In search of a novel scoring system for Tuberculous Meningitis: a Bayesian Latent Class Approach

Dr. Trinh Dong Huu Khanh1,2\*, Dr. Joseph Donovan1,2,3, Dr. Guy Thwaites1,2, Dr. Ronald Geskus1,2

2021-10-07

1. Oxford University Clinical Research Unit, Centre for Tropical Medicine, Ho Chi Minh City, Vietnam

2. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, United Kingdom

3. London School of Hygiene and Tropical Medicine, London, United Kingdom

\*Corresponding author

# Introduction

Tuberculous meningitis (TBM) is the most severe form of tuberculosis. Diagnosing TBM is notoriously challenging, with microbiological confirmation requiring identification of *Mycobacterium tuberculosis* in paucibacillary cerebrospinal fluid (CSF). In addition to widely used confirmatory methods of CSF testing for *M. tuberculosis* such as Ziehl-Neelsen (ZN) staining, GeneXpert MTB/RIF (Xpert), and mycobacterial culture, additional parameters may increase the likelihood of a diagnosis of TBM. Such parameters are illustrated in the uniform case definition for TBM[ref Marais 2010] where, in the absence of positive microbiological tests, an increased certainty of TBM is assigned in the presence of particular clinical, CSF, and imaging findings, or with evidence of non-neurological *M. tuberculosis*.

*Depending on what goes in the table, more text can be written here*

Estimated sensitivity and specificity of confirmation tests based on the TBM case definition (Marais et al. 2010) are listed in table (Nhu et al. 2013).

Table : Prior knowledge of sensitivity and specificity for TBM confirmation tests.

| **Test** | **Sensitivity** | **Specificity** |
| --- | --- | --- |
| ZN Smear | 100% | 78.6% (71.9%, 84.3%) |
| MGIT | 100% | 66.5% (59.1%, 73.3%) |
| Xpert | 99.5% (97.2 -100%) | 59.3% (51.8%, 66.5%) |
| **Numbers are in format of estimated (95% Confidence Interval)** | | |

# Methods

## Study design and population

Data used for this analysis were extracted from an ongoing observational study of brain infection conducted at the Hospital for Tropical Diseases (HTD), Ho Chi Minh City, Vietnam; a large centre providing secondary and tertiary treatment for a wide range of tropical infections (Thwaites et al. 2002). Participants included in this analysis were enrolled between August 2017 and January 2021.

Enrolled participants to the prospective observational study were ≥16 years old, with suspected brain infection, admitted to a neuro-infection ward, and undergoing lumbar puncture at baseline as a routine diagnostic procedure. Patients were ineligible for enrolment if performing a lumbar puncture was contraindicated, or if informed consent to join the study was not given (by the patient, or by a relative if the patient lacked capacity to consent). This study received ethical approvals from HTD and Oxford Tropical Research Ethics Committee. A subset of participants were included in a randomised diagnostic comparison study of Xpert and GeneXpert MTB/RIF Ultra (Xpert Ultra) for the diagnosis of TBM.(Donovan et al. 2020)

## Data collection and testing procedures

Demographic data and relevant medical history were recorded. HIV testing was performed on a case-by-case basis by the treating clinician. Patients with either a known HIV positive status, or a positive HIV test result, were considered HIV positive. Patients underwent clinical examination, routine blood testing, and analysis of sputum for *M. tuberculosis* where indicated. CSF was obtained by lumbar puncture, with CSF white blood cells (WBC), WBC differential, CSF protein, CSF glucose (with paired blood glucose), and CSF lactate, all routinely measured. Confirmatory mycobacterial testing was performed if TBM was suspected by the treating clinician. At least 6mls CSF was used for mycobacterial testing; however, on occasions where less CSF was available mycobacterial testing was still performed. Confirmatory tests for TBM performed on CSF and used for this analysis were ZN staining and smear, Xpert or Xpert Ultra, and mycobacterial culture using Mycobacteria Growth Indicator Tube (MGIT). Given Xpert and Xpert Ultra were considered diagnostically comparable in a recent comparative study from this centre using a subgroup of these data,[Donovan et al. 2020) Xpert and Xpert Ultra were considered the same in this analysis.

Methods of CSF processing have been described elsewhere (Nhu et al, Donovan et al. 2020), however, briefly; CSF samples were centrifuged at 3000g for 15 minutes, and most of the CSF supernatant was removed. The CSF deposit was resuspended in 500µL of remaining supernatant, with this resuspended pellet then used for ZN smear (100µL), MGIT (200µL), and either Xpert or Xpert Ultra (200 µL).

All patients received appropriate anti-TB chemotherapy regimens according to national and local guidelines, depending on the diagnosis, without any interference from the study. At the time of discharge or death, all patients received a final diagnosis. If at least one of ZN Smear, Xpert, Xpert Ultra, or MGIT, was positive at anything time during the follow-up, the patient would be considered confirmed TBM; otherwise, if TBM was clinically suspected and treated, with confirmatory microbiological tests negative, the patient would be considered suspected TBM. If the patient recovered without anti-TB chemotherapy they would be reassigned to not TBM.

In the instance of repeat CSF sampling (performed based on clinical need), only the first samples at baseline were used to evaluate test sensitivities. Results of haematological and biochemical parameters, and of ZN Smear, Xpert, Xpert Ultra, and MGIT, were taken only from the first baseline CSF sample (in addition to the paired blood glucose). In general, the choice of predictors is based on the TBM definition scoring system (Marais et al. 2010). CSF eosinophil count was additionally included as it is a strong bio-marker for eosinophilic meningitis, a condition usually caused by parasitic helminths; whilst CSF erythrocyte count was added as a marker for traumatic lumbar puncture.

