|  |  |
| --- | --- |
| Data | Description |
| **studies\_4agegroups.xlsx** | Data collected from incidence studies. Contains notes on some of the cells on how data was harmonized across studies. |
| **./00-data/data\_long.txt** | Data set up for JAGS analysis. |
| **./BC\_sensitivity.csv** | Blood culture diagnostic sensitivity data from nine studies. |
| **./03-LKOvalidation/locations.mat** | The locations that are left out of the analysis in each of the 7 iterations of leave-k-out validation. |
| **mapped\_age\_dist.mat**  **GDP\_CAP\_raster\_all.mat**  **mapped\_vars\_gdl.mat**  **MASK2.mat**  **mapped\_wdi\_vars.mat**  **blue\_water.mat**  **pop\_mapres.mat**  **country\_raster.mat**  **regions\_indicator.mat**  **world\_map.mat**  **floods\_highres.mat**  **world\_map\_regions\_FINAL.txt** | Data for to create the maps. |

The analysis was run using Matlab R2014b and Jags 4.0.

The data used for this analysis:

The code in this analysis:

|  |  |  |  |
| --- | --- | --- | --- |
| File Name | Dependencies | Description | Product |
| In directory 01-bcs |  |  |  |
| **sensitivity\_all.m** | **sensitivity\_nocorr\_re.jags** | Run the analysis to estimate the relationship between blood sample volume and culture sensitivity. | Gamma distribution to characterize the uncertainty of culture sensitivity. Produces **Figure S1**. |
| **sensitivity\_nocorr\_re.jags** |  | JAGS script to sample parameters |  |
| In directory 02-spike\_slab | | | |
| **jags\_spike\_slab\_locations\_1re\_4ages\_slopere.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_2levels.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_2levels\_gbd.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_NULL.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_NULL\_2levels.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_NULL\_2levels\_gbd.m** | **data\_long.txt**  **data\_long.m**  **prep\_jags\_input.m**  **prep\_jags\_initial\_sat.m**  **prep\_jags\_initial\_null.m**  **/jags\_code/tss\_int\_1re\_4ages\_hp\_slope\_taugamma4.jags** | Runs the model as well as alternative models that were run for sensitivity analysis. Takes about 10 hours in a PC with a 4.0 GHz Intel 7 processor. | Posterior samples of the model, **structArray**. The means, standard deviations, minima, maxima of the predictors in the model, in **normalize**. |
| **data\_long.m** |  | Formatting the text data to be useful with matlab functions. |  |
| **normalizing\_data.m** |  | Normalizes the predictor variables. Stores the mean, standard deviation, minima, and maxima of the predictors to transform the predictors for the prediction dataset. | The matlab structure **normalize.** |
| **prep\_jags\_input.m** |  | Formats the predictor data, the observed data, and initial parameters for JAGS. | The table **datajags**, which goes into the **matjags** function. |
| **prep\_jags\_initial\_sat.m** |  | Produces initial parameter estimates (for fixed and random effects) for the MCMC chain initiated with a model that includes a model with all the predictors of the intercept. | The structure **initial\_sat**. |
| **prep\_jags\_initial\_null.m** |  | Produces initial parameter estimates (for random effects) for the MCMC chain initiated with model that does not include any predictors. | The structure **initial\_null**. |
| **matjags.m** |  | Interfaces between matlab and JAGS. |  |
| **/jags\_code/tss\_int\_1re\_4ages\_hp\_slope\_NULL.jags** |  | Describes the model from which JAGS draws posterior distributions. This is the null model, which does not include any predictors, only terms for 1-level random effects and age-specific incidence. |  |
| **/jags\_code/tss\_int\_1re\_4ages\_hp\_slope\_taugamma4.jags** |  | Describes the model from which JAGS draws posterior distributions. This is the model with predictors, with 1 level of random effects (location). |  |
| **/jags\_code/tss\_int\_1re\_4ages\_hp\_slope\_NULL\_2levels.jags** |  | Describes the model from which JAGS draws posterior distributions. This is the null model, which does not include any predictors, only terms for 2-level random effects (location and continent) and age-specific incidence. |  |
| **/jags\_code/tss\_int\_1re\_4ages\_hp\_slope\_taugamma4\_2levels.jags** |  | Describes the model from which JAGS draws posterior distributions. This is the model with predictors, with 2 levels of random effects (location & continent). |  |
| **/jags\_code/tss\_int\_1re\_4ages\_hp\_slope\_NULL\_2levels\_gbd.jags** |  | Describes the model from which JAGS draws posterior distributions. This is the null model, which does not include any predictors, only terms for 2-level random effects (location and GBD region) and age-specific incidence. |  |
| **/jags\_code/tss\_int\_1re\_4ages\_hp\_slope\_taugamma4\_2levels\_gbd.jags** |  | Describes the model from which JAGS draws posterior distributions. This is the model with predictors, with 2 levels of random effects (location & GBD region). |  |
| In directory 03-LKOvalidation | | | |
| **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_LKO1.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_LKO2.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_LKO3.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_LKO4.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_LKO5.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_LKO6.m**  **jags\_spike\_slab\_locations\_1re\_4ages\_slopere\_LKO7.m** |  |  | The structure **structArray** with a list **out** of the locations that were excluded in each iteration of leave-k-out validation. |
| In directory 04-results | | | |
| **violin.m** |  | Makes violin plots. |  |
| **post\_spike\_slab\_summarize\_output.m** |  | Summarizes the posterior distributions of the JAGS analysis. | Produces the following tables and figures in the manuscript: **Figure 3, Figure S2,** and **Table S3.** |
| **post\_spike\_slab\_int\_4ages\_slopere.m** |  | Simulates the incidence predicted by each of the models | Produces predictions necessary to produce **Figure 5**. |
| **summarize\_LOO\_results.m** |  | Summarizes the posterior distributions of the JAGS analysis in each iteration of leave-k-out validation. | Produces the following tables and figures in the manuscript: **Figure S4** and **Figure S5**. Produces predictions necessary to produce **Figure 5**. |
| **post\_spike\_slab\_plots.m** | Predictions produced using **post\_spike\_slab\_int\_4ages\_slopere.m** and **summarize\_LOO\_results.m** | Graphs model-predicted incidence against observed incidence. | Produces the following tables and figures in the manuscript: **Figure 4, Figure 5**, and **Figure S3**. |
| **validation.m** | **tsap\_logincidence.mat** and **tsap\_simulations.mat** | Validation with TSAP data. | Produces the following tables and figures in the manuscript: **Figure 6**. |
| In directory 06-maps | | | |
| Code to simulate incidence in all LMIC and makes maps | Map data in 00-data. | Takes about 5 hours at in a PC with an 4.0 GHz Intel 7 processor. | Produces simulated incidence for each age group for all LMIC countries\* in the world.  Produces the following tables and figures in the manuscript: **Figure 7** and **Figure 8**. |

\*LMIC countries were defined as countries that fell outside the high-income lending bracket by the World Bank in the years 2011-2015: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.