University of Pittsburgh

Thesis Defense

Sanya Bathla Taneja
Intelligent Systems Program, School of Computing and Information
April 10, 2020; 3.30pm EST

Bayesian Networks for Diagnosing Childhood Malaria in Malawi

Sanya Bathla Taneja

Primary Advisor: Dr. Shyam Visweswaran

Can a *data-driven approach* to diagnosis of childhood illnesses address the challenges faced in *health centers in low-resource countries*?

Diagnostic support in high-income countries

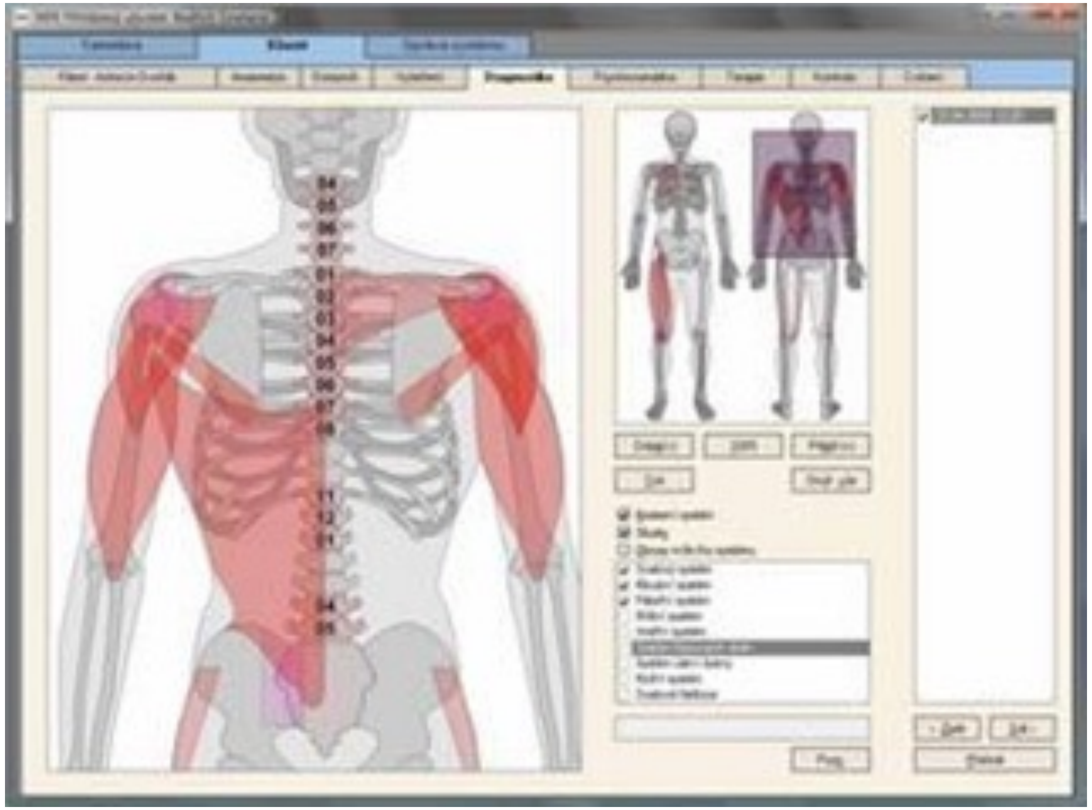
Diagnostic Expert Systems

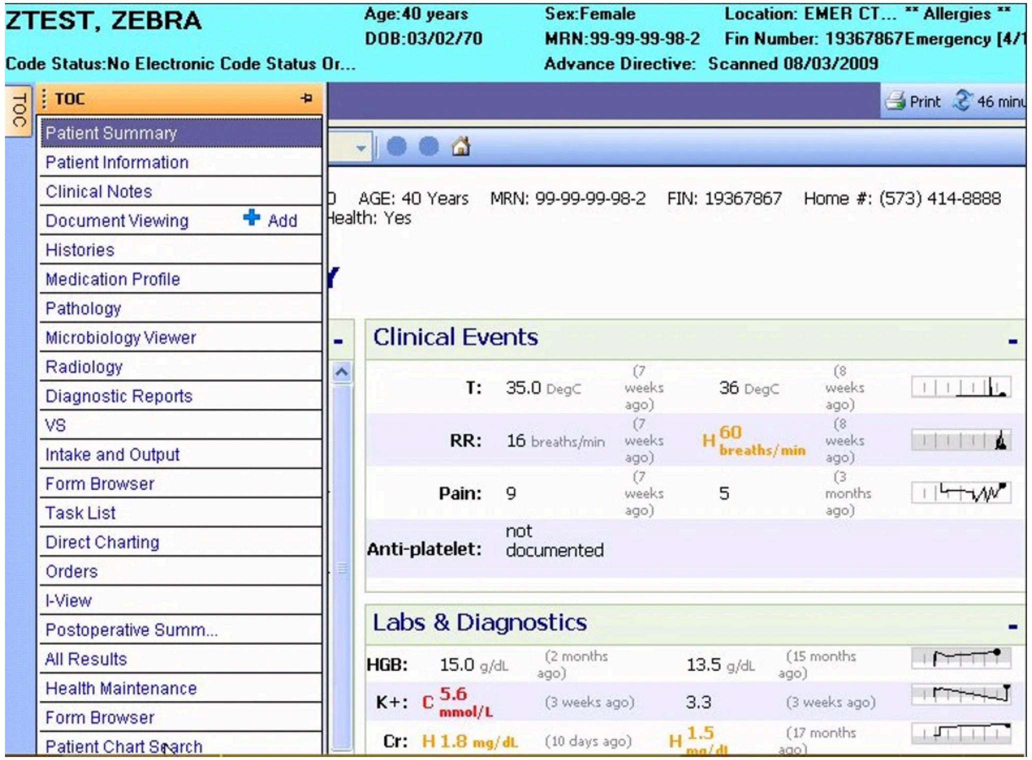
Medication Administration

Admission Registration

Computerized Physician Order Entry (CPOE)

Lab, Radiology, Pharmacy Systems



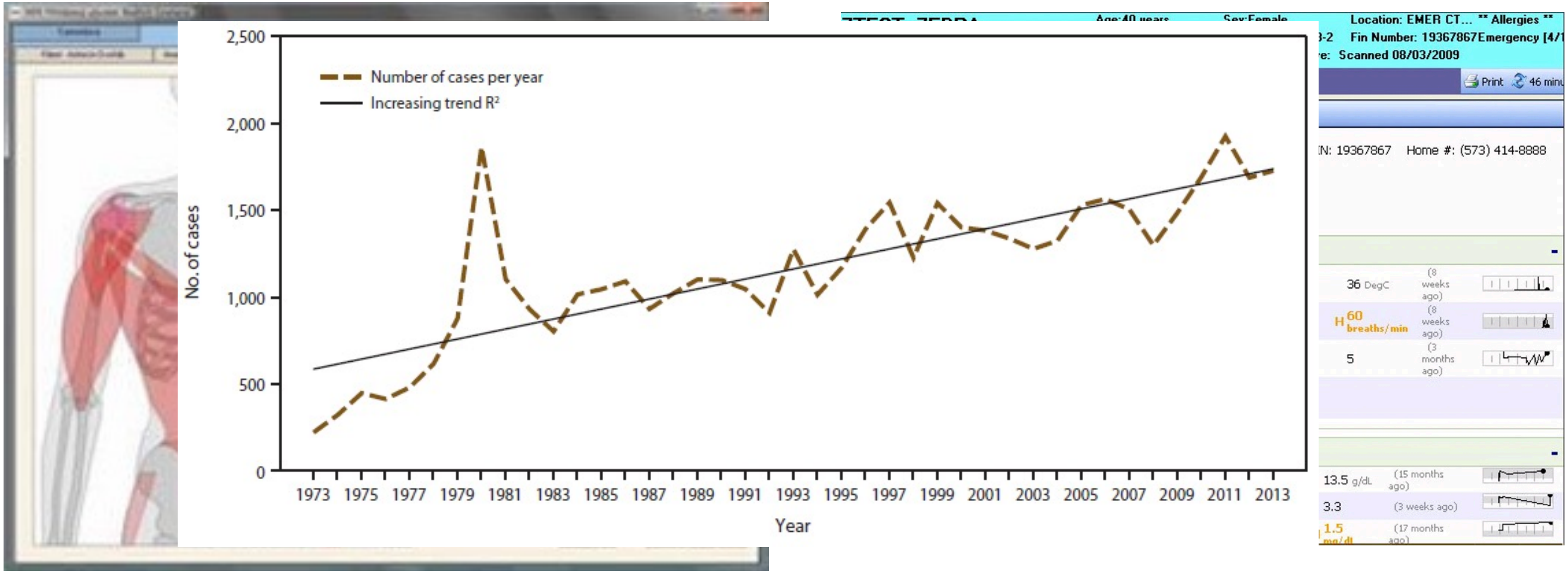


https://www.softwareadvice.com/medical/cerner-powerchart-profile/

https://www.ictlounge.com/html/expert_systems.htm

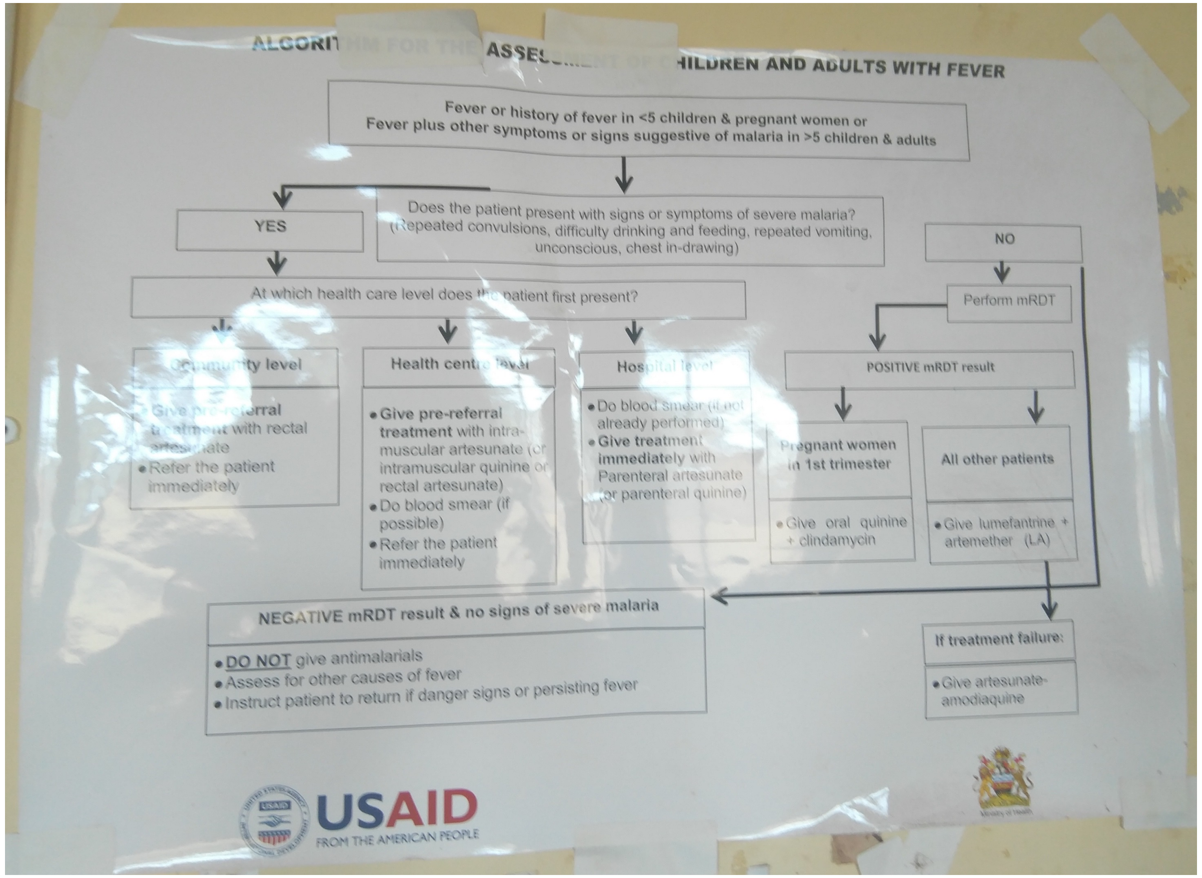
Diagnostic support in high-income countries

Number of malaria cases among U.S. military personnel and U.S. and foreign civilians — United States, 1973–2013*

