

UC Riverside

UC Riverside Previously Published Works

Title

The burden of typhoid fever in low- and middle-income countries: A meta-regression approach

Permalink

<https://escholarship.org/uc/item/3k7933sm>

Journal

PLOS Neglected Tropical Diseases, 11(2)

ISSN

1935-2727

Authors

Antillón, Marina
Warren, Joshua L
Crawford, Forrest W
[et al.](#)

Publication Date

2017

DOI

10.1371/journal.pntd.0005376

Peer reviewed

RESEARCH ARTICLE

The burden of typhoid fever in low- and middle-income countries: A meta-regression approach

Marina Antillón^{1*}, Joshua L. Warren², Forrest W. Crawford², Daniel M. Weinberger¹, Esra Kürüm³, Gi Deok Pak⁴, Florian Marks⁴, Virginia E. Pitzer¹

1 Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, Connecticut, United States of America, **2** Department of Biostatistics, Yale School of Public Health, New Haven, Connecticut, United States of America, **3** Department of Statistics, University of California Riverside, Riverside, California, United States of America, **4** International Vaccine Institute, Seoul, Republic of Korea

* marina.antillon@yale.edu



Abstract

Background

Upcoming vaccination efforts against typhoid fever require an assessment of the baseline burden of disease in countries at risk. There are no typhoid incidence data from most low- and middle-income countries (LMICs), so model-based estimates offer insights for decision-makers in the absence of readily available data.

Methods

We developed a mixed-effects model fit to data from 32 population-based studies of typhoid incidence in 22 locations in 14 countries. We tested the contribution of economic and environmental indices for predicting typhoid incidence using a stochastic search variable selection algorithm. We performed out-of-sample validation to assess the predictive performance of the model.

Results

We estimated that 17.8 million cases of typhoid fever occur each year in LMICs (95% credible interval: 6.9–48.4 million). Central Africa was predicted to experience the highest incidence of typhoid, followed by select countries in Central, South, and Southeast Asia. Incidence typically peaked in the 2–4 year old age group. Models incorporating widely available economic and environmental indicators were found to describe incidence better than null models.

Conclusions

Recent estimates of typhoid burden may under-estimate the number of cases and magnitude of uncertainty in typhoid incidence. Our analysis permits prediction of overall as well as age-specific incidence of typhoid fever in LMICs, and incorporates uncertainty around the

OPEN ACCESS

Citation: Antillón M, Warren JL, Crawford FW, Weinberger DM, Kürüm E, Pak GD, et al. (2017) The burden of typhoid fever in low- and middle-income countries: A meta-regression approach. *PLoS Negl Trop Dis* 11(2): e0005376. doi:10.1371/journal.pntd.0005376

Editor: Marília Sá Carvalho, Oswaldo Cruz Foundation, BRAZIL

Received: August 26, 2016

Accepted: January 31, 2017

Published: February 27, 2017

Copyright: © 2017 Antillón et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All files needed to implement the work described in the manuscript are available at <https://github.com/Marina-Antillon/typhoid-burden>. All other relevant data are contained in the Supporting Information files.

Funding: This work was funded by the Bill and Melinda Gates Foundation (OPP1116967). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

model structure and estimates of the predictors. Future studies are needed to further validate and refine model predictions and better understand year-to-year variation in cases.

Author summary

Typhoid fever is a bacterial enteric infection that continues to pose a considerable burden to the 5.5 billion people living in low- and middle-income countries (LMICs). We developed and validated a model incorporating widely available indicators of economic and social development and the environment to estimate the burden of typhoid fever across LMICs. Our analysis uses all available data to estimate the incidence of typhoid in key age groups, which is important for the design and implementation of optimal vaccination strategies, and it identifies regions of the world that have the most uncertainty in typhoid incidence. Across all LMICs, we estimated that the expected number of typhoid fever cases per year is 17.8 million (95% CI: 6.9–48.4 million). We also present the probability that incidence surpasses the criteria for low, medium, high, and very high incidence in each country, which could help guide policy in the face of uncertainty.

Introduction

Typhoid fever has been estimated to cause between 9.9 and 24.2 million cases and 75,000–208,000 deaths per year [1–3]. Typhoid fever is caused by infection with *Salmonella enterica* serovar Typhi, a gram-negative bacterium that invades the body via the small intestines and colonizes macrophages in the reticuloendothelial system, from where it is shed into the bloodstream [4,5]. Symptoms of the resulting disease typically include prolonged fever, frontal headache, malaise and marked loss of appetite, sometimes accompanied by abdominal pains, nausea, and (in severe cases) intestinal perforation and neurological complications [6]. Symptoms typically subside in 7–21 days, but mortality is estimated to occur in 1–5% of hospitalized patients [7–9]. In a small percentage of cases, the bacteria may also colonize the gall bladder, leading to a chronic carrier state [6].