## Statistical Model

All data preparation, cleaning, and processing were performed on statistical package , version 4.1.0 (R Core Team 2021). The model was developed on the probabilistic language via the interface , version 2.27 (Stan Development Team 2021a). Plotting was done using package (Gabry et al. 2019), (Defazio and Campbell 2020), and (Yan 2021). Some other packages used include: (Chang 2020), (Harrell Jr 2021).

In the model, three aforementioned TBM confirmation tests (ZN Smear, MGIT, and Xpert) were used as manifest variables. Linear predictors for the latent class prevalence were **history and demographic information** (HIV status, Age, Past noticeable TB contact), **clinical signs and symptoms** (Days from onset to admission, Systemic symptoms suggestive of tuberculosis, Focal neurological deficit, Cranial nerve palsy, Glasgow Coma Score - GCS), **Imaging chest data** (Pulmonary TB, Miliary TB), **CSF criteria (**lymphocyte count, neutrophil count, oeosinophil count, erythrocyte count, glucose - with corresponding blood glucose, protein, lactate), and **Cryptococcus Antigen/Indian Ink** in combination**.**

### Data pre-processing

Most continuous variables were transformed to logarithmic scale and subsequently centred. In all base models, no scaling was performed as they could potentially imputed. . Glasgow Coma Score and its components were not transformed, rather we translated them to **Loss of GCS** (LoGCS); so that a would be equivalent to , while would be translated to .

We also made an adjustment for Past TB contact, in which we hypothesise that by changing the question into “Past *noticeable* contact with TB patients within the past recent year”, we would be able to imply all “Unknown” answer as a “No”. In the main study, no brain imaging was taken, hence not included.

Binary variables were encoded into 0 and 1 and not centred.

### Latent class regression model

We created two-level hierarchical models each of which combines:e

* **Prevalence model**: A logistic regression model trying to estimate the prevalence of TBM amongst the study population
* **Latent class analysis**: estimating the probabilities of having positive results from three tests in each class. Similar to previous applications (Qu, Tan, and Kutner 1996; Hadgu and Qu 2002; Schumacher et al. 2016), we also corrected for individual bacillary burden and procedural variance between samples. Latent bacillary burden is regressed by individual Gaussian random variables which captures noisy fluctuation of test results, and fixed effects coming from covariates. We hypothesise that even if two patients were in the same TBM positive class, one who had lower bacillary burden would be less likely to be tested positive. This lifted the local independence assumptions of vanilla Latent Class Analysis .

Table : Contribution of different variables in each model

| **Predictor** | **Risk of TBM** | **Bacillary Burden** |
| --- | --- | --- |
| Age | + | 0 |
| HIV Status | ++ | + |
| Past TB contact | + | 0 |
| TB-suggested symptoms | + | 0 |
| Local motor deficit | + | 0 |
| Cranial nerve palsy | + | 0 |
| Days from onset | + | 0 |
| PTB/X-Ray | + | 0 |
| MTB/X-Ray | ++ | 0 |
| GCS | + | 0 |
| Cryptococcus Antigen/Indian Ink | - | 0 |
| Blood Glucose | - | 0 |
| CSF Glucose | - | ? |
| CSF Lymphocyte Count | + | ? |
| CSF Neutrophil Count | +/0/- | ? |
| CSF Protein | + | ? |
| CSF Lactate | + | ? |
| CSF Oeosinophil Count | - | 0 |
| CSF RBC Count | - | 0 |
| **"++" predictors were strongly believed to be positive risk factors** | | |
| **"0" predictors were not included in any model** | | |
| **"?" predictors were included and excluded in separated models.** | | |

The inclusion of covariates was performed according to prior knowledge of potential predictors that share association with the infection risk and test sensitivity, summarised in table ). Due to the lack of understanding about the mechanism of the disease and the main interest of the analysis, we did not investigate the causal relationship between different covariates.