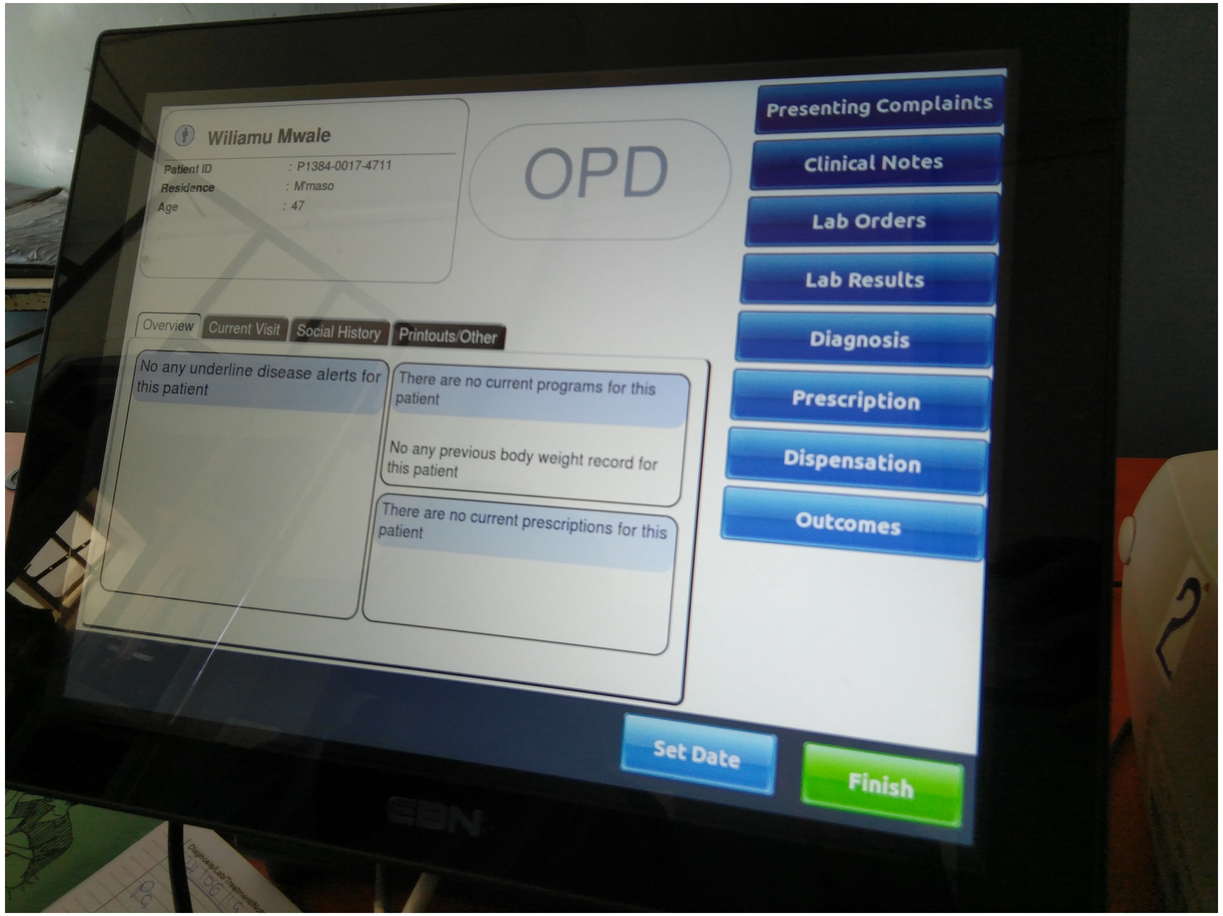


Diagnostic support in low-resource countries

Wall Chart: Algorithm for Assessment of Children and Adults with Fever (USAID)



System used for prescriptions, Malawi



Diagnostic support in low-resource countries

WHO and ITU establish benchmarking process for artificial intelligence in health

Thomas Wiegand • Ramesh Krishnamurthy • Monique Kuglitsch • Naomi Lee • Sameer Pujari • Marcel Salathé • et al. Show all authors

Published: March 29, 2019 • DOI: https://doi.org/10.1016/S0140-6736(19)30762-7 • Check for updates

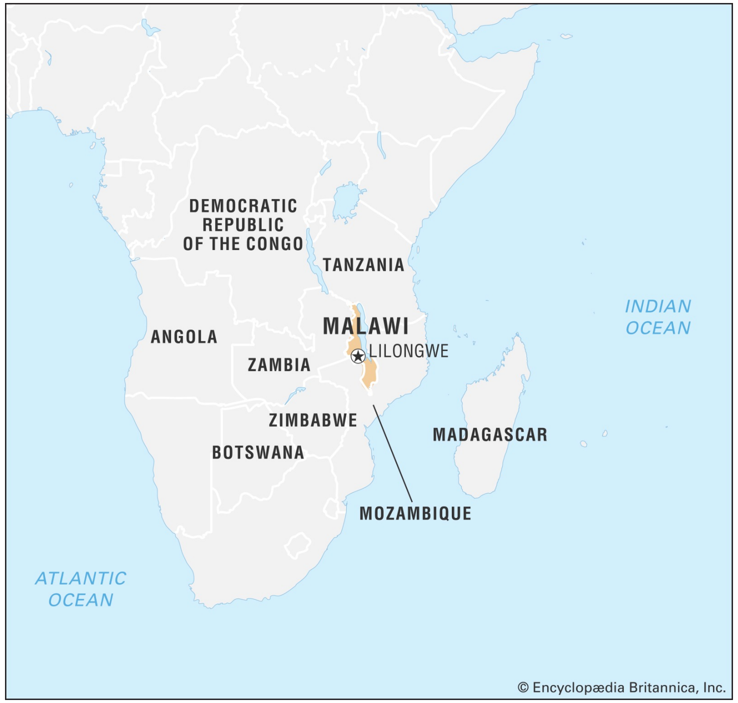
PERSPECTIVE | GLOBAL HEALTH

Artificial intelligence for global health

Ahmed Hosny[1,2], Hugo J. W. L. Aerts[1,2,3]
+ See all authors and affiliations

Science 22 Nov 2019:

Healthcare in Malawi



Malawi
Southeastern Africa

- Public and private health facilities
- Levels of public healthcare: free of cost
  - District and Central Hospitals
  - Health Centers and Health Posts
  - Village Clinics
- Health Centers and Posts are primary points of care
- High burden of infectious diseases in childhood population
  - Malaria
  - Pneumonia
  - Tuberculosis
  - Diarrheal disease

Challenges in health centers and posts

Infrastructure

- Limited laboratory facilities (microscopy and rapid tests)
- No admissions (except maternity)
- Limited network and connectivity
- May or may not have ambulance on site
- Under-5 clinics in building outposts

Ngoni Health Center, Malawi

Challenges in health centers and posts

Infrastructure

Clinical Support

- 1 health worker per 8000 people
- 1 medical officer + 2 nurses at health center usually
- Health Surveillance Assistants (HSA) with 12-week training – run village clinics and immunization camps
- Shortage of healthcare workers overall

Ngoni Health Center, Malawi

Challenges in health centers and posts

Infrastructure

Clinical Support

Resources

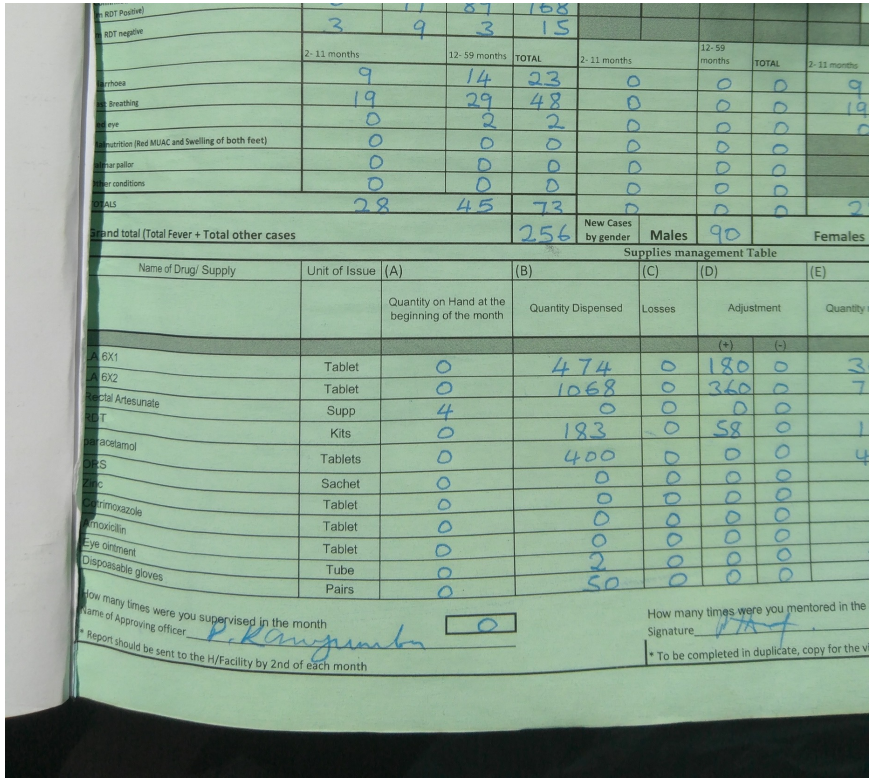
- Stock outs of rapid tests
- Erratic supply of drugs/ACTs
- Limited electronic record systems (pharmacy and lab)
- Paper register-based data entry

Ngoni Health Center, Malawi

Childhood Malaria in Malawi

- Prevalence
  - 24% overall
  - 48% in rural areas
- Primary points of care:
  - Under-5 clinics: twice a week
  - Health posts and health centers
- WHO-recommended diagnosis with malaria rapid diagnostic test (mRDT)
- Treatment with artemisinin-based combination therapy (ACT) drugs
- High burden of care
- Data collection not as important

Summary data from under-5 clinic, Malawi



RDT Positive)				168
RDT negative	3	9	3	15

	2- 11 months	12- 59 months	TOTAL	2- 11 months	12- 59 months	TOTAL	2- 11 months
iarrhoea	9	14	23	0	0	0	9
ast Breathing	19	29	48	0	0	0	19
ed eye	0	2	2	0	0	0	0
alnutrition (Red MUAC and Swelling of both feet)	0	0	0	0	0	0	
almar pallor	0	0	0	0	0	0	
ther conditions	0	0	0	0	0	0	
OTALS	28	45	73	0	0	0	2

Grand total (Total Fever + Total other cases 256 | New Cases by gender | Males 90 | Females

Supplies management Table

Name of Drug/ Supply	Unit of Issue	(A) Quantity on Hand at the beginning of the month	(B) Quantity Dispensed	(C) Losses	(D) Adjustment (+)	(D) Adjustment (-)	(E) Quantity
LA 6X1	Tablet	0	474	0	180	0	3
LA 6X2	Tablet	0	1068	0	360	0	7
Rectal Artesunate	Supp	4	0	0	0	0	
RDT	Kits	0	183	0	58	0	1
paracetamol	Tablets	0	400	0	0	0	4
ORS	Sachet	0	0	0	0	0	
Zinc	Tablet	0	0	0	0	0	
Cotrimoxazole	Tablet	0	0	0	0	0	
Amoxicillin	Tablet	0	0	0	0	0	
Eye ointment	Tube	0	2	0	0	0	
Disposable gloves	Pairs	0	50	0	0	0	

How many times were you supervised in the month 0

Name of Approving officer P. Kanyumba

* Report should be sent to the H/Facility by 2nd of each month

How many times were you mentored in the

Signature

* To be completed in duplicate, copy for the vi

Childhood Malaria in Malawi

Fever Present

Malaria risk: high or low

Other symptoms?
Stiff neck, runny nose, bacterial cause, measles

no → Malaria Rapid Diagnostic Test (mRDT)

yes → Stiff neck or danger sign

Classify Fever

Malaria Rapid Diagnostic Test (mRDT)
+ve → MALARIA
-ve → FEVER NO MALARIA

Stiff neck or danger sign → VERY SEVERE FEBRILE DISEASE

Treat for:

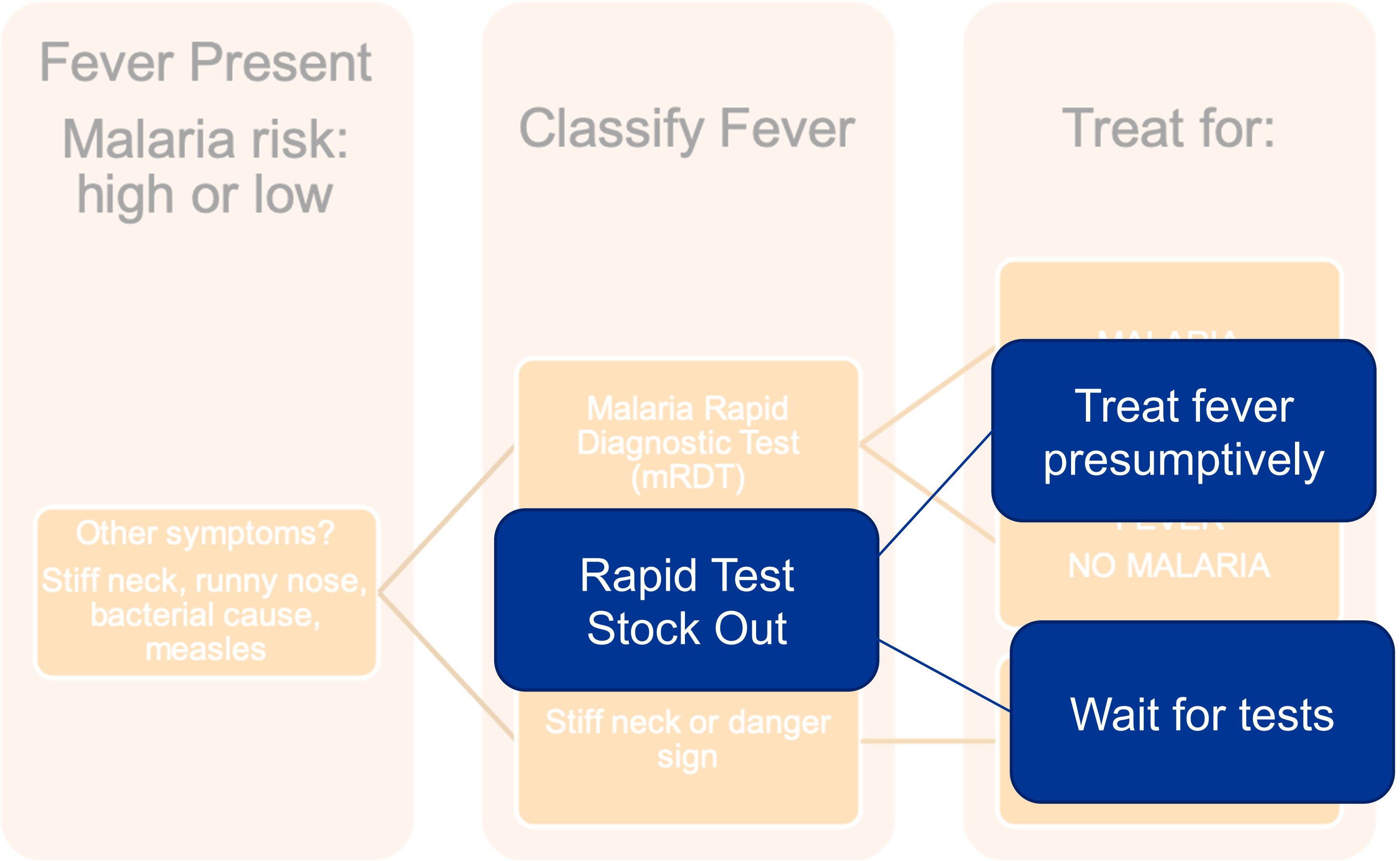
MALARIA

FEVER
NO MALARIA

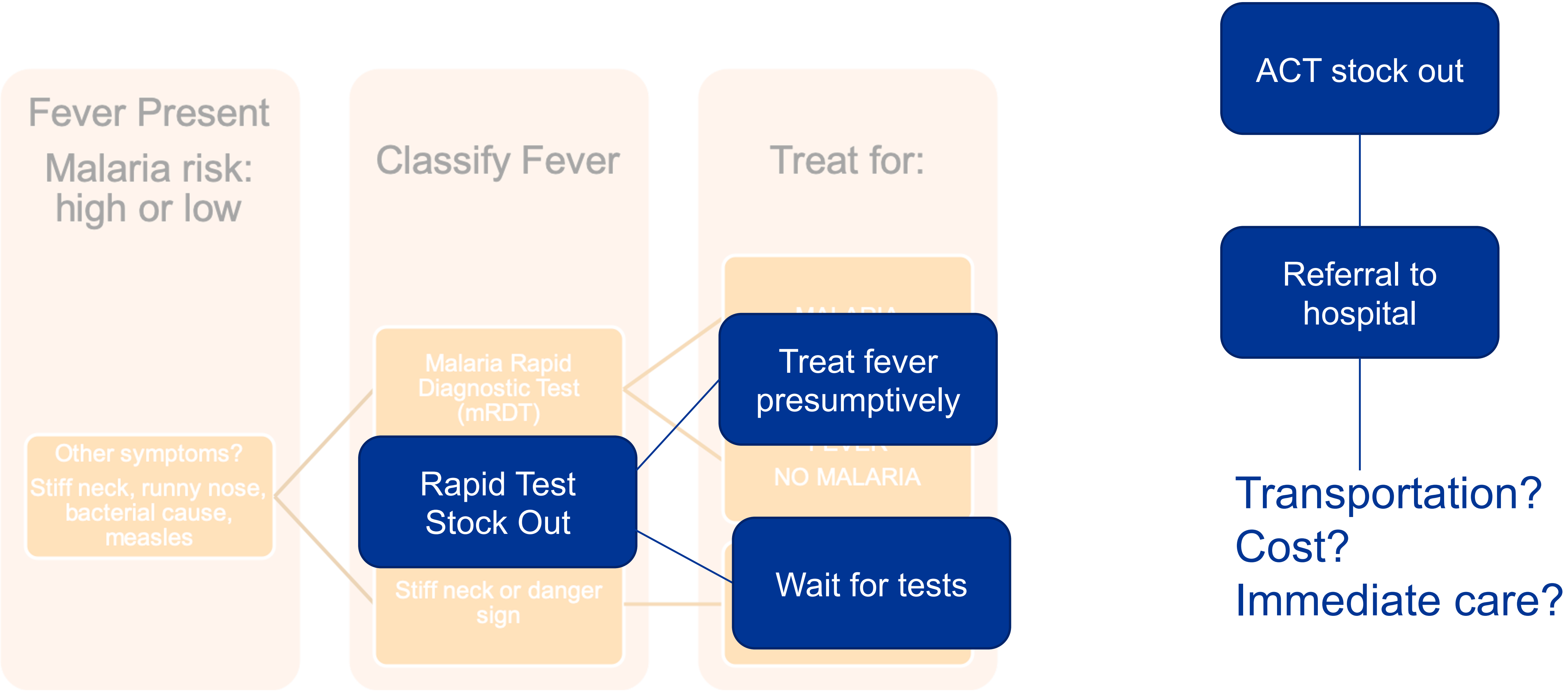
VERY SEVERE FEBRILE DISEASE

Batch of rapid tests

Childhood Malaria in Malawi



Childhood Malaria in Malawi



Can a *data-driven approach* to diagnosis of childhood malaria *improve care and resource utilization* in health centers and health posts in Malawi?

AIMS

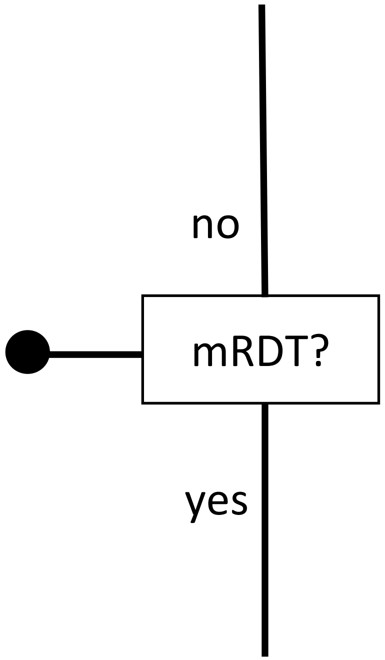
Develop and assess Bayesian diagnostic model using data

Use structure learning to find dependencies between variables

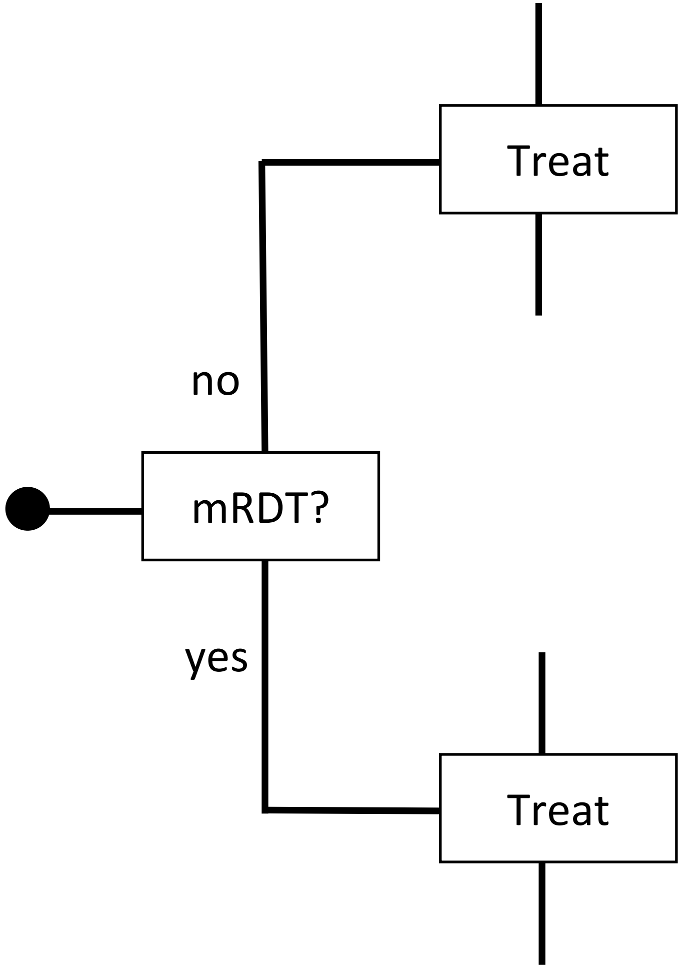
Compare model performances with supervised machine learning methods