Data on the incidence of typhoid fever are scarce in low- and middle-income countries (LMICs). The symptoms of typhoid fever resemble those of many other significant febrile diseases, precluding straightforward estimates of typhoid incidence [10,11]. Recent estimates have relied on key expert assumptions, primarily geographical groupings that coincide with UN development regions or pre-determined epidemiological regions [1,2,12]. The degree to which incidence may be attributable to geography as well as to indicators of poverty and socio-economic circumstances remains mostly unexamined [3]. Considering that typhoid incidence may vary both between and within countries, it is necessary to identify potential predictors of incidence that facilitate interpolation across LMICs, where the disease is suspected to remain endemic.

Furthermore, variation in the age distribution of typhoid fever across settings is not well understood. Cases tend to be concentrated in younger age groups in settings with higher transmission and distributed more equally among different ages in low-transmission settings [1,2,12]. However, recent studies have cast doubt on the generality of these age patterns in relation to overall incidence [10,13–15]. Identifying predictors of the age distribution of typhoid fever is of particular salience to the design and implementation of optimal vaccination strategies.

We explored the potential contribution of demographic, environmental, and socioeconomic indicators that serve as candidate predictors of the age-specific incidence of typhoid fever. We used a data-driven approach to predict the mean and variance in age-specific incidence while accounting for uncertainty in the underlying model structure, and validated our predictions against out-of-sample data.

Methods

Data sources

We carried out a literature search to identify population-based studies that reported incidence of culture-confirmed typhoid fever for the period of 1980–2014. We excluded all hospital-based or clinic-based studies that did not constitute exhaustive surveillance of typhoid in a well-defined population. Further details of the literature search are presented in the [S1 Text](#).

We gathered data on possible predictors of typhoid fever incidence from publicly available databases, aiming to identify indicators of environmental characteristics and socioeconomic development for all LMICs. Predictors were chosen for their ubiquity and relevance to water-borne disease transmission in consultation with typhoid experts. [Table 1](#) lists the predictors included in this analysis and the source of data; [S1 Text](#) provides more details. The predictors' values were matched as closely as possible to the time and location of each incidence study.

We validated our model predictions against previously unpublished data from the Typhoid Fever Surveillance in Africa Program (TSAP, see [S2 Table](#)), which consisted of passive, population-based surveillance in 9 of 13 sites in 10 countries across sub-Saharan Africa [24,25]. We extracted predictor data for the TSAP studies from the same databases used for the estimation sample ([Table 1](#)).

Model framework

The observed data on typhoid fever incidence result from both a disease process and an observation process. We aimed to take into account both of these processes in our modeling strategy, as illustrated in [Fig 1](#).

Table 1. Summary statistics for the candidate predictors included in the predictive model.

Covariate	Resolution	Mean and range in estimation sample	Mean and range in TSAP* sample	Mean and range in prediction sample	Source
Population density: pop/km ² †	1/24 x 1/24 degree	2,091 (17–47,024)	1,205 (46–11,592)	7.06 (0–98,928)	[16]
GDP per capita, 2005 international dollars †	1x1 degree	\$1,780 (\$500–\$5,509)	1,130 (841–1,608)	\$6,909 (694–53,836)	[17]
Gini coefficient	National	41 (31–59)	42 (37–50)	39 (29–57)	[18]
Access to improved source of water ‡	Subnational, national	35 (2–90)	25 (3–66)	61 (1–100)	[19,20]
Access to improved source of sanitation ‡	Subnational	38 (2–96)	17 (3–63)	57 (0.36–98)	[19,20]
Years of education for women over age 20 †	Subnational	4.7 (1.9–10.3)	3.8 (0.5–9.8)	1.7 (0.6–10.7)	[18,19]
% roads paved ‡	National	40 (10–87)	4 (4–28)	47 (1.8–91)	[18]
% population living on <\$2/day ‡	National	59 (18–93)	78 (52–93)	25 (0.05–95)	[18]
Prevalence of stunting ‡	Subnational, national	42 (6–60)	34 (18–51)	17 (0.02–82.7)	[18,19]
Number of major floods 1985–2011 ²	50m ²	21 (4–99)	11 (1–25)	3.7 (0–32)	[21]
HIV prevalence ‡	National	0.958 (0.013–6.2)	2.46 (0.3–6)	1.42 (0.1–28.8)	[22]
Water stress †	½ x ½ degree	1.7 (0.0021–6.8)	1.17 (0.07–3.46)	1.37 (0.10–10)	[23]

* TSAP: Typhoid Fever Surveillance in Africa Program

† Log transformed values were used. Geometric means were reported in this case.

