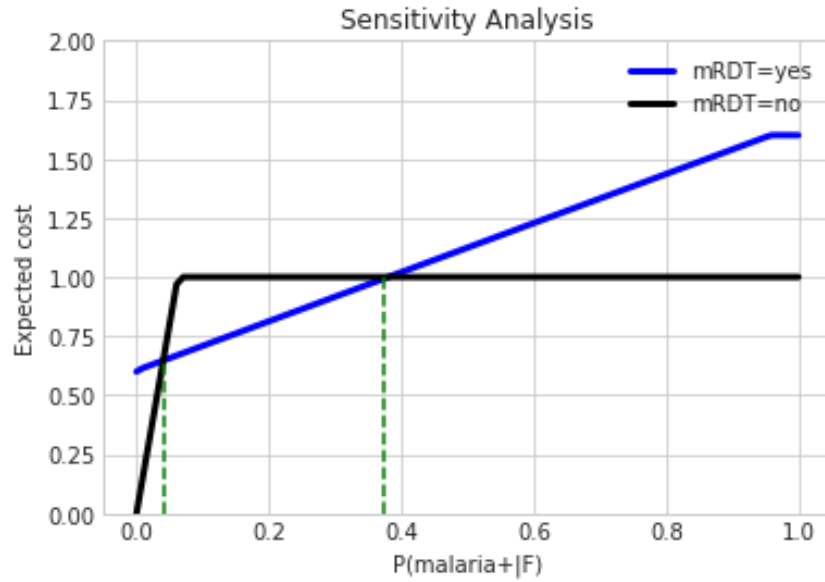


Figure 6: Sensitivity analysis of the decision model in Figure 1.

x-axis is the probability of malaria given the findings (that can be obtained from a BN model) and y-axis is the corresponding expected cost of the decision to use or not to use mRDT. The blue line represents the expected cost of using mRDT and the black line shows the expected cost of not using mRDT. In the probability range 0.04 to 0.38 (green lines) the expected cost of using mRDT is lower, and outside of that range the expected cost of using the BN model (and not using mRDT) is lower.

4.0 Discussion

To the best of our knowledge, this study was the first attempt at developing BN models for diagnosis of childhood malaria in Malawi as well as an LMIC. We used clinical signs and symptoms from a publicly available dataset to predict the malaria diagnosis using models that can be utilized in resource-constrained settings. We implemented two variations of BNs, one created manually where the parameters are learnt from the dataset, and a hybrid model which combined an Augmented Naïve Bayes approach with expert knowledge, and learnt both the structure and the parameters using the dataset.

The current practice for management of suspected malaria in children either involves the mRDT (if available) or prescription of ACT drugs for all children presenting with fever. With increasing evidence of the limited availability of both mRDTs and ACT drugs in rural health centers, this approach is not sustainable in terms of resource use. The BN models developed in this study for classification of malaria demonstrate a unique approach for clinical decision support for childhood illness. We envision the use of these models as an aid to HSAs in Malawi at under-5 clinics and health posts.

The BN models provide further advantages over the current malaria management approach. Fever alone has been found to be a poor indicator of childhood malaria (29), while CDB, anemia status, malnourishment and diarrhea have been previously associated with malaria risk (30,31). The BN models are enriched with variables with known associations to malaria in addition to fever to gain a more accurate picture of individual-level disease risk. These models are designed to work under uncertainty that is inherent in diagnosis of illnesses such as unavailability of certain information. Moreover, BNs take advantage of the value of information concept to reduce uncertainty in the probability of the disease diagnosis in the most efficient way (26). At each observation stage, the model orders nodes in the network such that observing the next node or variable will decrease the uncertainty (and thereby lead to a timely diagnosis) in the network the most. In addition to this, both BN models give interpretable diagnostic support such that the user is able to observe the evidence that leads to the diagnosis through the arcs between the various nodes. While both BN structures offer the above advantages, the choice of the model depends on the performance and the clinical environment for use.

The accuracy, precision and AUC scores are similar for all five models on the test dataset. While the logistic regression and random forest models achieved high specificity at the default threshold, their sensitivity values were extremely low at 9.6% and 3.6% respectively. This implies that these models are unable to correctly classify cases of positive malaria diagnosis as compared to the BN models. The Naïve Bayes model, which assumes independence between all variables, had good consistent performance across all the metrics, although its specificity was lower than that of the BNs. Between the two BNs that are the main focus of this study, we observed that these models were able to classify malaria diagnosis using clinical signs and symptoms with a reasonably good performance on the test dataset. The structure created manually outperformed the hybrid model on the basis of precision and sensitivity, while the hybrid model achieved a higher specificity value. There was no clear winner among the two BN models as each of them has advantages over the other. It is clear, however, that the poor sensitivity values across all the models (as compared to specificity) indicate that classification of positive malaria diagnosis is more challenging than classification of negative malaria cases in the given dataset. The accuracy and AUC scores likewise suggest that there is room for improvement in the models.