Sensitivity analysis for diagnostic support at health posts

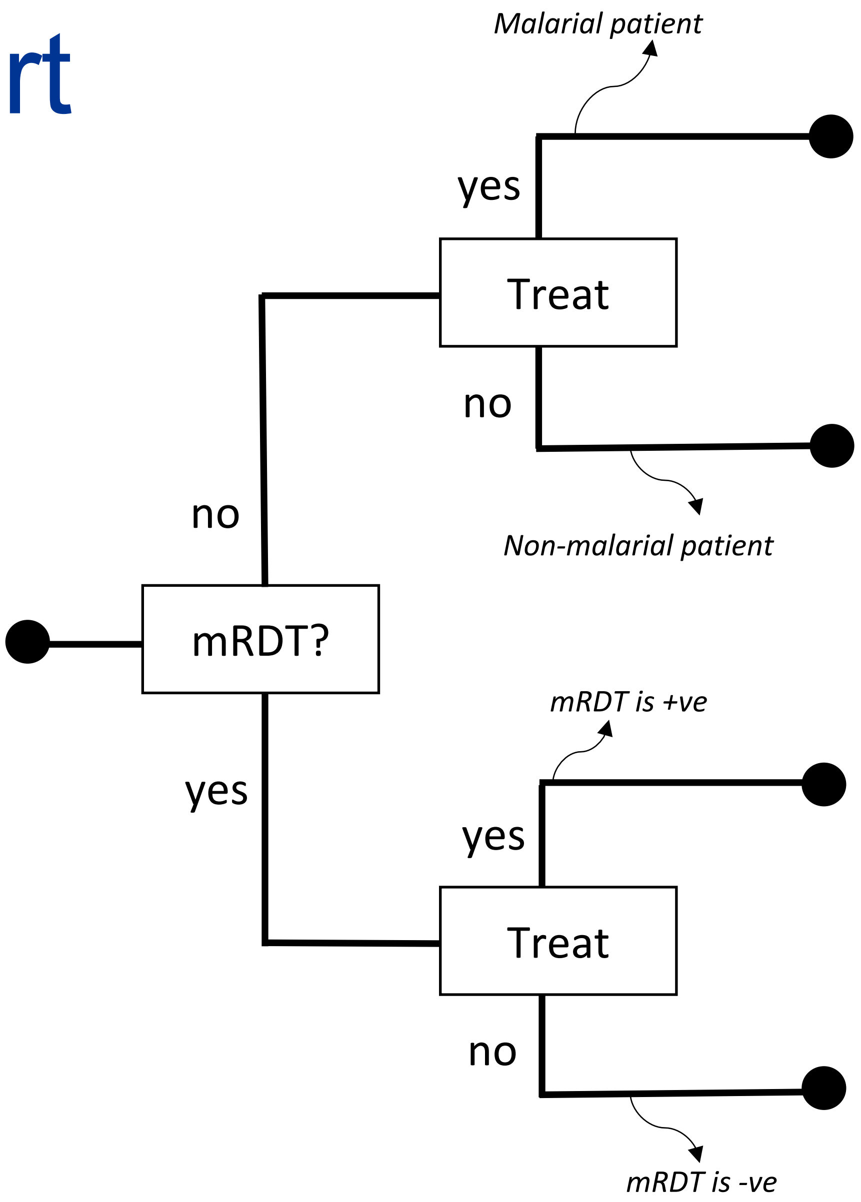
Decision Support



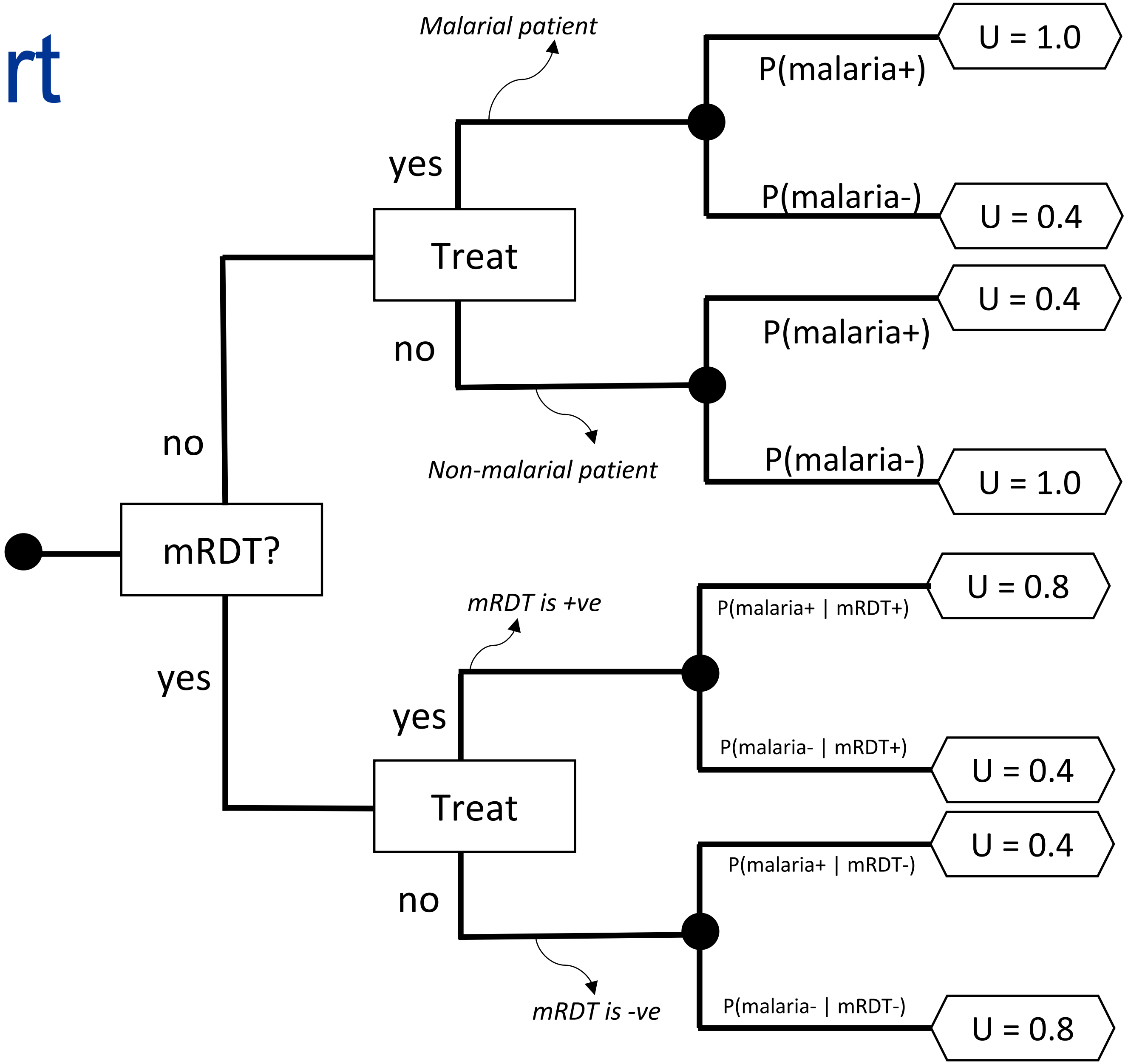
Decision Support



Decision Support



Decision Support



Decision Support

Expected utility of [mRDT?=no] =

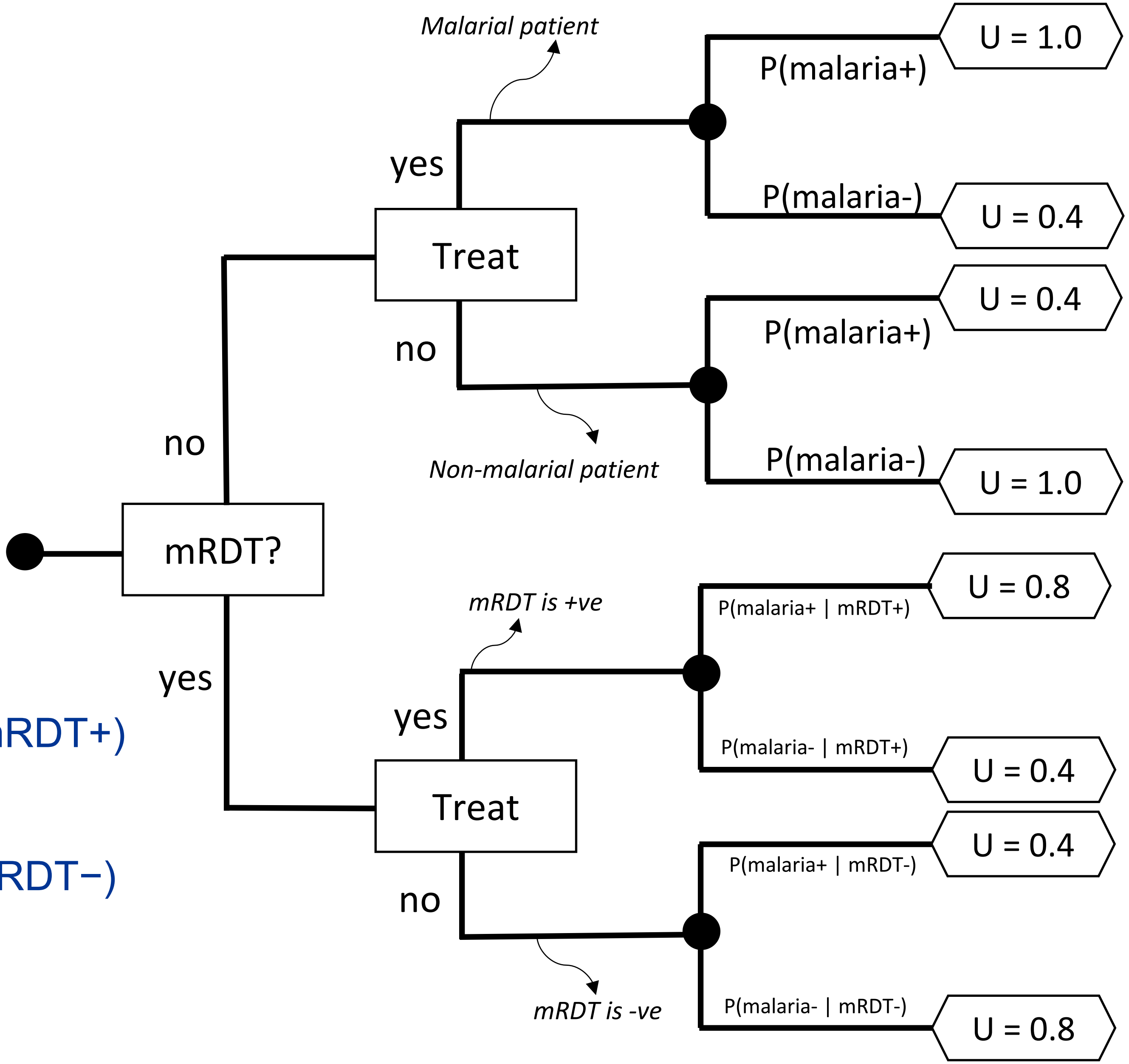
$$\max \begin{cases} 1.0 * P(\text{malaria}+) + 0.4 * P(\text{malaria}-) \\ 0.4 * P(\text{malaria}+) + 1.0 * P(\text{malaria}-) \end{cases}$$

Expected utility of [mRDT?=yes] =

$$\max \begin{cases} 0.8 * P(\text{malaria}+|\text{mRDT}+) + 0.4 * P(\text{malaria}-|\text{mRDT}+) \\ 0.4 * P(\text{malaria}+|\text{mRDT}-) + 0.8 * P(\text{malaria}-|\text{mRDT}-) \end{cases}$$

P(malaria+|mRDT+) = sensitivity of mRDT

P(malaria−|mRDT−) = specificity of mRDT



Bayesian Networks

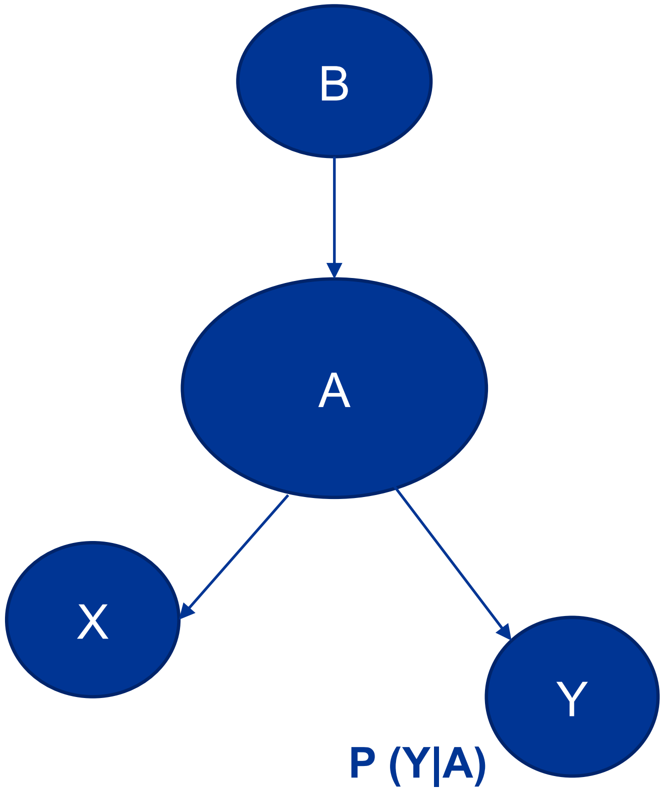
- Probabilistic graphical models
- Nodes as variables and arcs represent relationships

- Joint probability distribution given the evidence

  P (malaria+) == P(mRDT+)

- Structure: manual or automated

- Parameters or conditional probability distributions: data or experts



P (Y|A)

A	T	F
T	0.7	0.3
F	0.1	0.9

Bayesian Networks and Data

- Individual-level diagnosis or risk prediction
- Probabilistic modeling
  - Get rank of diseases
  - Allow concurrent illnesses
- Value of information
  - Eliminate steps based on likelihood of critical observations
- Region-specific prevalence and analysis
- Visualize interactions and contributions

Service Provision Assessment (SPA) survey

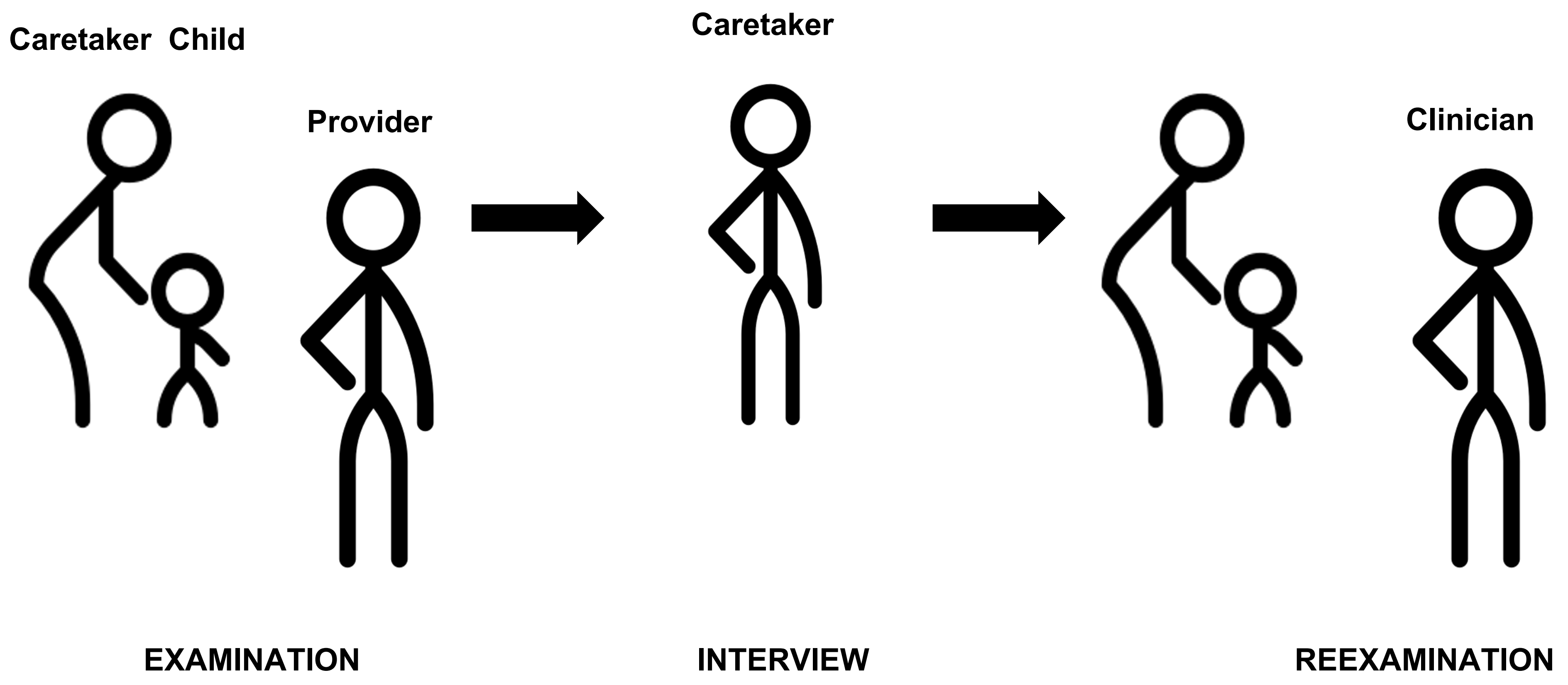
- Malawi Ministry of Health and DHS program, 2013-14

- 977 facilities in 3 major regions
  - Hospitals, health centers, dispensaries, clinics and health posts

- Primary goal: assess facilities and healthcare workers

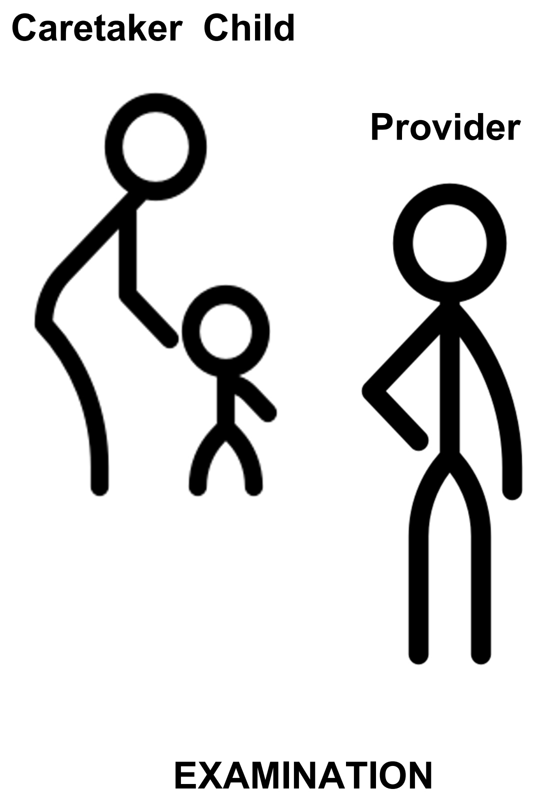
- Secondary:
  - 3,441 sick child observations
  - 1,139 (33%) with known mRDT result
    - Demographics
    - Signs and symptoms