‡ Logistic transformed values were used.

doi:10.1371/journal.pntd.0005376.t001

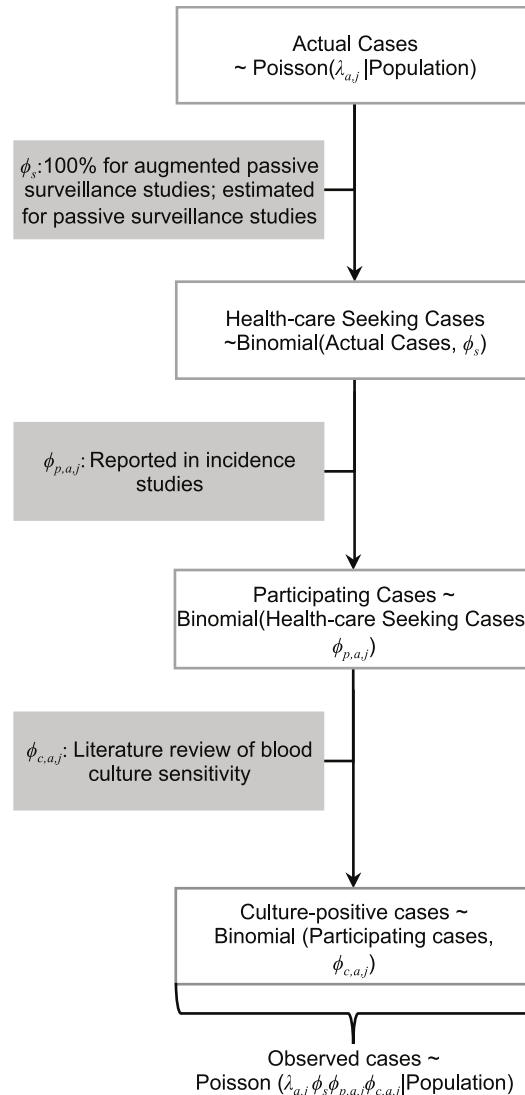


Fig 1. Flowchart of the disease and observation process reflected in our data. The top of the flowchart consists of all cases of typhoid fever in the population, which is what we are ultimately interested in estimating. However, three intervening (observation) processes result in a considerable difference between the “actual” cases of typhoid fever and the observed cases of typhoid fever. In all, the observation process is made up of the type of surveillance employed in each study (ϕ_s), the participation rate of patients seen at each of the study clinics ($\phi_{p,a,j}$, adjusted for age a and study site j), and the sensitivity of blood culture used to confirm typhoid infection ($\phi_{c,a,j}$, adjusted for age a and study site j). We modeled the observed cases as representing the successful trials of a binomial process (the detection process), where the number of trials is Poisson distributed (with a rate parameter equal to the “true” incidence of typhoid cases, $\lambda_{a,j}$, at age a and study site j); thus, the number of observed cases is a thinned Poisson distribution with rate parameter equal to the product of the disease rate and the probability of a case being detected at each step of the observation process. We used Bayesian priors to account for the observation process and adequately estimate the underlying disease process.

doi:10.1371/journal.pntd.0005376.g001

Disease process. We employed a generalized linear mixed-effects model with a log link function to characterize the true underlying incidence of typhoid ($\lambda_{a,j}$) in setting j and age group a : <2 years, 2–4 years, 5–14 years, ≥ 15 years and older. We estimated the age-specific incidence as a function of fixed effects (modeled by the inclusion of predictors for slope coefficients of the age variables) and location-specific random effects. The age groups were chosen

based on the resolution of the available data and their salience in typhoid vaccine policymaking: current vaccines are licensed for children over 2 years of age, and some vaccine programs have employed school-based campaigns targeting children 5–14 years of age [26,27]. Observations from some sites were only available for a combination of age groups, so we adjusted our analysis where necessary to accommodate the lessened granularity of the observed age patterns (see [S1 Text](#)).

First, we considered a null model in which annual incidence was based on age-group fixed effects and random effects for each age group in each location. The intercept, B_0 , represents the incidence in the referent age group (designated as 5–14 years olds) in each setting, and the slopes B_a represent the incidence rate ratio between the three other age groups and the referent group. An offset equal to the log of person-time was included to adjust for the size of the population under observation:

$$\text{Actual Cases}_{a,j} \sim \text{Poisson}(\lambda_{a,j} | \text{person-time}_{a,j})$$

$$\log(\lambda_{a,j}) = B_{0,j} + B_{a,j} + \log(\text{person-time}_{a,j})$$

$$B_{0,j} = \beta_0 + \alpha_{0,j}$$

$$B_{a,j} = \beta_a + \alpha_{a,j}$$

where $\alpha_{0,j} \sim \text{MVN}(0, \Sigma)$ with Σ as the covariance structure of the random effect terms. $B_{a,j} = 0$ for 5–14 year olds.