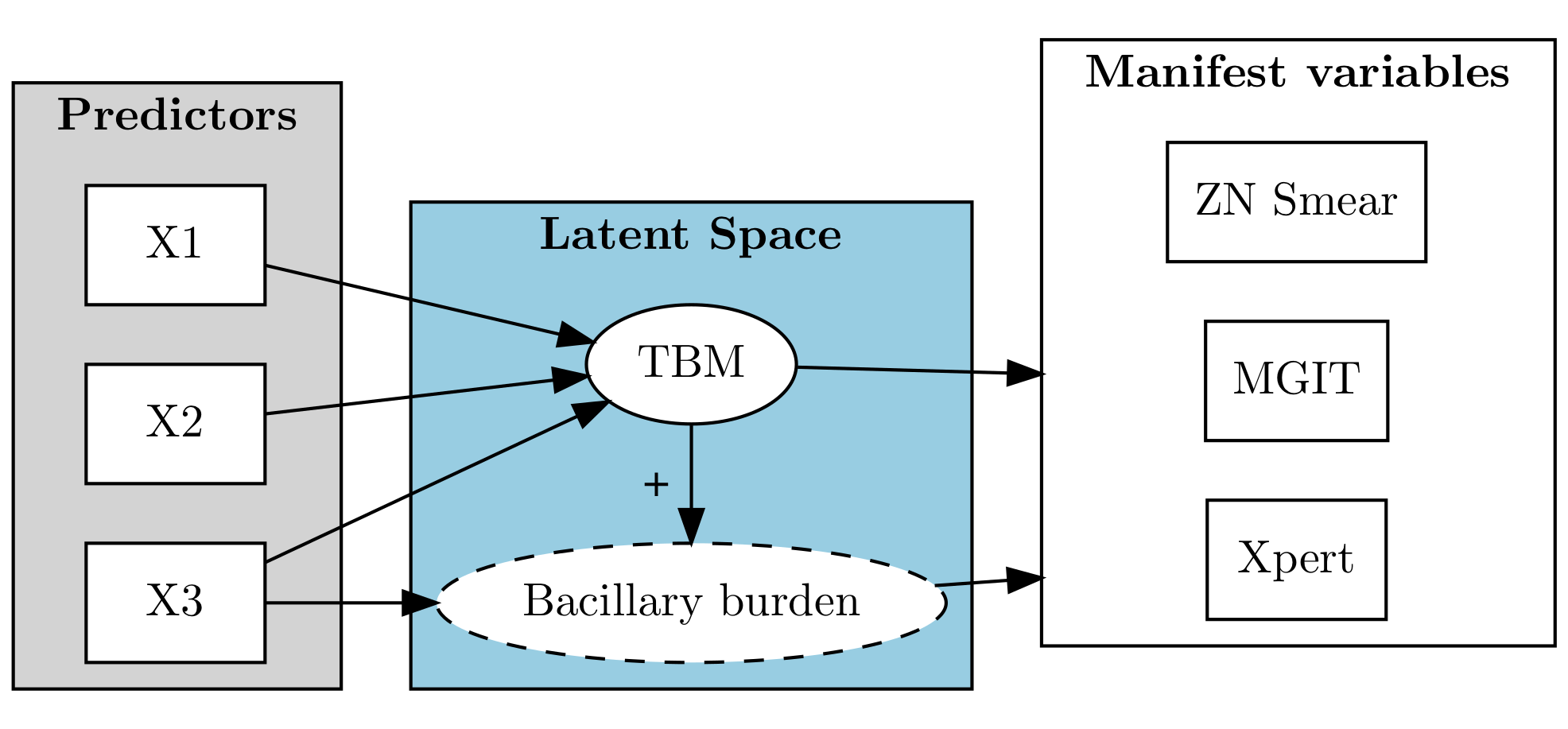


Figure : Model skeleton. Bacillary Burden is only available in model 2+

The skeleton for all models is shown in figure . We use a stepwise approach where we incrementally added up more flexibility and lifted more constraints. We also include several extensions to explore many possibilities than can improve performance. In the main analysis, only CSF neutrophil count had quadratic effect as suggested by the TBM case definition (Marais et al. 2010). A summary of all architectures an extensions are shown in table , whereas technical formulations are detailed in appendix .

Table : Model architectures and extensions

| **Model** | **Base definition (compared to the lower one)** | **Non-linear effects in Prevalence model** | **Covariates in bacillary burden** |
| --- | --- | --- | --- |
| 1 | No baccillary burden; everyone in the same class has equal risk of tested positive | - | - |
| 2 | Added individual bacillary burden; impacts of bacillary burden on test results are the same | CSF Neutrophil count | HIV status   CSF Glucose   CSF Lymphocyte count   CSF Neutrophil count   CSF Protein   CSF Lactate |
| 3 | Impacts of bacillary burden are different for different tests | CSF Neutrophil count | HIV status   CSF Glucose   CSF Lymphocyte count   CSF Neutrophil count   CSF Protein   CSF Lactate |
| 4 | Added technical fluctuation as a second random effect; fixed effects only contributes to bacillary burden | CSF Neutrophil count | HIV status   CSF Glucose   CSF Lymphocyte count   CSF Neutrophil count   CSF Protein   CSF Lactate |
| 5 | Added fixed effects for technical fluctuation | CSF Neutrophil count | HIV status   CSF Glucose   CSF Lymphocyte count   CSF Neutrophil count   CSF Protein   CSF Lactate |

### Missing data handling

#### Manifest variables

By design, most patients with very high chance of and/or evidently diagnosed with different diseases were not tested with TBM confirmation assays (namely ZN Smear, MGIT, and Xpert), unless there were excessive amount of CSF samples. However, as the tests’ sensitivities were all firmly believed to be almost perfect (Nhu et al. 2013), we assumed that patients who had no TBM confirmation tests are all negative. Apart from one premature death, we have yet to find any patients left un-tested and un-diagnosed.

#### Predictors

In this analysis, missing predictors’ values were assumed to be Missing At Random (MAR) and imputed within sampling programmes, together with the main model. Composite predictor variables, such as *TB-suggested symptoms* and *Glasgow Coma Score (GCS)* were imputed by compartments; while potentially correlated variables were grouped and imputed together. Due to Stan not supporting Multivariate Logistic Regression, in favour of method consistency, all binary predictors were imputed using (Multivariate) Probit models. Continuous variables are imputed using Multivariate Linear Regression. HIV status was included in most imputation model as predictors.

As HIV tests were not mandatory in the main study, their chance of missingness were mostly dependent on whether or not they were to enrol in a TBM study. Hence, it is safely to assume that HIV status are Missing at Random (MAR). Accordingly, we imputed HIV using probit regression, corrected for Blood Lymphocyte and Neutrophil counts.