Service Provision Assessment (SPA) survey



https://www.iconfinder.com/iconsets/stick-figures-in-motion

Service Provision Assessment (SPA) survey



Malnutrition

Malaria Rapid Diagnostic Test (mRDT) result

Service Provision Assessment (SPA) survey

Caretaker

Age of child

Diarrhea

History of Fever

Convulsions

Duration of Illness

Unable to Feed

Lethargy

Vomiting

INTERVIEW

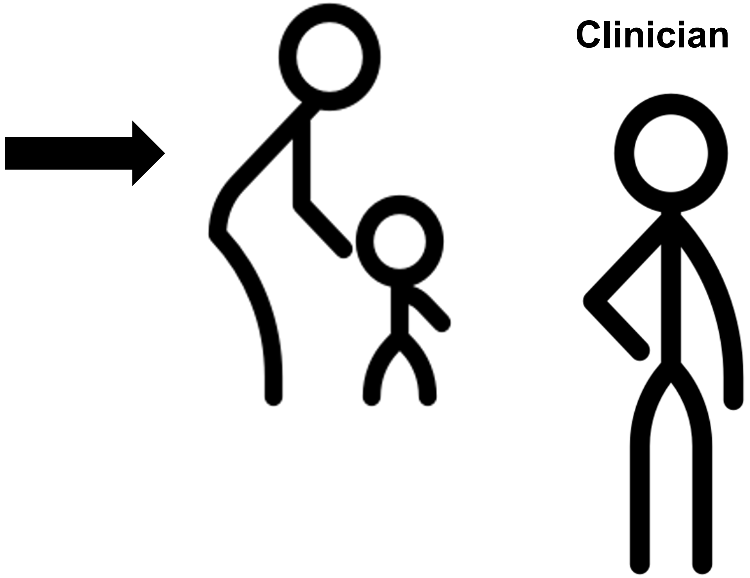
Service Provision Assessment (SPA) survey

Fever (temperature)

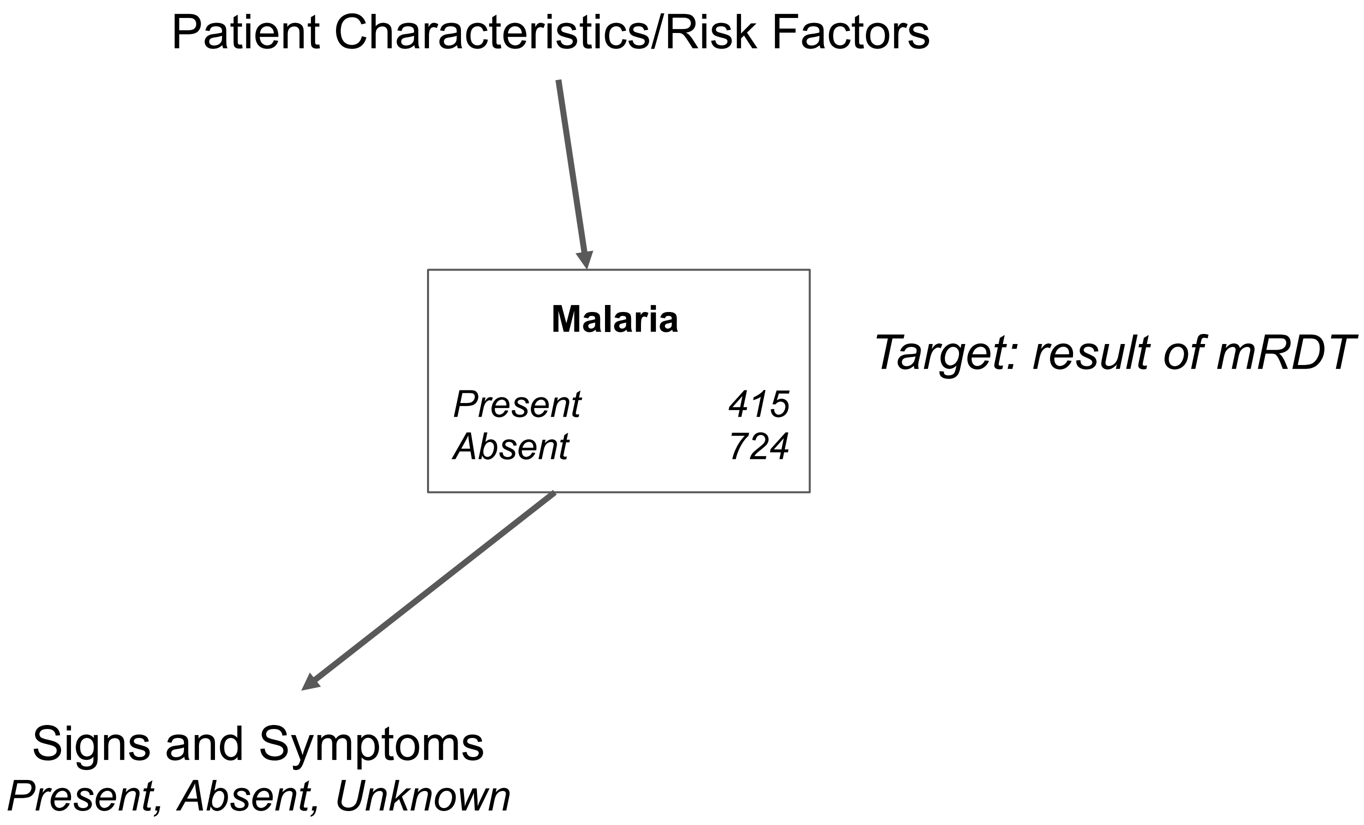
Cough or Difficulty Breathing (CDB)