To test the assumption that typhoid fever incidence correlates with indicators of environmental characteristics and socioeconomic development, we evaluated whether including additional covariates (X_j) to model both the intercept and the slope improved our ability to predict the incidence of typhoid fever:

$$\text{Actual Cases}_{a,j} \sim \text{Poisson}(\lambda_{a,j} | \text{person-time}_{a,j})$$

$$\log(\lambda_{a,j}) = B_{0,j} + B_{a,j} + \log(\text{person-time}_{a,j})$$

$$B_{0,j} = \beta_0 + \gamma X_j + \alpha_{0,j}$$

$$B_{a,j} = \beta_a + \eta_a X_j + \alpha_{a,j}$$

where γ and η_a are the effect sizes corresponding to each predictor for the intercept and slope, respectively; again, $\alpha_{0,j} \sim \text{MVN}(0, \Sigma)$ and $B_{a,j} = 0$ for 5–14 year olds. Variable selection for the predictors was based on the spike-and-slab method described below.

Observation process. Three additional sub-processes that impact our observed data were taken into account ([Fig 2](#), details in [S1 Text](#)). First, we adjusted for the difference in case ascertainment between active and passive surveillance. We assumed active surveillance would be capable of identifying all cases and estimated a fixed effect for the relative ascertainment under passive surveillance (see [S1 Text](#)). Second, we included the participation rate of patients at each location (by age, if possible) as a fixed parameter based on the reported proportion of patients meeting the case definition who had blood drawn for diagnosis. Third, we adjusted for the sensitivity of blood culture to detect typhoid infection. We estimated the test sensitivity as a function of age group for each location using strong prior distributions informed by a

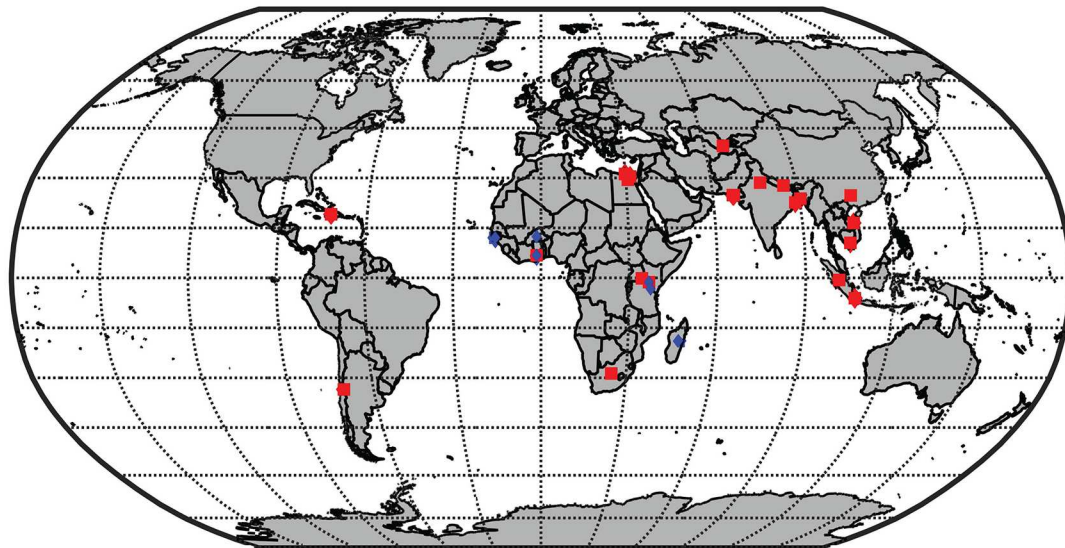


Fig 2. Map of the location of studies in our dataset. Studies used in the estimation sample are depicted in red and the studies used in the validation sample are depicted in blue. The studies in the validation sample come from the Typhoid Fever Surveillance in Africa Program (TSAP).

doi:10.1371/journal.pntd.0005376.g002

meta-analysis of the relationship between sample volume and blood culture sensitivity (See [S1 Text](#)).

Model selection and parameter estimation. We used a Bayesian framework for model estimation, which allowed us to incorporate prior information on diagnostic test sensitivity, as described above. We used a stochastic search variable selection algorithm employing spike-and-slab priors to estimate the fixed effects of the predictors while incorporating uncertainty in model structure (see [S1 Text](#)) [28,29]. To assess convergence on the space of predictor combinations, we ran two chains, one initiated with a null model and one initiated with a saturated model; if each covariate was selected for inclusion with approximately equal chance in each chain, we concluded that the algorithm had converged. The model was fit using the JAGS (Just Another Gibbs Sampler) software, version 3.4.0, in conjunction with MATLAB 2014b via the MATJAGS interface [30,31].