Table : Rationale and method of missing values handling

| **Variable** | **N  missing** | **Expected Reason of Missingness** | **Mechanism** | **Handling method** |
| --- | --- | --- | --- | --- |
| ZN Smear | 311 | Not suspected TBM | MNAR | Set = 0 |
| MGIT | 314 | Not suspected TBM | MNAR | Set = 0 |
| Xpert | 305 | Not suspected TBM | MNAR | Set = 0 |
| HIV Status | 215 | Test not mandatory | MAR/MCAR | Imputation |
| TB-suggested symptoms | 273 | Unmeasured / Unnoticed / Unconscious | MAR/MNAR | Imputation |
| Local neuro-deficit | 17 | Unconscious | MAR/MNAR | Imputation |
| Glasgow Coma Score | 21 | Unconscious (GCSV) / Blinded (GCSE) / Unmeasured | MAR | Imputation |
| Age | 1 | Input error | MCAR/MAR | Imputation |
| Illness days | 29 | Patients forget / Unconscious | MAR | Imputation |
| Blood Lymphocyte | 1 | Unmeasured (premature death) | MAR | Imputation |
| Blood Neutrophil | 1 | Unmeasured (premature death) | MAR | Imputation |
| Blood Glucose | 7 | Most likely input error / Unmeasured (premature death) | MAR/MCAR | Imputation |
| CSF glucose | 1 | Unmeasured (premature death) | MAR/MCAR | Imputation |
| CSF lymphocyte count1 | 2 | Very low or zero / Input error / Unmeasured (premature death) | MNAR/MAR | Manually set/Imputation |
| CSF neutrophil count1 | 21 | Very low or zero / Input error / Unmeasured (premature death) | MNAR/MAR | Manually set/Imputation |
| CSF protein | 2 | Data input error / Unmeasured(premature death) | MAR/MCAR | Imputation |
| CSF lactate | 1 | Unmeasured(premature death) | MAR/MCAR | Imputation |
| CSF oseosinophil count1 | 1 | Zero cell count | MNAR | Set = 0 |
| **1CSF Cell counts are calculated as CSF white-cell count \* Pct of Cell Type / 100; if very low, then either lymphocytes or neutrophils had values, the other were left missing; Missing oeosinophil count implies 0 cell.** | | | | |

We summaries our rationales and corresponding handling strategies in table . In the case where missing values were imputed, figure depicts how those were sampled, together with their potential hyper-predictors. Note that due to Hamiltonian Monte Carlo (HMC)’s limitation, imputed LoGCS were treated as-is in sampling process, but were rounded when estimating model performance.

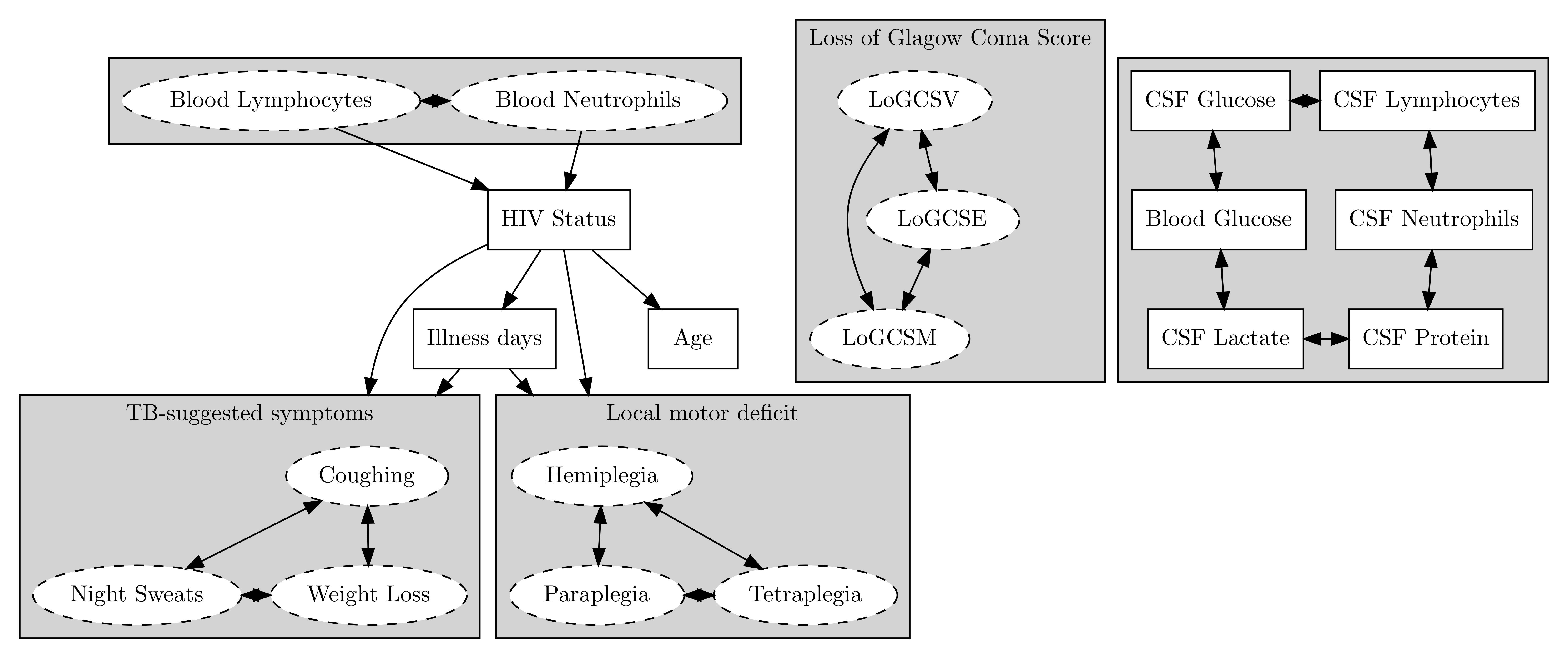


Figure : Imputation strategy for predictors. Variables in rectangular solid boxes were used in the model, in oval dashed were either compartments of composite ones or only contributed in the imputation model and were not included in the main model. Clustered covariables were imputed together in a multivariate regression. Arrows demonstrate a predictor-response correlation.