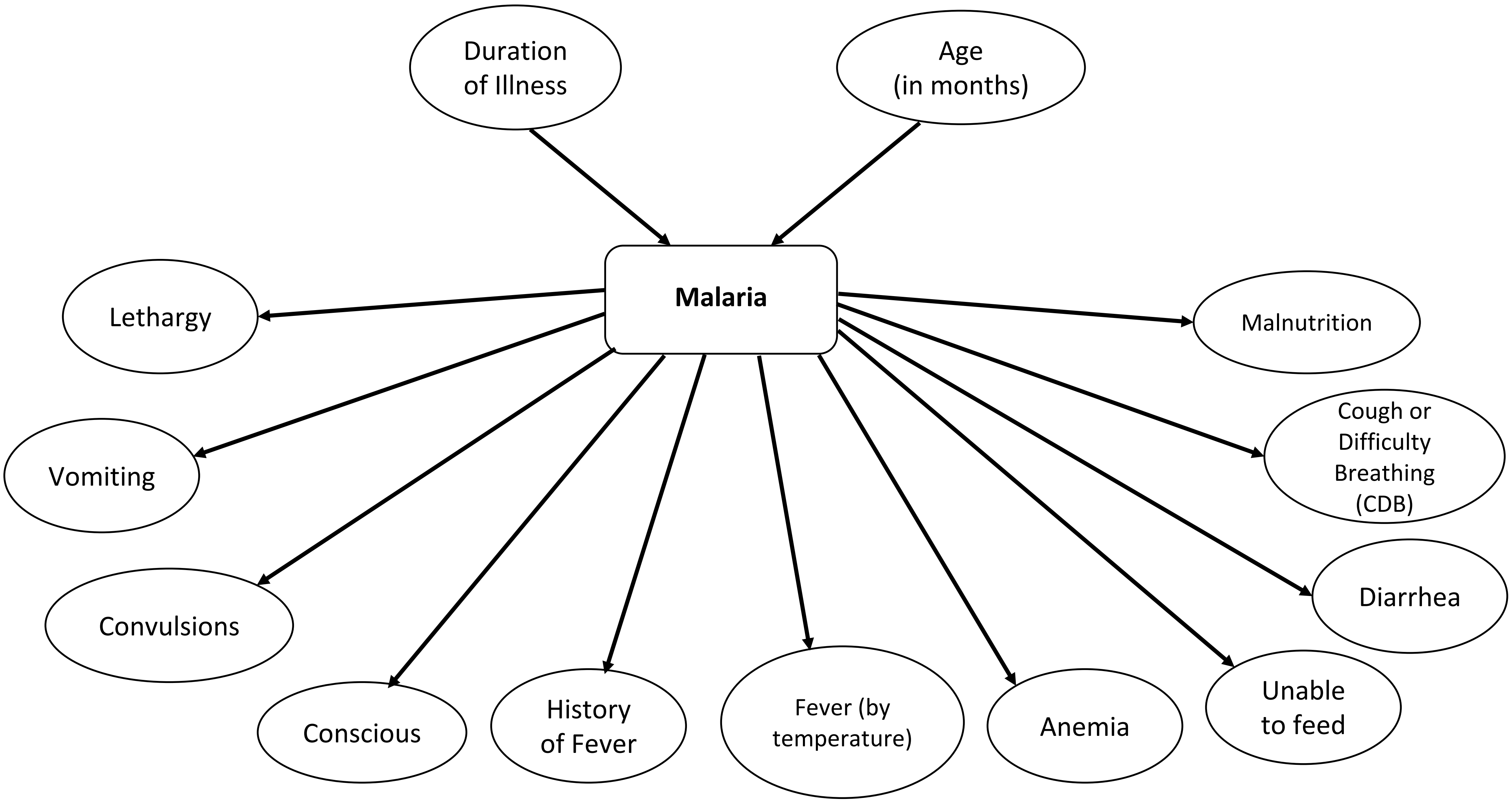
Conscious

Anemia

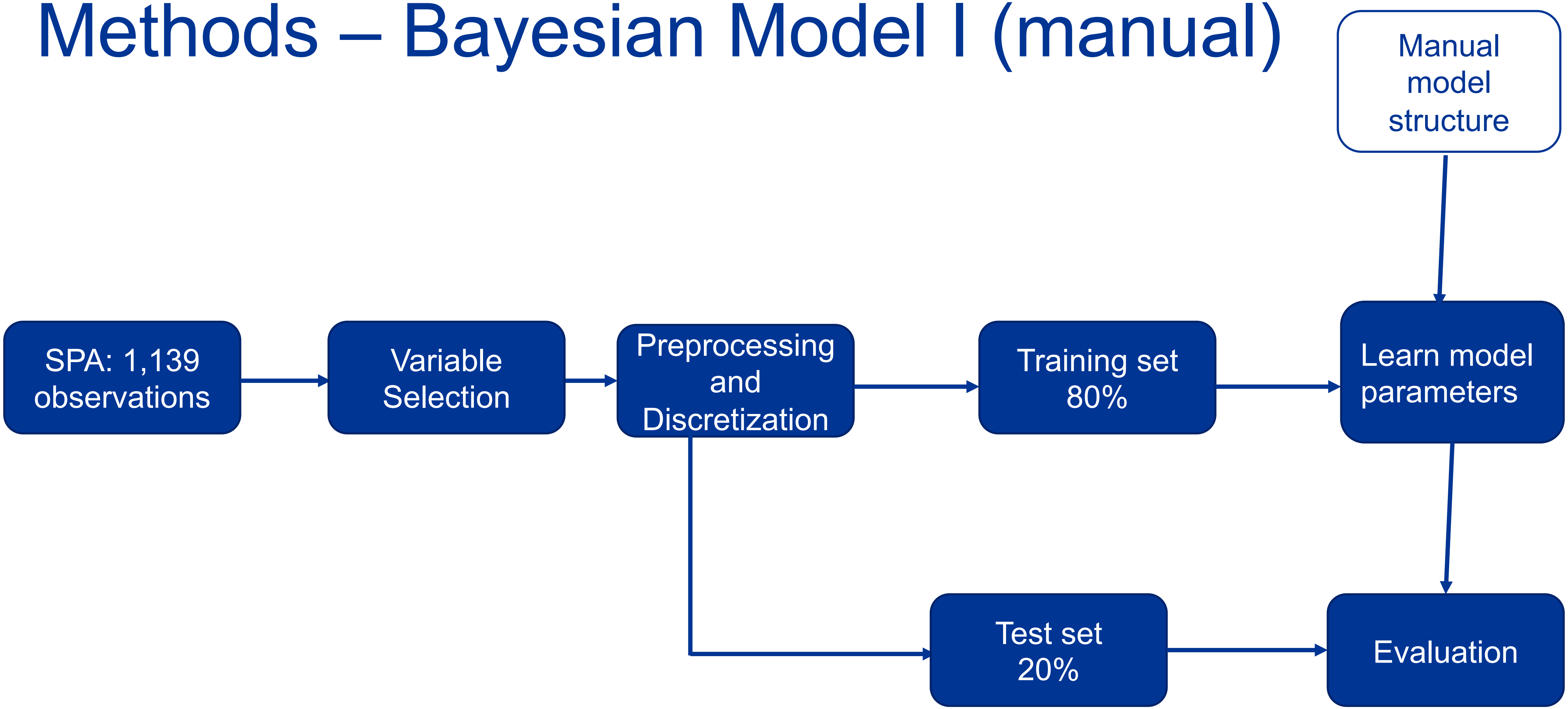


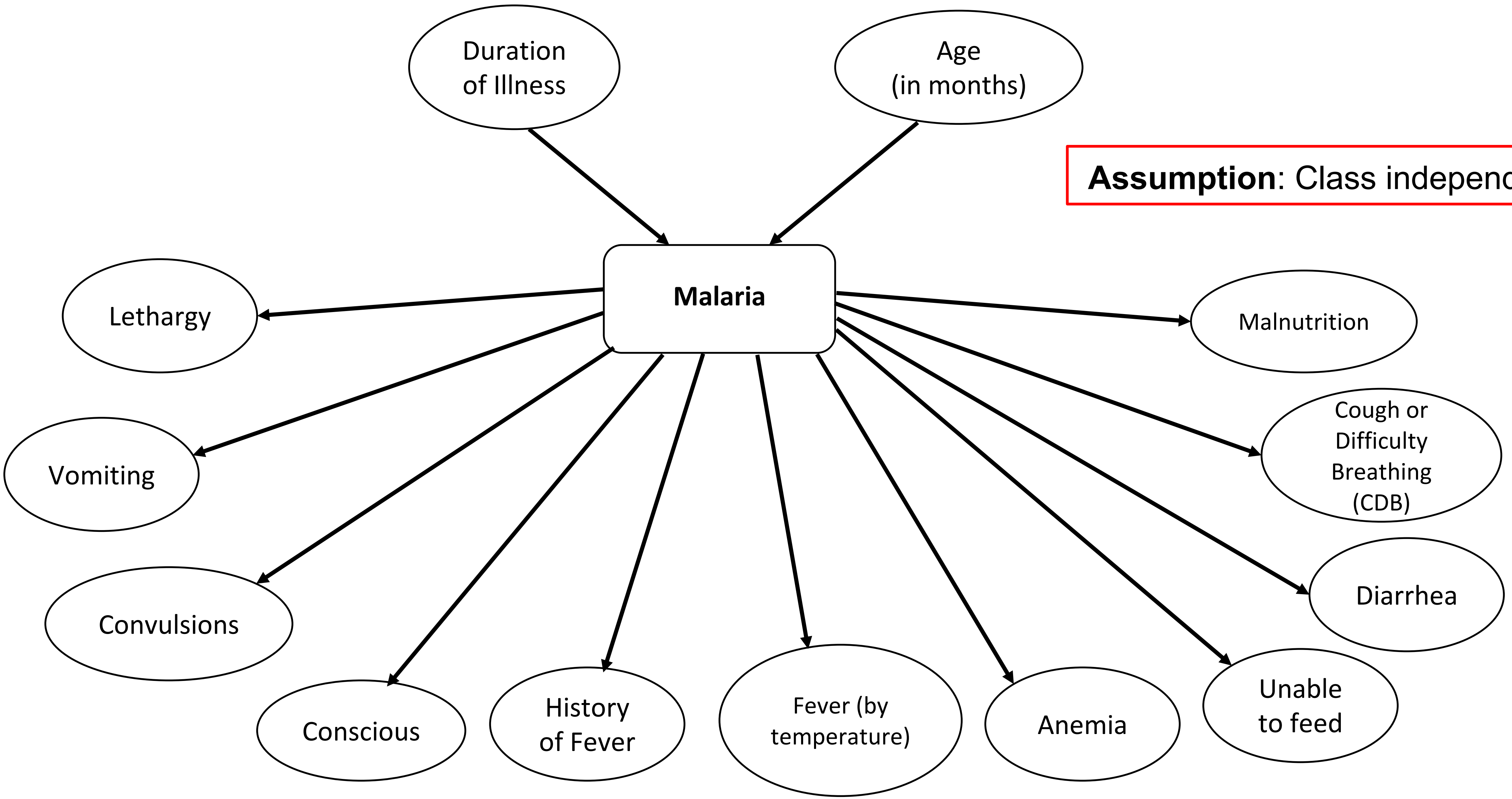
Methods – Bayesian Model I




Duration of Illness
Age (in months)
Malaria
Lethargy
Malnutrition
Vomiting
Cough or Difficulty Breathing (CDB)
Convulsions
Diarrhea
Conscious
History of Fever
Fever (by temperature)
Anemia
Unable to feed

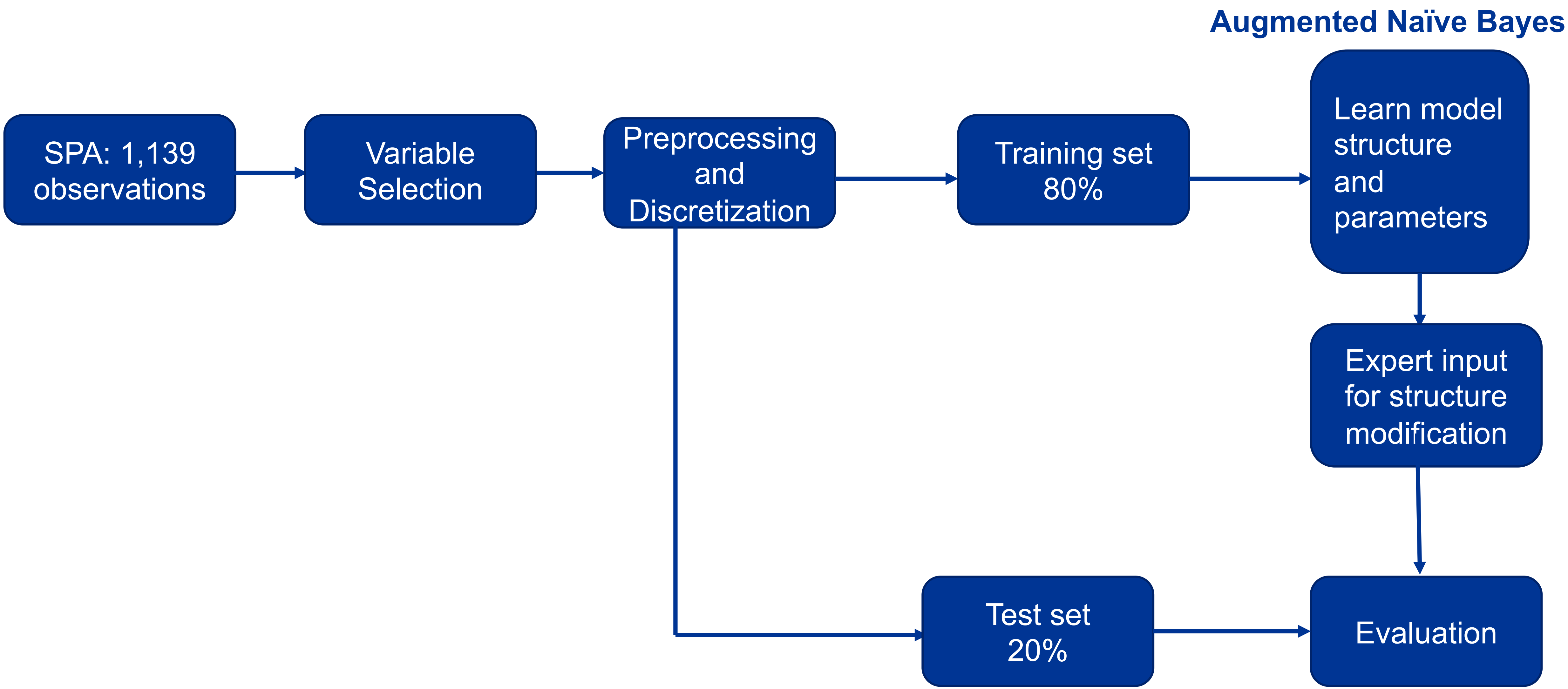
Methods – Bayesian Model I (manual)