Model validation

We sought to validate the predictive ability of our model in two ways: using a leave-k-out validation method and by comparing model predictions to out-of-sample data from the recent TSAP studies [24,32]. For the leave-k-out validation, we randomly partitioned the dataset into seven sets of three locations each. We fit the model to data from six of the seven sets of locations, and used the fitted model to predict the incidence in the seventh set. We re-sampled the model seven times, excluding one of the groups each time. To assess the improvement in model fit, we compared the covariance in the predictions produced by the null model and the model with predictors, as well as predictions from the leave-k-out validation.

Posterior model predictions

We drew 1,000 samples from the estimated model in order to obtain posterior predictions of the incidence rate across all countries classified as lower income, lower-middle income, or upper-middle income at least once in the period of 2011–2015 by the World Bank. We mapped

the median estimates of the predicted incidence by age using a map resolution of 0.1 degrees. We capped the predicted incidence at 10,000 per 100,000 person-years. To obtain estimates of the incidence for each region, we took the population-weighted sum of the estimated incidence over the raster surface. To characterize the uncertainty in the incidence, we calculated the proportion of the posterior predictive sample in each of four incidence categories: <10 , $10-<100$, $100-<500$, and ≥ 500 cases per 100,000 person-years, designated as low, medium, high, and very high incidence.

Results

We identified 32 studies from 22 sites located in 14 countries (Fig 2); these studies are detailed in S1 Table. We extracted data on case counts, person-time of observation, and details of study design. In total, the studies observed 2,668 cases of typhoid in 3,329,183 person-years of observation. The validation dataset observed 140 cases of typhoid fever during 212,312 person-years of observation.

Predictor selection and age-specific incidence estimates

We evaluated twelve potential predictors of typhoid fever incidence (Table 1). Fig 3A shows the posterior distribution of the probability of inclusion for each predictor. Both chains provided equivalent probabilities of variable inclusion as well as equivalent distributions for the size of the underlying model (Fig 3B), indicating that our algorithm converged.

No single covariate was included in all models. The percent of roads paved, prevalence of stunting, and percent of the population living in extreme poverty were the most sampled covariates (present in 99%, 97%, and 91% of all models in both chains); in almost all models in which they were present, the covariates were useful to predict both the overall incidence as well as the age-specific incidence rate ratios in each setting. HIV prevalence and flood risk were the next most sampled predictors (present in 85% and 45% of all models in both chains), but these indices were only useful to predict the overall incidence. Indices for income inequality (Gini coefficient), access to flush toilets, and GDP per capita were sampled in just over a third of all models, while the rest of the predictors were each included in 21–30% of all models. In total, the models had a median of six predictors, and 95% of models had between three and nine predictors; the null model was never sampled (Fig 3B).

The model estimated a lower incidence rate (on average) among <2 year olds and ≥ 15 year olds, and a slightly higher incidence rate among 2–4 year olds compared to 5–14 year olds (Fig 3C). Furthermore, the model reproduced the heterogeneity in age-specific incidence (Fig 4). A weak relationship between overall incidence and the shape of the age distribution is evident: a higher overall incidence is associated with a peak in incidence among children 2–4 years of age instead of a more uniform burden across different ages, with a slight peak among 5–14 year olds in lower incidence settings.

Model validation and predictions

The posterior incidence predictions are shown in Fig 5 for the null model, the model with predictors, and the leave-k-out validation. Including the predictors improved the model fit, although the models tended to underestimate the incidence in <5 year olds in high incidence settings and overestimate it in low incidence settings. The model including random effects explained the residual variance and provided a good fit to the data (S3 Fig). Importantly, leave-k-out validation shows that most out-of-sample credible intervals contained the observed incidence (Fig 5C). Moreover, estimating the models using subsets of the data (all