### Prior choices

Following Gelman’s recommendations (Stan Development Team n.d.), we chose as prior distributions for all intercepts and coefficients in the imputation model, except for LoGCS (E, V, and M) as their supports were constrained to [0,1]; in the latter case, was chosen instead for the means, and for the standard deviation.

Cholesky decomposition of the covariance matrix was sampled from a *LKJ Correlation Cholesky* prior with (Stan Development Team 2021b):

In every linear sub-model of main one, we used for the intercept on which we impose our weak expectation that its absolute value cannot be higher than 10 (Boonstra, Barbaro, and Sen 2019). For the covariates, we considered several sets of prior representing a spectrum of penalties bestowed upon the model (van Erp, Oberski, and Mulder 2019):

* Weakly informative prior:
* Ridge-equivalent prior:
* LASSO-equivalent prior:

Coefficient for known strongly positive risk factors (marked **++** in table were specially imposed a positive Half Normal distribution. Individual random effects representing unmeasured bacillary burden was sampled from a distribution.

For manifest variables, we used highly informative priors for *specificity (Spc)* basing on previous study (Nhu et al. 2013) and weakly informative priors for *sensitivity (Sen)* on the logit scale. These choices were visualised on figure .

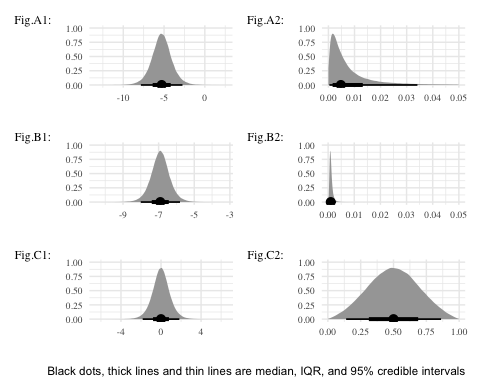


Figure : Density plots for different priors. A: Logistic(logit(.005),.7), B: Logistic(logit(.001),.3), C: Logistic(0,.5). 1: Logit scale, 2: Linear scale.

### Model performances

Models performances were defined and compared by three metrics:

* Expected log point-wise predictive density (elpd) of hold-out observations (Vehtari, Gelman, and Gabry 2016): Models with higher elpd arw supposed to have more predictive values for untrained observations.
* Visualised model calibration between estimated predicted and observed probabilities of positive confirmation tests (Harrell Jr 2021), using a non-parametric *loess* fit. We also used the diagnosis at discharge as a pseudo-gold standard to visualise the calibration of TBM prevalence. A model with good calibration would correctly estimate the observed values.
* Receiver Operating Characteristic (ROC) curve and corresponding Areas Under the Curve (AUC) . Confidence interval for AUC was estimated by a 2000-time bootstrapping process ***<I might change this to a fully bayesian estimation. shall I? >***. A model with good discriminative value would be better to distinguish between two class.
* Additionally, we also visualised class-wise predicted probability density plots to visualise how much separable the classes are based on the models. This demonstrate how predicted probabilities distributed between two classes.

All metrics are based on 5 repetitions of 20-fold cross validated datasets (i.e. 100 fits, as suggested by Harrell (Harrell Jr 2021)). Accordingly, we selected three best-performed models and re-estimated their parameters using the full dataset.

We also test for local independence assumption.

### Exploratory and sensitivity analysis

#### Complete-case analysis and imputation under MNAR assumption

To check the level of impact from our imputation method, we did a complete-case analysis and a “missing-as-a-category” analysis in which we considered missing values as a level for binary variables. Suspected MNAR variables as listed in table were also tested for MNAR where we randomly allocated values based on experts’ opinions.

Under MNAR assumption, we did a pattern-mixture method(Mason et al. 2017; White et al. 2007), where we inquired prediction offsets . represents the difference between unobserved part and observed part of each variables, after correction for all hyper-predictors listed in . The offsets were collected from interviews with experts working at the Viet Anh Ward at HTD, where they were supposed to provide an estimation and 95% confidence intervals (95% CI), based on which s were then sampled from.

*where* *and* *are estimation and* *of 95% CI provided by interviewed experts*.

The main model underwent a relief of strict priors for test specificities in @ref(eq:priors-response), the lifted model used the same prior for all three tests which cover specificity from at least 90%:

*where* for ZN Smear and MGIT, for Xpert.

Lastly, as recent studies suggested a sup-optimal specificity of Xpert test on CSF samples(Nhu et al. 2013; Chen et al. 2020), our assumptions made in table might not completely valid. To tackle this, we considered a MAR scenario, where observation chance of confirmation tests depend on the unknown TBM status and locally independent to the value of confirmation tests. The observation status was then included in the model as a separated manifest variables **(to Ronald: should I left this in the sensitivity or include this in the main analysis, for Xpert only or for all three?).** The validity of this method was depicted in a simulation study in .