While the manually created model is simple and interpretable with good performance, the hybrid model is advantageous in terms of the associations between the various signs and symptoms and is more representative of the clinical scenario. The model in Figure 3 gives the associations between the variables and has been confirmed to be clinically accurate for childhood malaria diagnosis. This model is able to incorporate the known influences of other variables when making a diagnosis. This may be more important in clinical practice when the performance of models is similar or varies only slightly such as in the study by Onisko et al. (15).

Perhaps the greatest advantage of data-driven approaches such as BNs is the ability to do individual-level risk prediction (26). The experimental framework for a decision model in Figure 1 shows a possible extension to the BN diagnostic model that can be used for decision support at the health posts. While we included speculative cost values in this framework, the translation of the BN model to clinical practice in this way allows us to analyze the potential use of the models based on threshold values and expected cost of the decisions. When the probability of malaria given the signs and symptoms of the patient lies between a certain threshold (0.04 to 0.38 with the test costs in Figure 1), and mRDTs and ACT drugs are sufficiently available at the health post, the lower branch in Figure 1 will have a lower expected cost value and the system will advise the

healthcare worker to use the mRDT in more cases (especially when the BN model is uncertain of the outcome). However, in case of limited or no supply of mRDTs (modeled through utility/cost values in the lower branch), the expected cost of the upper branch where mRDT is not used will be lower. In this scenario, the BN model informs the decision on whether or not to treat based on the symptoms observed and the probability of the outcome. Further, evidence against malaria from the model can be indicative of the need for a differential diagnosis of the illness.

The costs in the decision model can be varied depending on the resources available at each healthcare facility. As previously developed tools have not focused on a data-driven approach to diagnosis (13), we believe that the BN model for malaria diagnosis in this study is the first step toward developing an evidence-based eCDA to provide diagnostic support at the primary points of care. Since the dataset used to learn the BN models was representative of the population in Malawi, the same approach can be generalized to any number of infectious diseases in other LMIC. The BN model provides a simple, interpretable and powerful way to diagnose diseases, and at the same time allows transparency in what goes into the model and how the decision is taken. Consequently, it also enables automated data collection for each child whose symptoms are reported using the model. We can further improve the model iteratively by updating the parameters and interactions based on the data collected during use and expert judgment.

The models developed in this study, however, must be considered in light of data limitations. The choice of the mRDT result as the target diagnosis in the dataset reduced the size from the original dataset substantially. A smaller dataset produces more unreliable estimates of the parameters of the BN models as compared to models developed with larger datasets. Further, the selected dataset may yield biased predictions that are not representative of the outcomes in the remainder of the dataset. However, this is the only dataset that we know of with both gold-standard diagnosis and symptoms of childhood malaria available. Further, our choice of variables was constrained by the information collected in the SPA survey which did not include details such as the immunization and HIV status of the children. The proportion of children with malnutrition in the dataset was also much less than the reported prevalence in the country (32). Thus, there may exist latent associations in the models in this study. We believe that the choice of mRDT result as the gold standard diagnosis was the best possible decision given the dataset, and the WHO reported sensitivity and specificity of the test. As the type of mRDT and procedure of the test was not made

available with the dataset, we cannot verify the reported outcome. Information about prior exposure to anti-malarial drugs would also be useful in this context.

While this study developed and validated the models with the same dataset, external validation with appropriate feedback from the healthcare providers in Malawi would be valuable to guide the next steps to refine and use the model for diagnosis. Additionally, incorporation of the model into an mHealth application will make it easier for the healthcare workers to access it along the patient's clinical course. We further recognize that the model will be strengthened by adding more diseases in the differential diagnosis. We continue to search for datasets for other common childhood illnesses that will make it possible for us to do so.

Bibliography

1. National Malaria Control Programme (NMCP), ICF. Malawi Malaria Indicator Survey. 2017;2.
2. Chart Booklet Integrated Management of Childhood Illness. 2014.
3. Health Organization W. World malaria report 2015.
4. Lufesi NN, Andrew M, Aursnes I. Deficient supplies of drugs for life threatening diseases in an African community. *BMC Health Serv Res.* 2007;7:1–7.
5. Klootwijk L, Chirwa AE, Kabaghe AN, Van Vugt M. Challenges affecting prompt access to adequate uncomplicated malaria case management in children in rural primary health facilities in Chikhwawa Malawi. *BMC Health Serv Res.* 2019 Oct 22;19(1):735.
6. MoH. Malawi Service Provision Assessment (SPA) 2013-14. 2014;1.
7. Kabaghe AN, Phiri MD, Phiri KS, Van Vugt M. Challenges in implementing uncomplicated malaria treatment in children: A health facility survey in rural Malawi. *Malar J.* 2017 Oct 18;16(1):419.
8. Shillcutt S, Morel C, Goodman C, Coleman P, Bell D, Whitty CJM, et al. Cost-effectiveness of malaria diagnostic methods in sub-Saharan Africa in an era of combination therapy. *Bull World Health Organ.* 2008 Feb;86(2):101–10.
9. Bjornstad E, Preidis GA, Lufesi N, Olson D, Kamthunzi P, Hosseinipour MC, et al. Determining the quality of IMCI pneumonia care in Malawian children. *Paediatr Int Child Health.* 2014 Feb 6;34(1):29–36.
10. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Heal.* 2018 Aug 1;3(4):e000798.
11. Hosny A, Aerts HJWL. Artificial intelligence for global health. Vol. 366, *Science.* 2019. p. 955–6.
12. Connor YO, Sullivan TO, Gallagher J, Heavin C, Donoghue JO. Developing eXtensible mHealth Solutions for Low Resource Settings. In: Prasath R, O'Reilly P, Kathirvalavakumar T, editors. *Mining Intelligence and Knowledge Exploration.* Cham: Springer International Publishing; 2014. p. 361–71.