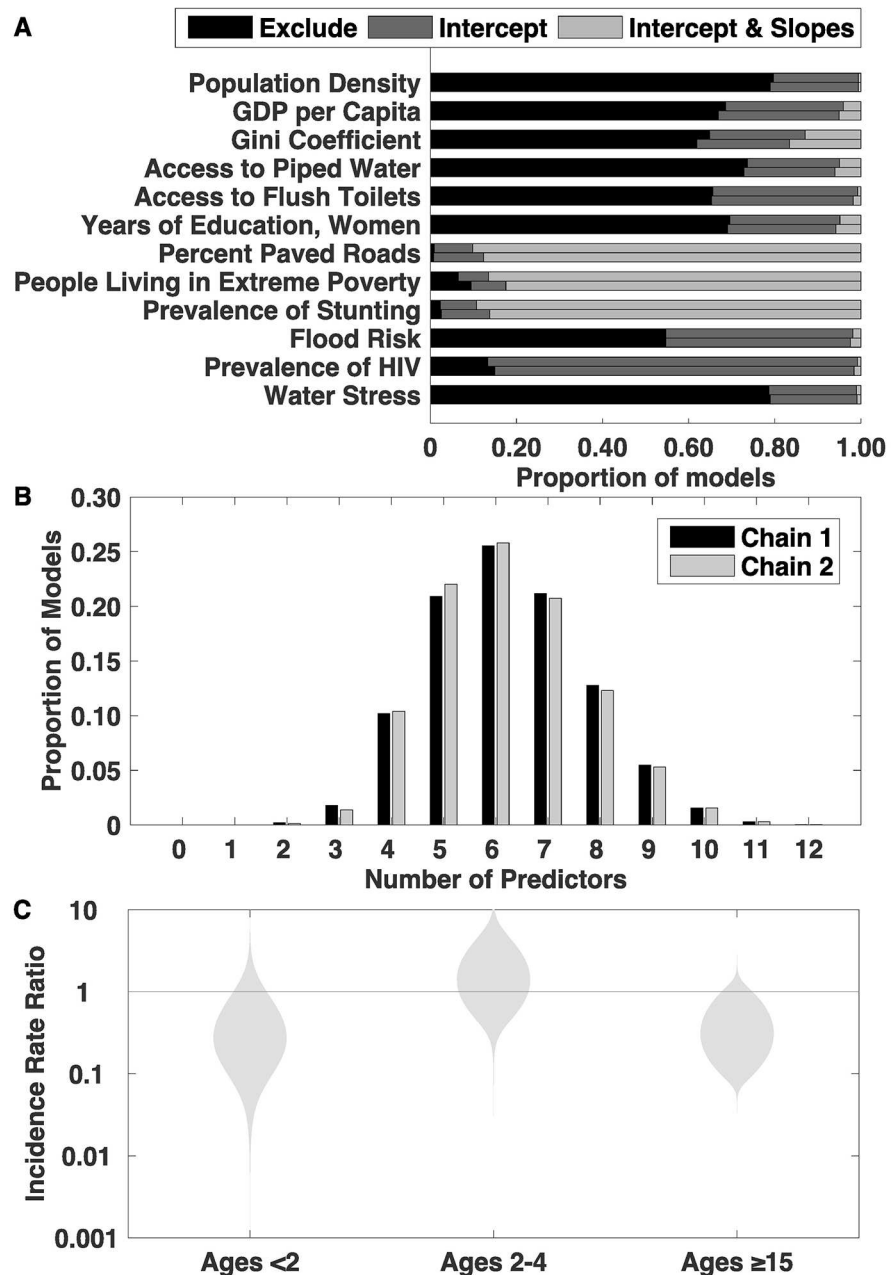


Fig 3. Model summary. A) The posterior marginal probability that each variable was excluded from the model (black) or included as a predictor of the intercept (dark grey) or intercept and slope (light grey) is shown for two chains. Our stochastic search variable selection algorithm could include variables either as a predictor of the intercept (the incidence in 5–14 year olds) or as a predictor of the intercept as well as the slopes (the incidence rate ratios between the other age groups and the referent age group of 5–14 year olds). B) Distribution of the average number of covariates in the model. Chain 1 was initiated using a model that included all the covariates as predictors of the main effect, while chain 2 was initiated as the null model. The null model was never sampled, implying that the models including at least one predictor better described the data than the null model. C) Posterior distributions of age-specific incidence rate ratios between the referent age group (5–14 years of age) and other age groups: <2 years, 2–4 years, ≥15 years old.