In this analysis, we expected the chance in which at least one confirmation test was done are 95% for TBM-positive patients, and 50% for TBM-negative, hence led to two conservative priors:

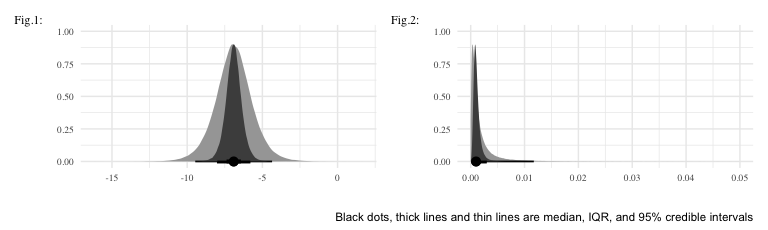


Figure : Lifted prior for ZN Smear and MGIT specificities, compared to the old one (dark grey). 1: logit scale, 2: linear scale

#### Non-linearity and Dimensionality Reduction

As part of the exploratory analysis, we considered two models:

* Non-linearity for all CSF bio-markers: We added quadratic effects for all CSF bio-markers in the prevalence model. LASSO-based variable selection was implemented for this analysis.
* Non-linearity for Glasgow coma scores: As GCS is not a continuous variable but rather an ordinal one de facto, it is possible that there is non-linear correlation between GCS and TBM risk. In this analysis, we employed a quadratic effect for GCS to capture this potential of non-linearity.
* Probabilistic Principal Component Analysis (PPCA): Instead of performing Variable selection, we performed an implementation of Probabilistic Principal Component Analysis for dimensionality reduction, especially amongst potentially collinear predictors. In this analysis, we only implemented this for collinearity-prone CSF bio-markers.
* Test accuracy with respect to CSF volume: By including the volume of sample collected, we can further investigate the effect size of CSF volumes on the sensitivity of each confirmation test.

### Simplified approximation of TBM risk

As the full model requires a plethora of predictors and measurements, it might not be pragmatic in some limited contexts. We hence developed a simplified version of the prevalence sub-model which exclude all laboratorial features. The aim of this exploratory analysis is to provide a decent approximation of TBM risk yet needs only a minimal amount of information.

In this analysis, we took out the posterior probabilities of TBM from the best-performed model, on the logit scale, fed them into subsequent model where the number of predictors were reduced. The error term follows Logistic distribution with .

*s.t.* and is an approximation of .