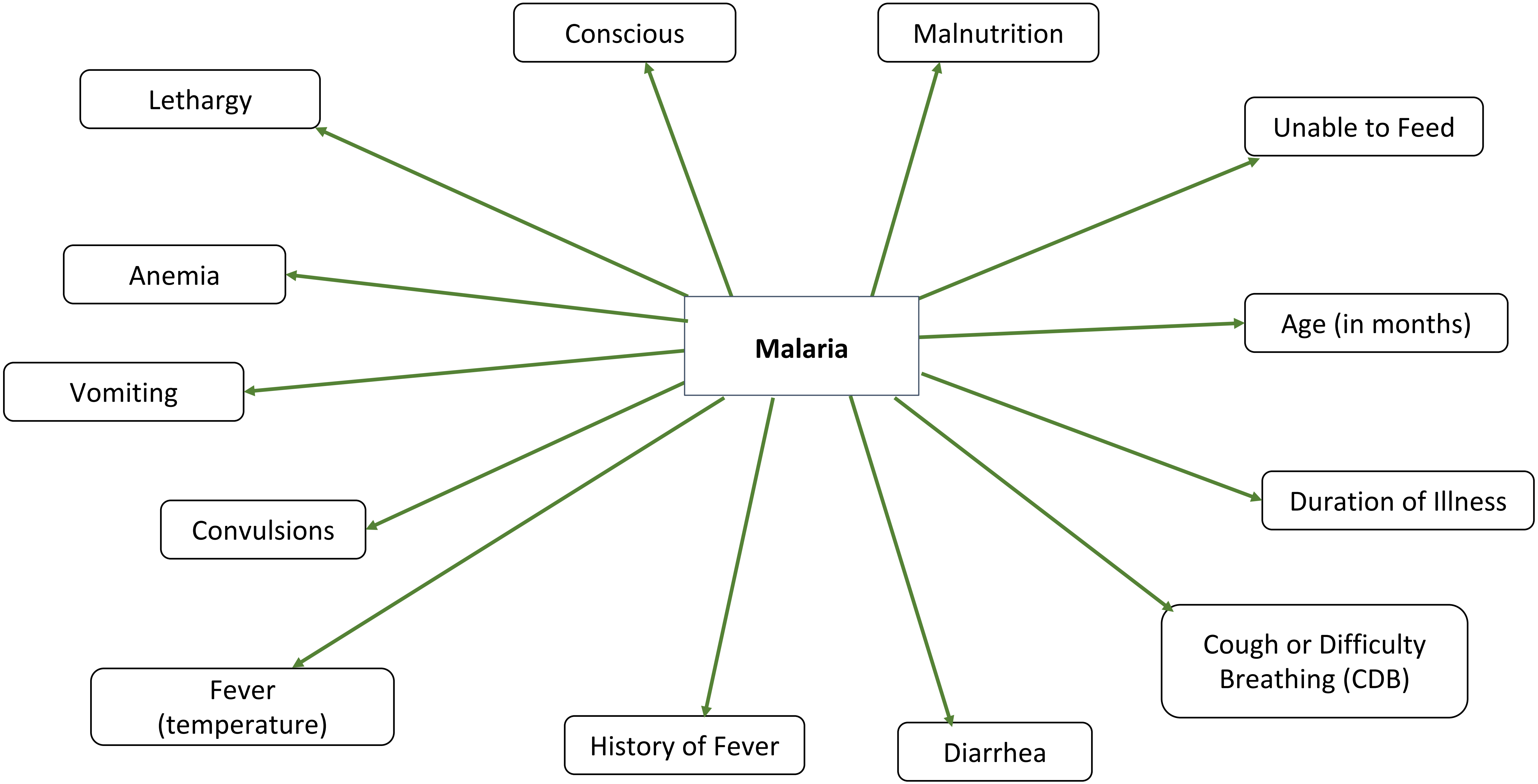


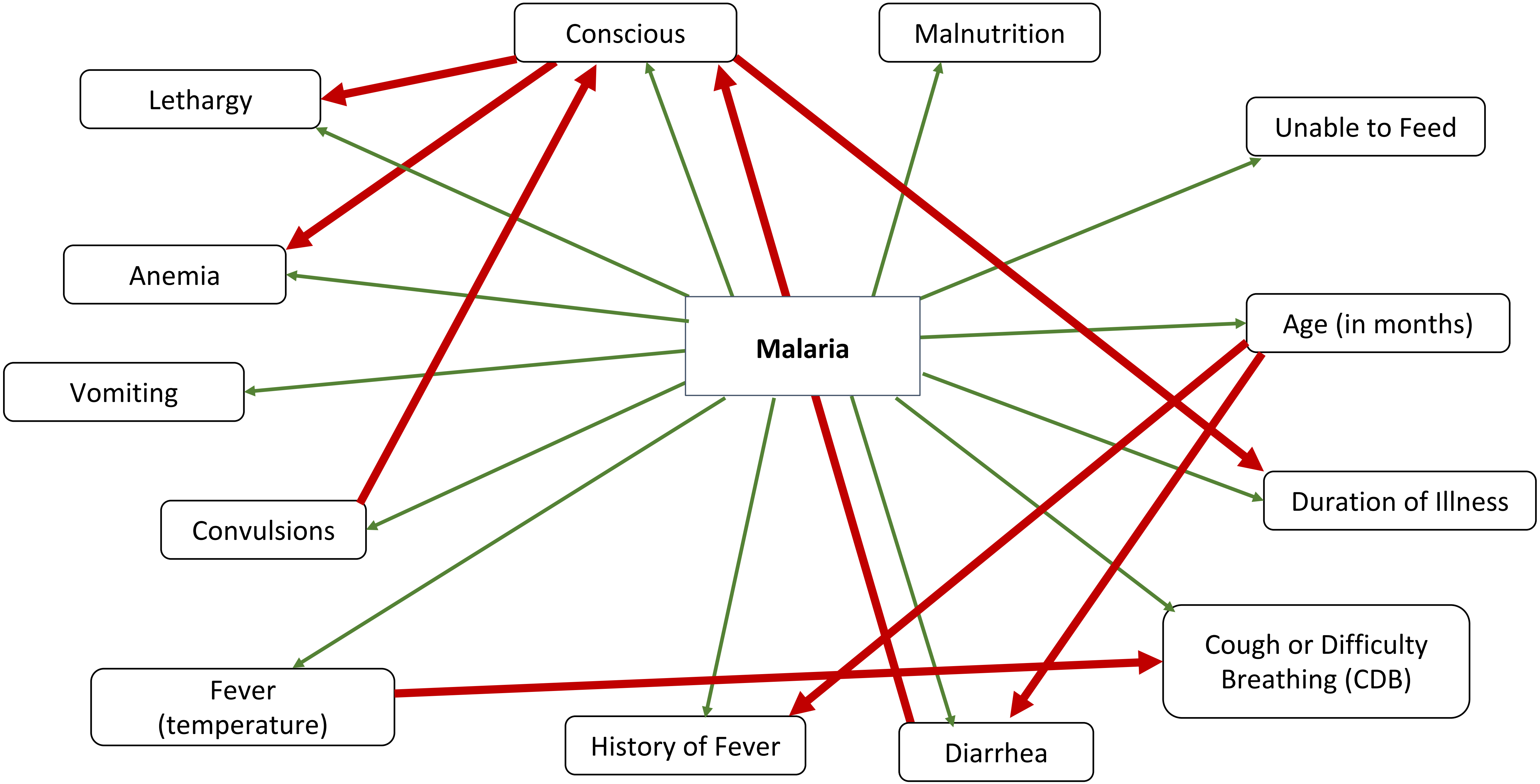


Assumption: Class independence

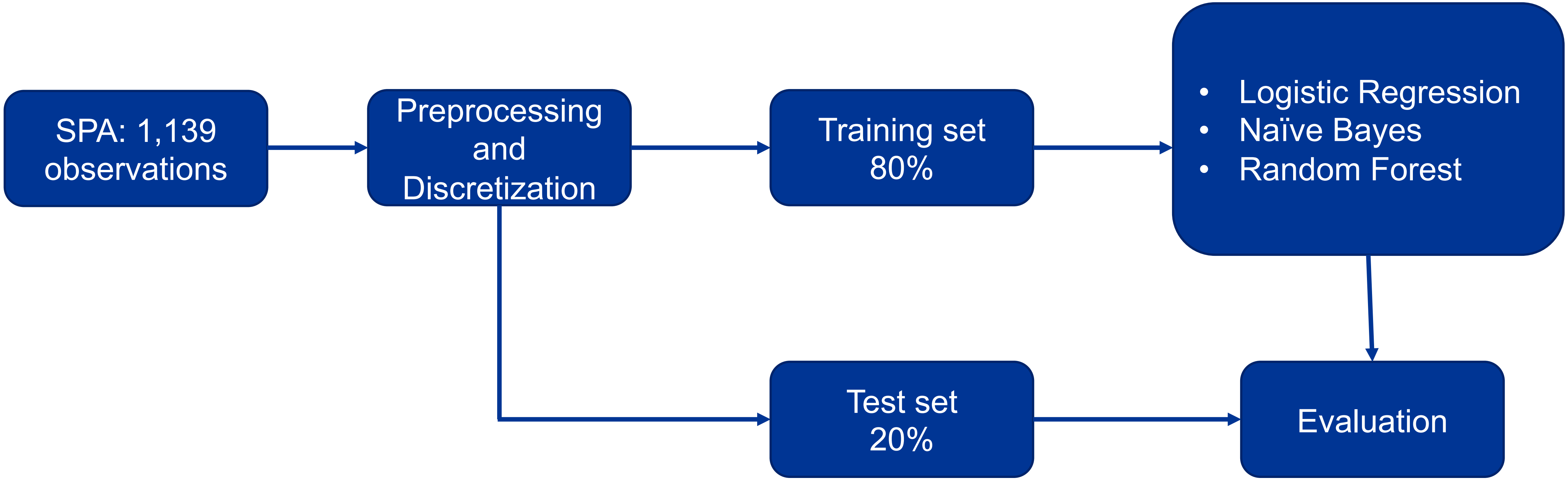
Methods – Bayesian Model II (Hybrid)



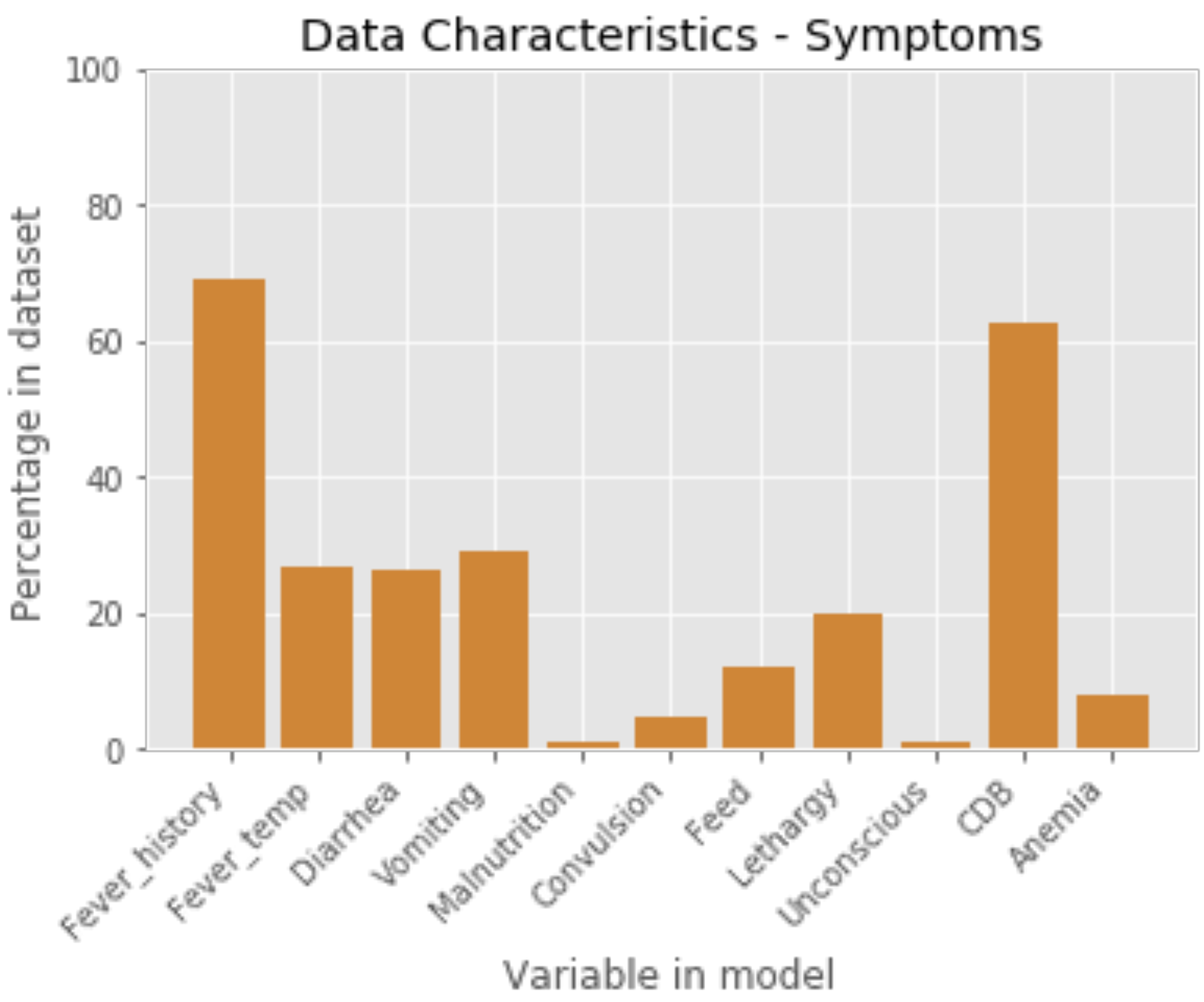

Malaria
Conscious
Malnutrition
Lethargy
Unable to Feed
Anemia
Age (in months)
Vomiting
Convulsions
Duration of Illness
Cough or Difficulty Breathing (CDB)
Fever (temperature)
History of Fever
Diarrhea


Conscious
Malnutrition
Lethargy
Unable to Feed
Anemia
Malaria
Age (in months)
Vomiting
Convulsions
Duration of Illness
Cough or Difficulty Breathing (CDB)
Fever (temperature)
History of Fever
Diarrhea

Methods - Machine Learning Pipeline



Results



Data Characteristics - Symptoms
100
80
60
40
20
0
Percentage in dataset
Fever_history
Fever_temp
Diarrhea
Vomiting
Malnutrition
Convulsion
Feed
Lethargy
Unconscious
CDB
Anemia
Variable in model

Results

Classifier	AUC	Accuracy	Precision	Sensitivity	F1	Specificity
Manual Bayesian Model	0.581	0.636	0.500	0.325	0.394	0.814
Hybrid Bayesian 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

Sensitivity (recall): proportion of positive malaria cases classified correctly
Specificity: proportion of negative malaria cases classified correctly

Results

Contingency table for classification of test set with manual Bayesian model

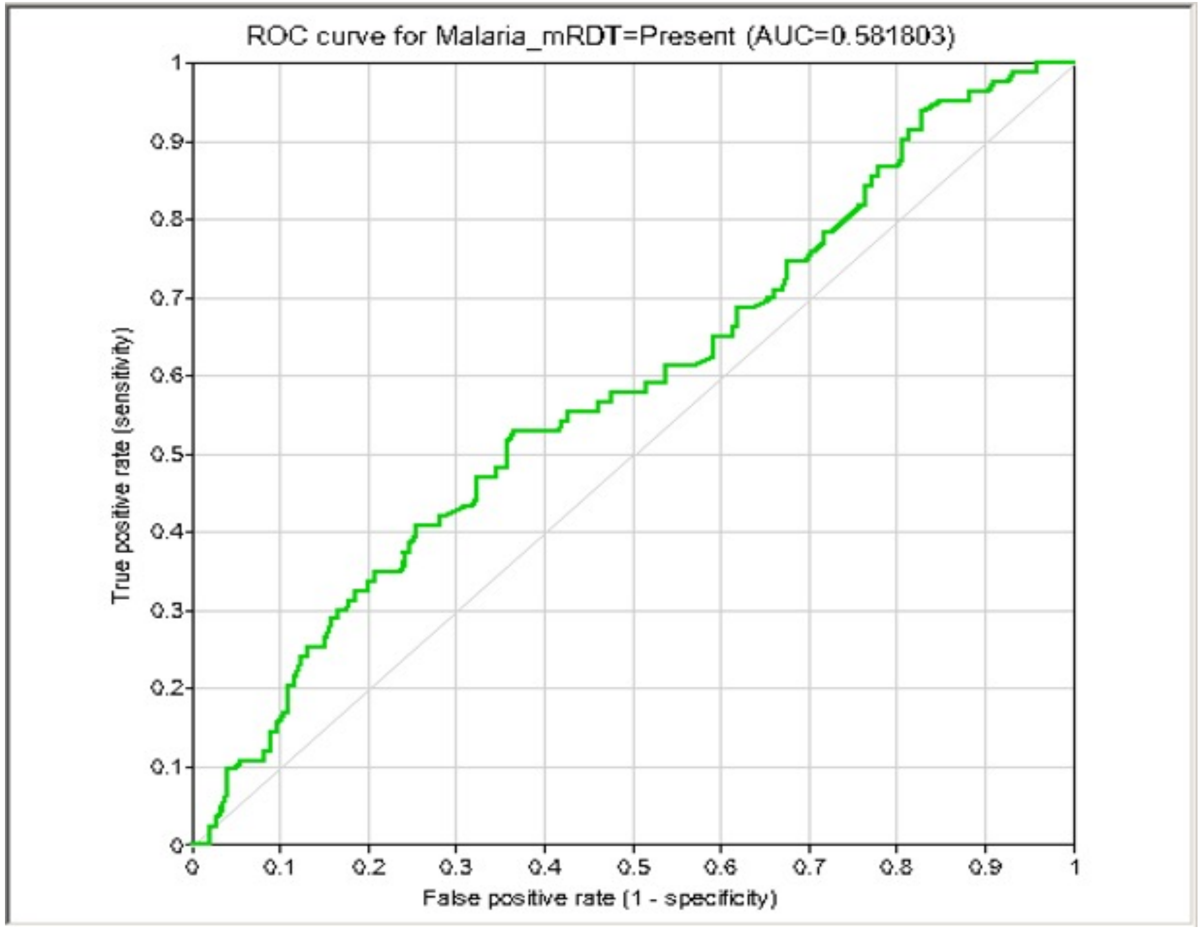
	Malaria Present	Malaria Absent
Predicted Present	27	27
Predicted Absent	56	118
Total	83	145