doi:10.1371/journal.pntd.0005376.g003

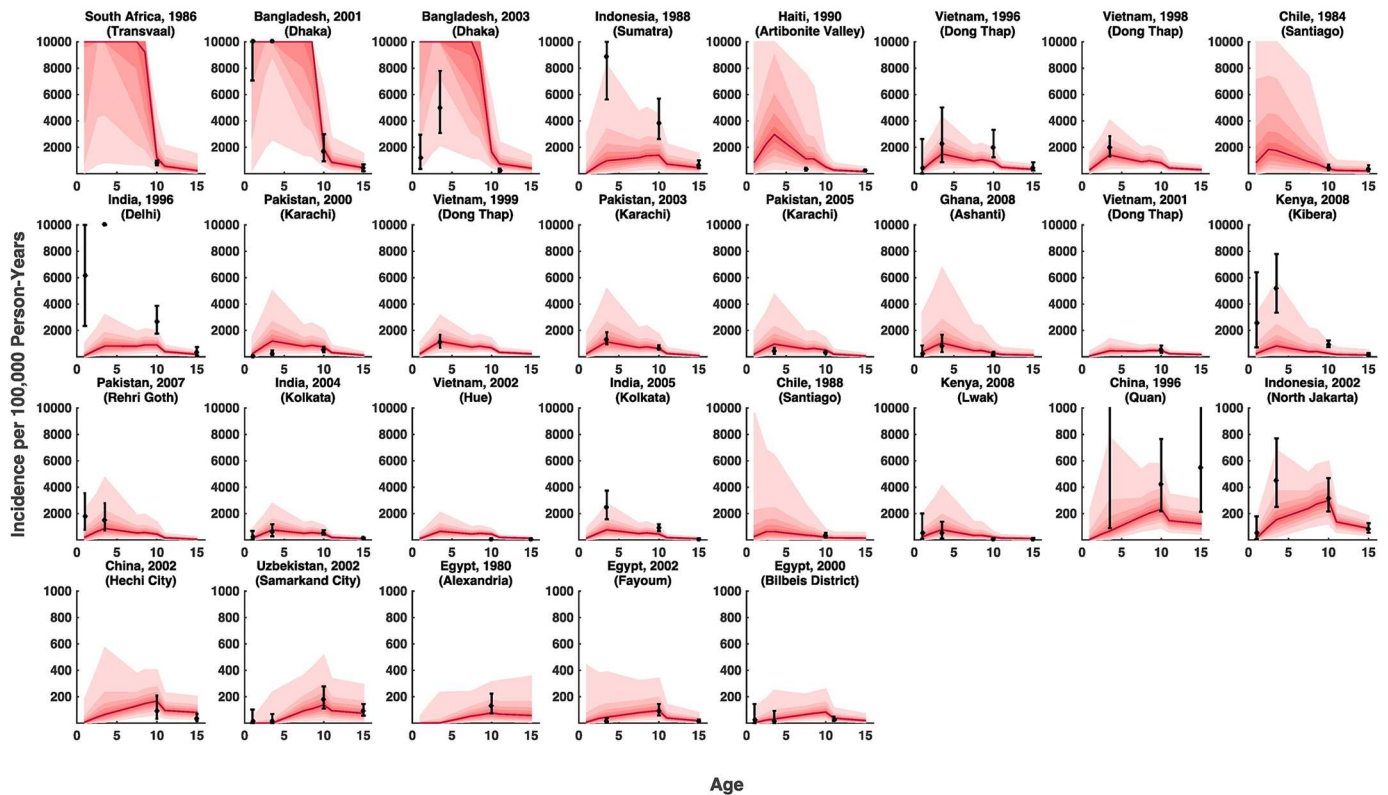


Fig 4. Observed versus predicted age-specific incidence rates. Sites are labeled by location and year, and plots are ordered by decreasing overall model-predicted incidence. The red line and regions represent the model fits—median and 95% credible interval of the expected incidence estimated by the joint posterior distribution of model parameter (excluding study specific random effects and the impact of the observation process). The black symbols are the observed incidence with the 95% credible intervals after adjusting for the observation process: surveillance type (active/augmented passive versus passive surveillance), the participation rate, and blood culture sensitivity. Only studies that reported age-specific incidence are featured here.

doi:10.1371/journal.pntd.0005376.g004

but 3 observations) would yield similar profiles in terms of the variables selected for inclusion and the size of the model (S4 and S5 Figs).

Model validation against the TSAP data performed well for some sites, but showed large variance in posterior estimates of incidence, as well as notable overestimates of incidence in several locations (Fig 6). The 95% credible intervals (CIs) overlapped with the observed incidence in some locations, but not others. However, the 95% CI of the observed incidence also showed a large amount of uncertainty in directly measured incidence estimates (sometimes spanning three orders of magnitude). No clear geographic pattern emerged to distinguish between observations that did and did not match model predictions. The model successfully predicted the incidence in Kibera, Kenya (the site of a previous study) for <15 year olds, as well as locations for which there was no previous data, e.g. Bandim (Guinea Bissau), rural Moshi (Tanzania) for ≥ 2 year olds, Polesgo (Burkina Faso) for all ages except 2–4 year olds, and Nioko II in Burkina Faso for <15 year olds. Successful prediction in more than one setting, as well as more than one age group, indicates that our models were able to capture some of the within-country heterogeneity in incidence.

Given the environmental and socioeconomic composition of LMICs, our model predicts that typhoid fever incidence should be highest in parts of Central Africa, Turkmenistan and Uzbekistan in Central Asia, as well as Mongolia and western China (Fig 7). Incidence is predicted to be highest in 2–4 year olds or 5–14 year olds, but the age of peak incidence varies