# Results

Venn diagram for confirmation test results are demonstrated in

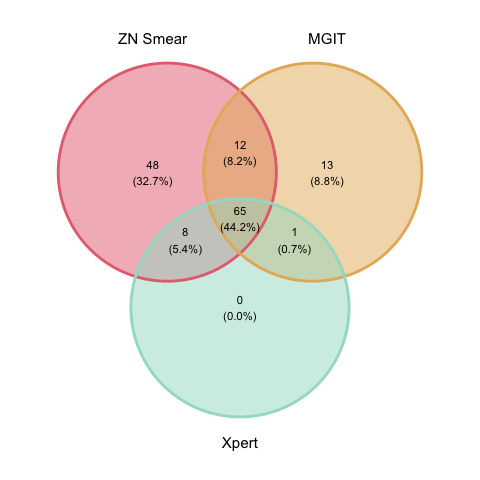


Figure : Venn diagram for ZN Smear, MGIT, and Xpert

Blahblahblah

Table : History and baseline information

|  | **MGIT** | | **Xpert (Ultra) MTB/RIF** | | **ZN Smear** | |
| --- | --- | --- | --- | --- | --- | --- |
| **Characteristic** | **Negative, N = 5961** | **Positive, N = 911** | **Negative, N = 6131** | **Positive, N = 741** | **Negative, N = 5541** | **Positive, N = 1331** |
| Age | 42 (27, 56) | 41 (30, 50) | 42 (27, 56) | 40 (29, 48) | 42 (26, 56) | 42 (29, 53) |
| HIV Positive | 26 (6.7%) | 25 (29%) | 27 (6.7%) | 24 (34%) | 18 (5.1%) | 33 (27%) |
| - Missing | 208 | 6 | 211 | 3 | 202 | 12 |
| Diabetes | 37 (6.2%) | 2 (2.2%) | 39 (6.4%) | 0 (0%) | 35 (6.3%) | 4 (3.0%) |
| Day from onset | 10 (4, 11) | 19 (10, 20) | 10 (4, 11) | 20 (11, 22) | 10 (4, 11) | 16 (8, 17) |
| - Missing | 21 | 8 | 22 | 7 | 19 | 10 |
| TB-suggested symptoms | 82 (23%) | 39 (67%) | 86 (24%) | 35 (71%) | 66 (20%) | 55 (64%) |
| - Missing | 240 | 33 | 248 | 25 | 226 | 47 |
| Past noticeable contact TB within 12 months | 7 (1.2%) | 3 (3.3%) | 6 (1.0%) | 4 (5.4%) | 4 (0.7%) | 6 (4.5%) |
| Focal neurological deficit | 64 (11%) | 20 (22%) | 67 (11%) | 17 (23%) | 59 (11%) | 25 (19%) |
| - Missing | 15 | 2 | 16 | 1 | 13 | 4 |
| Crain nerve palsy | 61 (10%) | 26 (29%) | 64 (10%) | 23 (31%) | 58 (10%) | 29 (22%) |
| Glasgow Coma Score | 12 (10, 14) | 11 (10, 14) | 12 (10, 14) | 11 (9, 14) | 12 (10, 14) | 12 (10, 15) |
| - Missing | 19 | 1 | 19 | 1 | 19 | 1 |
| X-Ray Pulmonary TB | 18 (3.0%) | 19 (21%) | 18 (2.9%) | 19 (26%) | 13 (2.3%) | 24 (18%) |
| X-Ray Miliary TB | 2 (0.3%) | 3 (3.3%) | 2 (0.3%) | 3 (4.1%) | 2 (0.4%) | 3 (2.3%) |
| CSF Lymphocyte Count | 283 (9, 274) | 225 (77, 305) | 284 (10, 289) | 201 (54, 236) | 291 (8, 286) | 210 (65, 271) |
| - Missing | 2 | 0 | 2 | 0 | 2 | 0 |
| CSF Neutrophil Count | 1,108 (3, 201) | 241 (53, 223) | 1,080 (3, 171) | 270 (54, 301) | 1,170 (2, 173) | 256 (31, 231) |
| - Missing | 2 | 0 | 2 | 0 | 2 | 0 |
| CSF Oeosinophil Count | 8 (0, 0) | 0 (0, 0) | 8 (0, 0) | 0 (0, 0) | 8 (0, 0) | 0 (0, 0) |
| - Missing | 1 | 0 | 1 | 0 | 1 | 0 |
| CSF Red blood cell Count | 1,649 (0, 700) | 1,439 (0, 132) | 1,721 (0, 700) | 787 (0, 87) | 1,861 (0, 1,000) | 620 (2, 80) |
| CSF Protein | 1.74 (0.42, 1.97) | 2.39 (1.34, 2.43) | 1.69 (0.43, 1.99) | 3.00 (1.52, 2.35) | 1.65 (0.41, 1.90) | 2.57 (1.21, 2.43) |
| - Missing | 1 | 1 | 1 | 1 | 1 | 1 |
| CSF Lactate | 5.0 (2.3, 5.4) | 6.5 (4.5, 8.1) | 5.0 (2.3, 5.4) | 6.9 (5.2, 8.3) | 5.0 (2.2, 5.4) | 5.9 (3.9, 7.8) |
| - Missing | 1 | 0 | 1 | 0 | 1 | 0 |
| CSF Glucose | 3.43 (2.51, 4.30) | 2.04 (1.23, 2.46) | 3.42 (2.49, 4.30) | 1.77 (1.12, 2.16) | 3.51 (2.60, 4.42) | 2.13 (1.27, 2.81) |
| - Missing | 1 | 0 | 1 | 0 | 1 | 0 |
| Corresponding Blood Glucose | 7.01 (5.46, 7.72) | 7.01 (5.74, 7.47) | 7.03 (5.47, 7.80) | 6.87 (5.70, 7.10) | 7.06 (5.44, 7.91) | 6.81 (5.72, 7.26) |
| - Missing | 6 | 1 | 6 | 1 | 6 | 1 |
| Cryptococcal Antigen/Indian Ink | 19 (3.2%) | 1 (1.1%) | 19 (3.1%) | 1 (1.4%) | 19 (3.4%) | 1 (0.8%) |
| **1Mean (IQR); n (%)** | | | | | | |

# Discussion

# Conclusion

# References

Boonstra, Philip S., Ryan P. Barbaro, and Ananda Sen. 2019. “Default Priors for the Intercept Parameter in Logistic Regressions.” *Computational Statistics & Data Analysis* 133 (May): 245–56. <https://doi.org/10.1016/j.csda.2018.10.014>.

Chang, Winston. 2020. *R6: Encapsulated Classes with Reference Semantics*. <https://CRAN.R-project.org/package=R6>.

Chen, Yuan-Zhi, Li-Chang Sun, Yao-Hong Wen, Zhong-Wei Li, Shu-Jin Fan, Hong-Kun Tan, Min Qiu, et al. 2020. “Pooled Analysis of the Xpert MTB/RIF Assay for Diagnosing Tuberculous Meningitis.” *Bioscience Reports* 40 (1). <https://doi.org/10.1042/bsr20191312>.

Defazio, Aaron, and Huw Campbell. 2020. *Classifierplots: Generates a Visualization of Classifier Performance as a Grid of Diagnostic Plots*. <https://CRAN.R-project.org/package=classifierplots>.

Donovan, Joseph, Do Dang Anh Thu, Nguyen Hoan Phu, Vu Thi Mong Dung, Tran Phu Quang, Ho Dang Trung Nghia, Pham Kieu Nguyet Oanh, et al. 2020. “Xpert MTB/RIF Ultra Versus Xpert MTB/RIF for the Diagnosis of Tuberculous Meningitis: A Prospective, Randomised, Diagnostic Accuracy Study.” *The Lancet Infectious Diseases* 20 (3): 299–307. <https://doi.org/10.1016/s1473-3099(19)30649-8>.

Gabry, Jonah, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. 2019. “Visualization in Bayesian Workflow.” *J. R. Stat. Soc. A* 182: 389–402. <https://doi.org/10.1111/rssa.12378>.

Hadgu, A., and Y. Qu. 2002. “A Biomedical Application of Latent Class Models with Random Effects.” *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 47 (4): 603–16. <https://doi.org/10.1111/1467-9876.00131>.

Harrell Jr, Frank E. 2021. *Rms: Regression Modeling Strategies*. <https://CRAN.R-project.org/package=rms>.

Marais, Suzaan, Guy Thwaites, Johan F Schoeman, M Estée Török, Usha K Misra, Kameshwar Prasad, Peter R Donald, Robert J Wilkinson, and Ben J Marais. 2010. “Tuberculous Meningitis: A Uniform Case Definition for Use in Clinical Research.” *The Lancet Infectious Diseases* 10 (11): 803–12. <https://doi.org/10.1016/s1473-3099(10)70138-9>.