13. Keitel K, D'Acremont V. Electronic clinical decision algorithms for the integrated primary care management of febrile children in low-resource settings: review of existing tools. Vol. 24, *Clinical Microbiology and Infection*. Elsevier B.V.; 2018. p. 845–55.
14. Friedman N, Geiger D, Goldszmidt M. Bayesian Network Classifiers. *Mach Learn*. 1997;29(2–3):131–63.
15. Onisko A, Druzdzel MJ, Wasyluk H. A probabilistic causal model for diagnosis of liver disorders. *Proc Seventh Symp Intell Inf Syst*. 1998;(September 2016):379–87.
16. Kraisangka J, Druzdzel MJ, Benza RL. A Risk Calculator for the Pulmonary Arterial Hypertension Based on a Bayesian Network. *Work Notes 13th Annu Bayesian Model Appl Work*. 2016;1–59.
17. Berkan Sesen M, Nicholson AE, Banares-Alcantara R, Kadir T, Brady M. Bayesian networks for clinical decision support in lung cancer care. *PLoS One*. 2013;8(12):82349.
18. Uwemedimo OT, Lewis TP, Essien EA, Chan GJ, Nsona H, Kruk ME, et al. Distribution and determinants of pneumonia diagnosis using Integrated Management of Childhood Illness guidelines: A nationally representative study in Malawi. *BMJ Glob Heal*. 2018;3(2):1–12.
19. The DHS Program - Quality information to plan, monitor and improve population, health, and nutrition programs [Internet]. [cited 2020 Mar 14]. Available from: <https://dhsprogram.com/>
20. Organization WH, others. Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 8 (2016--2018). 2018;
21. Ministry of Health Malawi. Guidelines for the treatment of malaria in Malawi. 4th ed. Malawi: Malawi Government; 2013.
22. Cheng J, Bell DA, Liu W. An Algorithm for Bayesian Belief Network Construction from Data. *Proc Sixth Int Work Artif Intell Stat (AI STAT '97)*. 1997;83–90.
23. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2011;12(Oct):2825–30.
24. Kjaerulff UB, Madsen AL. Bayesian networks and influence diagrams. *Springer Sci Bus Media*. 2008;200:114.
25. Khuluza F, Heide L. Availability and affordability of antimalarial and antibiotic medicines in Malawi. Postma M, editor. *PLoS One*. 2017 Apr 18;12(4):e0175399.

26. Arora P, Boyne D, Slater JJ, Gupta A, Brenner DR, Druzdzal MJ. Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine. *Value Heal.* 2019 Apr 1;22(4):439–45.
27. Koiter JR. Visualizing inference in Bayesian networks. *Man-machine Interact Gr.* 2006;Master of.
28. D’Acromont V, Kilowoko M, Kyungu E, Philipina S, Sangu W, Kahama-Marro J, et al. Beyond Malaria — Causes of Fever in Outpatient Tanzanian Children. *N Engl J Med.* 2014;370(9):809–17.
29. Mfueni E, Devleeschauwer B, Rosas-Aguirre A, Van Malderen C, Brandt PT, Ogutu B, et al. True malaria prevalence in children under five: Bayesian estimation using data of malaria household surveys from three sub-Saharan countries. *Malar J.* 2018;17(1):1–7.
30. Reithinger R, Ngondi JM, Graves PM, Hwang J, Getachew A, Jima D, et al. Risk factors for anemia in children under 6 years of age in Ethiopia: analysis of the data from the cross-sectional Malaria Indicator Survey, 2007. *Trans R Soc Trop Med Hyg.* 2013 Dec 1;107(12):769–76.
31. Feasey NA, Everett D, Faragher EB, Roca-feltrer A, Kang A, Denis B, et al. Modelling the Contributions of Malaria , HIV , Malnutrition and Rainfall to the Decline in Paediatric Invasive Non-typhoidal Salmonella Disease in Malawi. 2015;1–12.
32. ICF. Malawi Demographic and Health Survey 2015-16 . Zomba, Malawi : National Statistical Office and ICF ; 2017.