Contingency table for classification of test set with hybrid Bayesian model

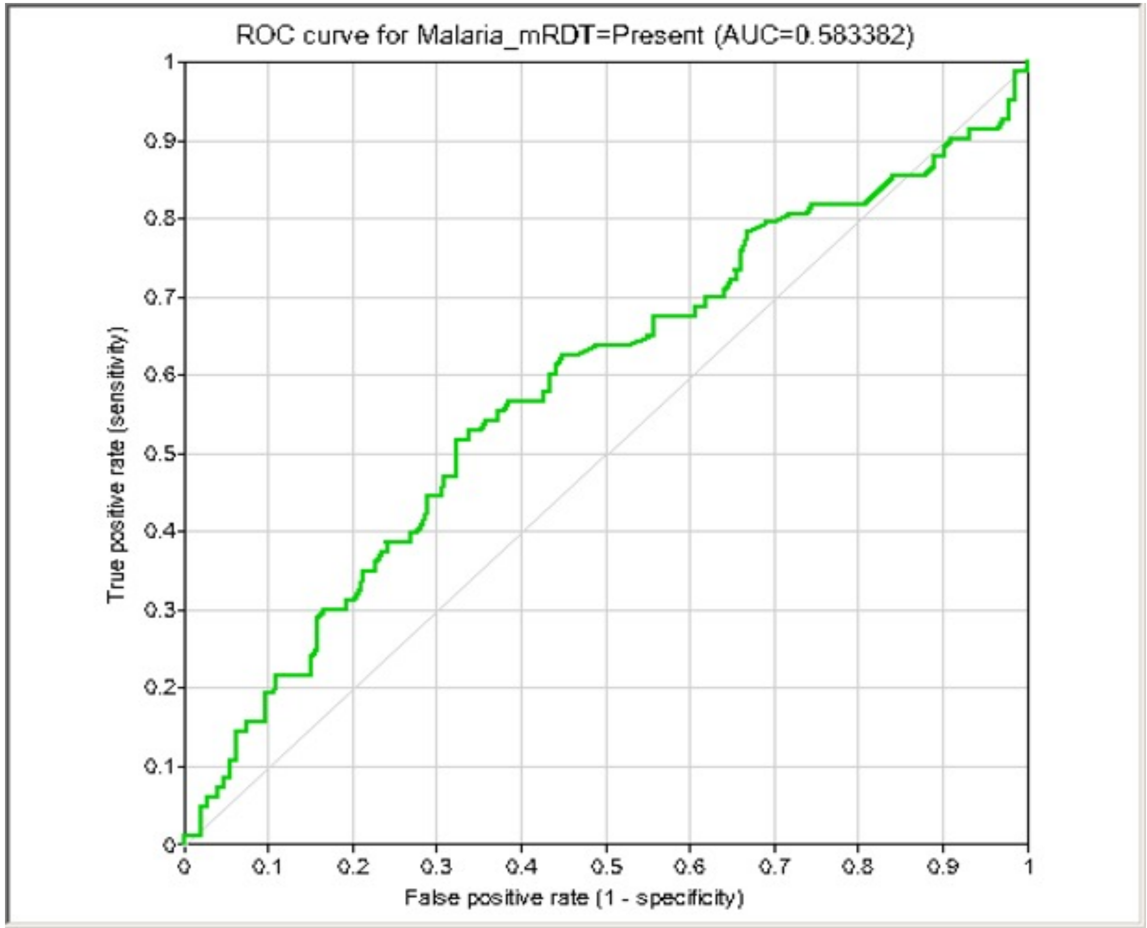
	Malaria Present	Malaria Absent
Predicted Present	20	22
Predicted Absent	63	123
Total	83	145

Results

Manual Bayesian Model



Hybrid Bayesian Model



Results

Performance

Interpretability

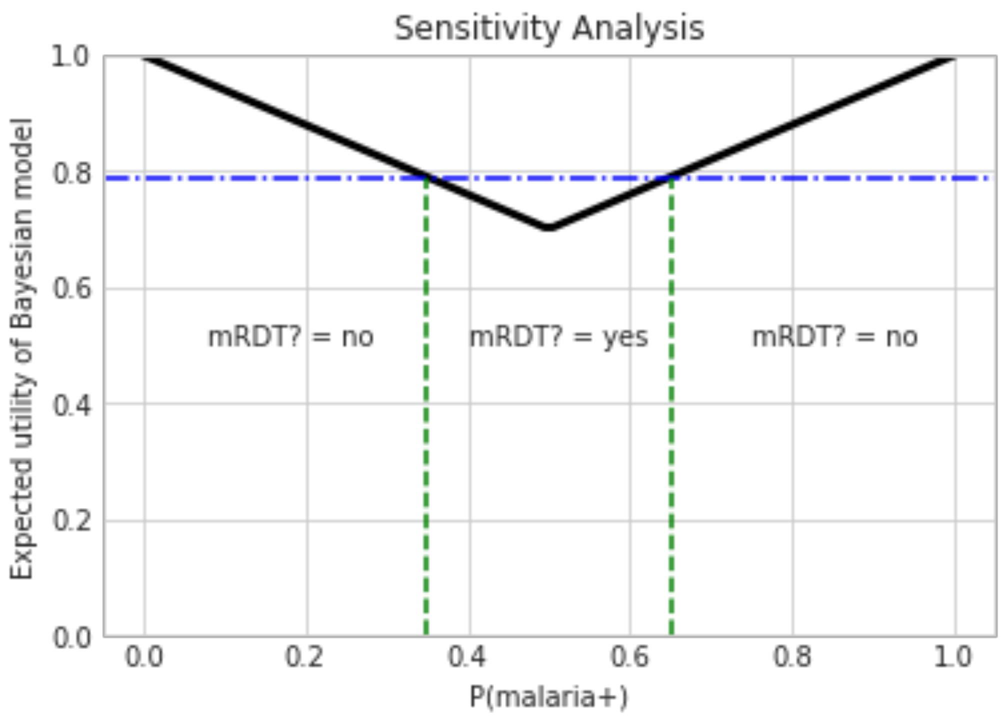
Explainability

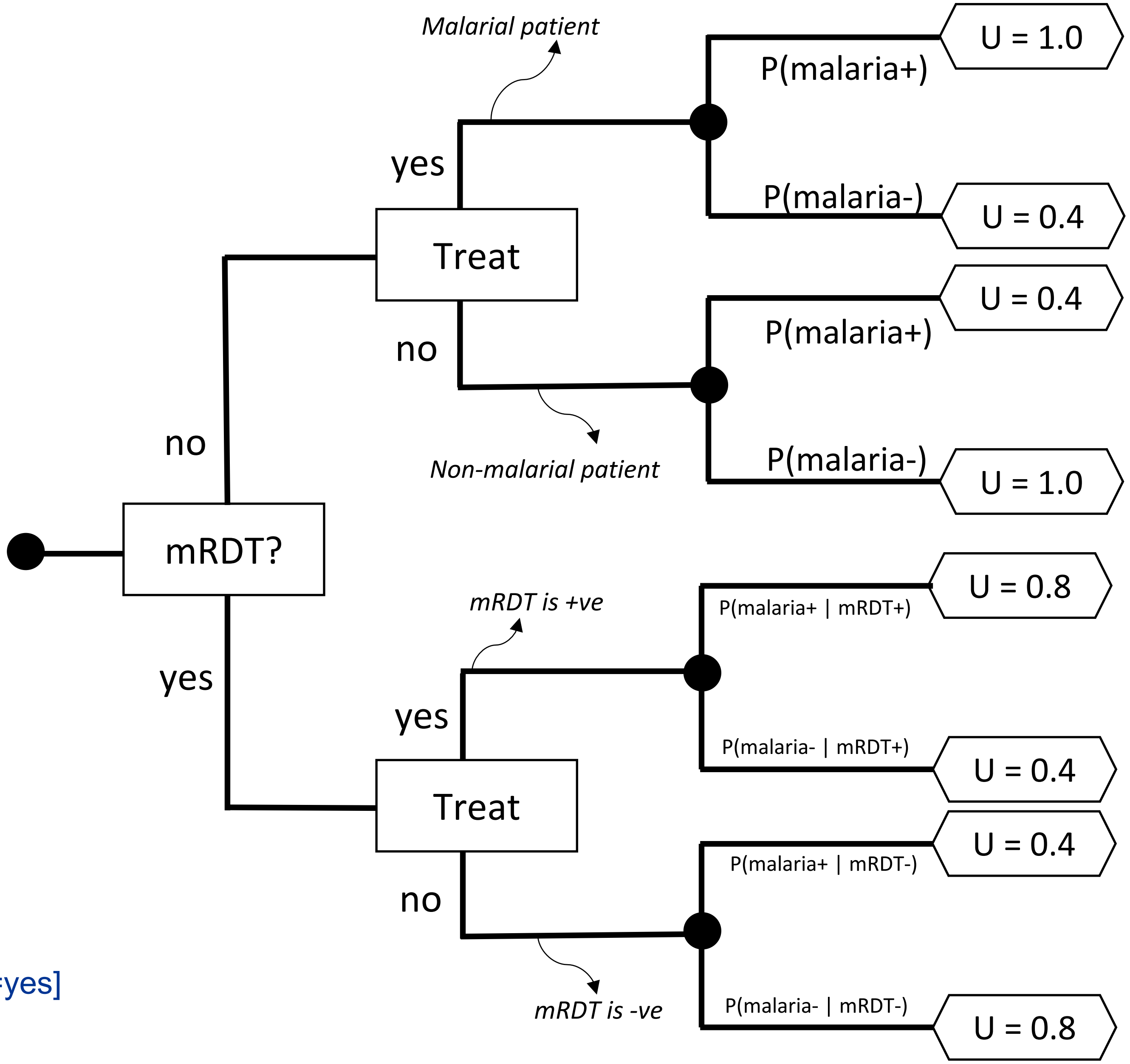
Modifiability

Both models give:

- Automated data collection
- Iterative improvement
- Diagnosis under uncertainty
- Sustainable resource use
- Better alternative to presumptive treatment

Results





Line at 0.79 representing value of expected utility of [mRDT=yes]

Results

History of Fever	Malaria Present	Malaria Absent	Total
Present	313	476	789
Absent	78	199	277
Unknown	24	49	73
Total	415	724	

History of Fever

- Positive Predictive Value = 39.6%
- Negative Predictive Value = 71.8%

Conclusion

- First attempt at Bayesian model for diagnosis in low- and middle-income countries using dataset

- Performance shows potential as an efficient diagnostic tool and includes relevant diagnostic factors in addition to current practice and guidelines

- Decision framework to implement in typical health center in Malawi as mHealth application

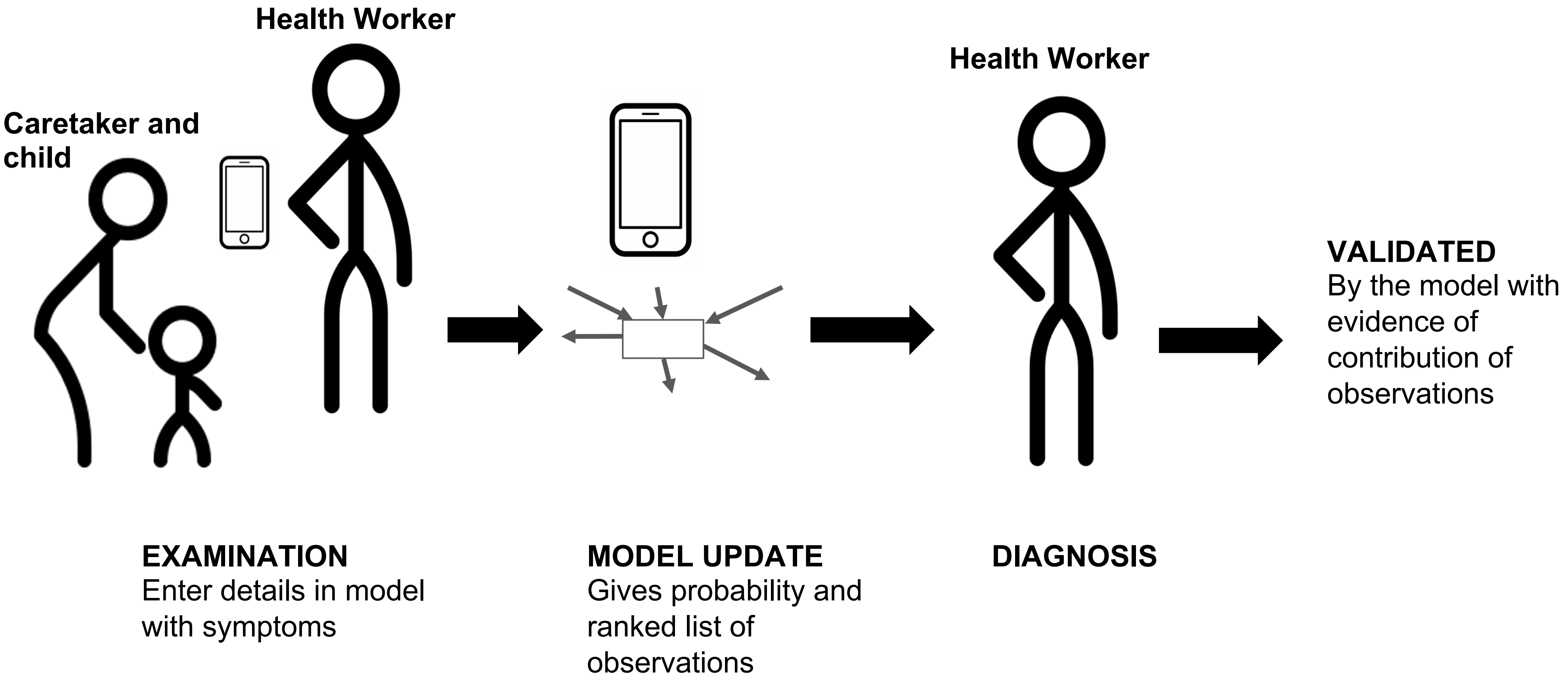
- Approach is generalizable to other diseases and other countries, given appropriate datasets

Limitations

- Only subset with gold standard diagnosis
- Choice of variables
- Other factors: HIV status, immunization, prior malarial infection etc
- Models require external validation, preferably on-site
- Room for improvement in model performances (AUC)

Innovation toward a holistic diagnostic system

Can a *data-driven approach* to diagnosis of childhood illnesses address the challenges faced in *health centers in low-resource countries*?



https://www.sccpre.cat/show/hixRo_mobile-phone-icon-vector-png-black-and-white/

Thank you! Questions?

Thank you for collaboration and guidance:

Dr. Shyam Visweswaran, University of Pittsburgh
Dr. Gerald Douglas, University of Pittsburgh
Dr. Gregory Cooper, University of Pittsburgh
Dr. Marian Michaels, UPMC Children’s Hospital of Pittsburgh
Dr. Harry Hochheiser, University of Pittsburgh
Dr. Marek Druzdzel, Bialystok University of Technology, Poland
Rashid Deula and Global Health Informatics Institute, Malawi
All Malawi clinical workers and staff for their help