Mason, Alexina J, Manuel Gomes, Richard Grieve, Pinar Ulug, Janet T Powell, and James Carpenter. 2017. “Development of a Practical Approach to Expert Elicitation for Randomised Controlled Trials with Missing Health Outcomes: Application to the IMPROVE Trial.” *Clinical Trials* 14 (4): 357–67. <https://doi.org/10.1177/1740774517711442>.

Nhu, N. T. Q., D. Heemskerk, D. D. A. Thu, T. T. H. Chau, N. T. H. Mai, H. D. T. Nghia, P. P. Loc, et al. 2013. “Evaluation of GeneXpert MTB/RIF for Diagnosis of Tuberculous Meningitis.” *Journal of Clinical Microbiology* 52 (1): 226–33. <https://doi.org/10.1128/jcm.01834-13>.

Qu, Yinsheng, Ming Tan, and Michael H. Kutner. 1996. “Random Effects Models in Latent Class Analysis for Evaluating Accuracy of Diagnostic Tests.” *Biometrics* 52 (3): 797. <https://doi.org/10.2307/2533043>.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Schumacher, Samuel G., Maarten van Smeden, Nandini Dendukuri, Lawrence Joseph, Mark P. Nicol, Madhukar Pai, and Heather J. Zar. 2016. “Diagnostic Test Accuracy in Childhood Pulmonary Tuberculosis: A Bayesian Latent Class Analysis.” *American Journal of Epidemiology* 184 (9): 690–700. <https://doi.org/10.1093/aje/kww094>.

Stan Development Team. 2021a. “RStan: The R Interface to Stan.” <https://mc-stan.org/>.

———. 2021b. *Stan Modeling Language Users Guide and Reference Manual*. <https://mc-stan.org/>.

———. n.d. “Prior Choice Recommendations stan-dev/stan Wiki GitHub.” Accessed August 12, 2021. <https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations>.

Thwaites, GE, TTH Chau, K Stepniewska, NH Phu, LV Chuong, DX Sinh, NJ White, CM Parry, and JJ Farrar. 2002. “Diagnosis of Adult Tuberculous Meningitis by Use of Clinical and Laboratory Features.” *The Lancet* 360 (9342): 1287–92. <https://doi.org/10.1016/s0140-6736(02)11318-3>.

van Erp, Sara, Daniel L. Oberski, and Joris Mulder. 2019. “Shrinkage Priors for Bayesian Penalized Regression.” *Journal of Mathematical Psychology* 89 (April): 31–50. <https://doi.org/10.1016/j.jmp.2018.12.004>.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2016. “Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC.” *Statistics and Computing* 27 (5): 1413–32. <https://doi.org/10.1007/s11222-016-9696-4>.

White, Ian R, James Carpenter, Stephen Evans, and Sara Schroter. 2007. “Eliciting and Using Expert Opinions about Dropout Bias in Randomized Controlled Trials.” *Clinical Trials* 4 (2): 125–39. <https://doi.org/10.1177/1740774507077849>.

Yan, Linlin. 2021. *Ggvenn: Draw Venn Diagram by ’Ggplot2’*. <https://CRAN.R-project.org/package=ggvenn>.

# (APPENDIX) Appendix

# Model Formulation Details

To be written

# Simulation study

The purpose of this study is to estimate the perfomance an LCA model in the scenario where manifest variables were partially missing.

In this study, we created as three covariates for latent class . For simplicity, , and were three manifest variables whose data generation processes are locally independent w.r.t. . , however were not completely observed; its observation rate (represented by the binary variable obs\_rate) are different for and .

The data generation process was done in version 4.1 (R Core Team 2021) as below:

# Misc function  
generate\_Y = \(C, probs){  
 Cs = sort(unique(C))  
 sapply(C, \(c) rbinom(1,1, probs[Cs==c]))  
}  
  
# Sample size  
N = 1000   
  
# Create data with 3 predictors, X, X2, and X3  
X = rnorm(N, 0, 1 )  
X2 = rnorm(N, 0, 1 )  
X3 = rbinom(N, 1, .3)  
  
# Create latent class  
probs = plogis(3\*X+X2+5\*X3-1)  
C = sapply(probs, \(p) rbinom(1,1,p))  
  
# Manifest variables  
Y1 = generate\_Y(C, c(.1 ,.3))  
Y2 = generate\_Y(C, c(.01,.5))  
Y3 = generate\_Y(C, c(.05,.8))  
  
# Simulate class-aware missing data for Y3. If obs3==1, Y3 is observed  
obs\_rate = c(.1, .95)  
obs3 = generate\_Y(C, obs\_rate)

The likelihood function of the model shall be

Consider an individual : If :

(as are conditionally independent on ).

Similarly, for those whose :

We considered five sets of prior distribution for , and corresponding for the level of prior knowledge. The formulation of priors follow the syntax in @ref(eq:priors-response).

1. Weak prior:
2. Good but conservative knowledge for observation rate, weak prior for manifest variables:
3. Good but conservative knowledge for observation rate and manifest variables:
4. Very strong knowledge for observation rate, weak prior for manifest variables:
5. Very strong knowledge for observation rate and manifest variables:

All models were fitted under the probabilistic language version 2.26 (Stan Development Team 2021b) via inferface (Stan Development Team 2021a).

**Fit results: Hmmm, I remember there is a package showing estimations for one parameter from different models side-by-side in one ggplot. Do you remember?**