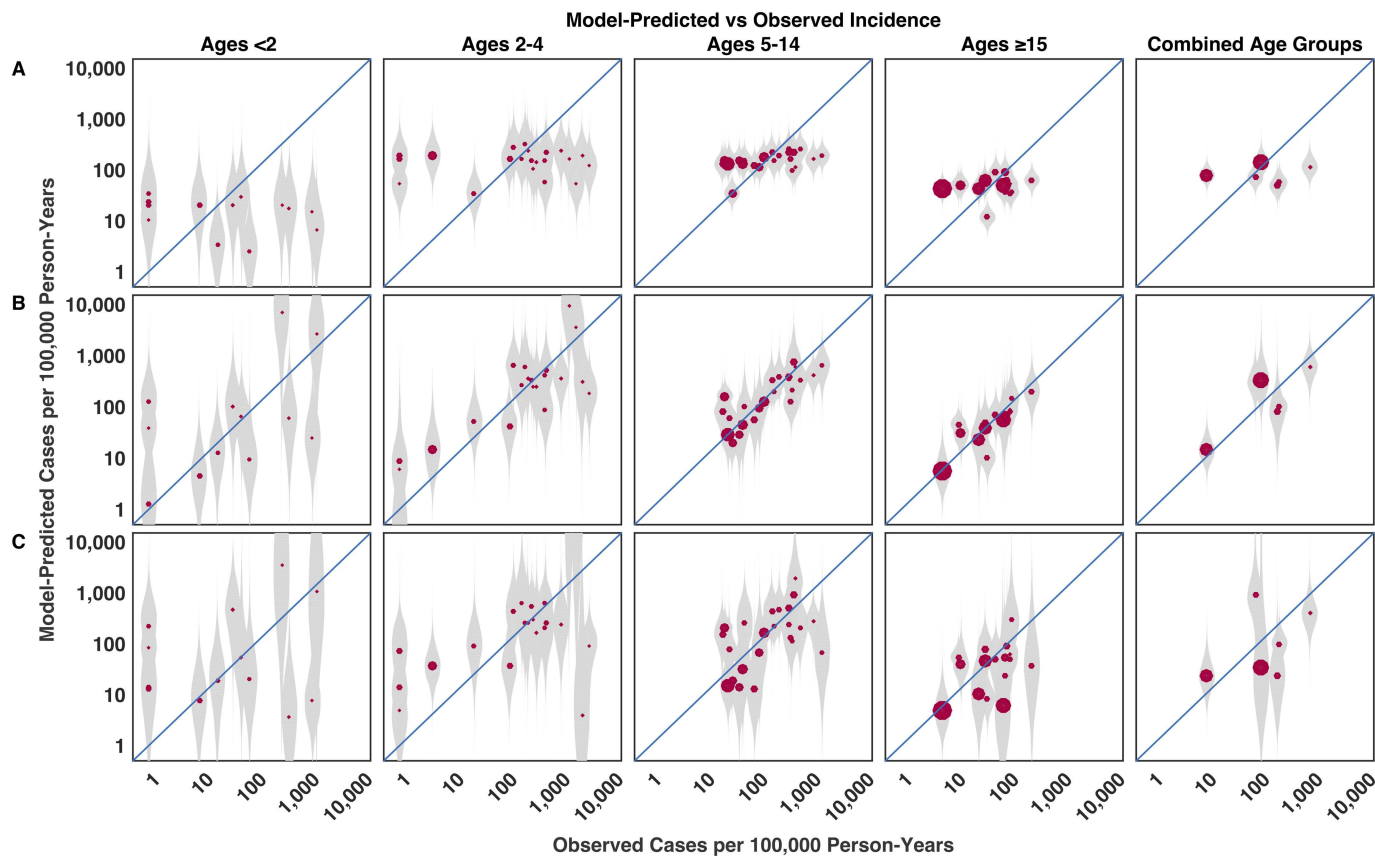


Fig 5. Observed versus model-predicted incidence. (A) Posterior predictions from the null model, which only adjusts for age and the observation process. (B) Posterior predictions from the model using fixed effects for the predictors. (C) Leave-3-out validation results. The gray markers represent the density of model-predicted posterior distributions of incidence, while the red dots represent the median posterior predicted incidence. The size of the red circular markers is proportional to the number of person-years of observation in each study. All predictions are of the mean incidence and were generated using only the fixed-effect terms of the model, and hence do not account for unmeasured location-specific differences, e.g. in healthcare-seeking behavior.

doi:10.1371/journal.pntd.0005376.g005

from place to place (Fig 7). The incidence is lower in children <2 years of age and adults in most settings.

Approximately 6 billion people lived in all LMICs in 2015. We estimated that the expected number of typhoid fever cases per year is 17.8 million across all LMICs (95% CI: 6.9–48.4 million) (Table 2). According to our analysis, almost 40% of all cases occur in sub-Saharan Africa (7.2 million, 95% CI: 2.2–30.2 million), although the uncertainty around our estimates is considerable.

Geographic heterogeneity

Although South Asia has the second largest case count among the regions, the region has the third highest expected incidence rate after sub-Saharan Africa and North Africa/Middle East (Table 2). Considerable heterogeneity in the expected incidence exists within each region, as well. Within sub-Saharan Africa, Central Africa has the highest expected incidence and Southern Africa has the lowest expected incidence. The incidence in Andean Latin America outpaces the incidence in other parts of Latin America by a factor of six or more. Oceania has the highest incidence of any sub-region in the world (Table 2).

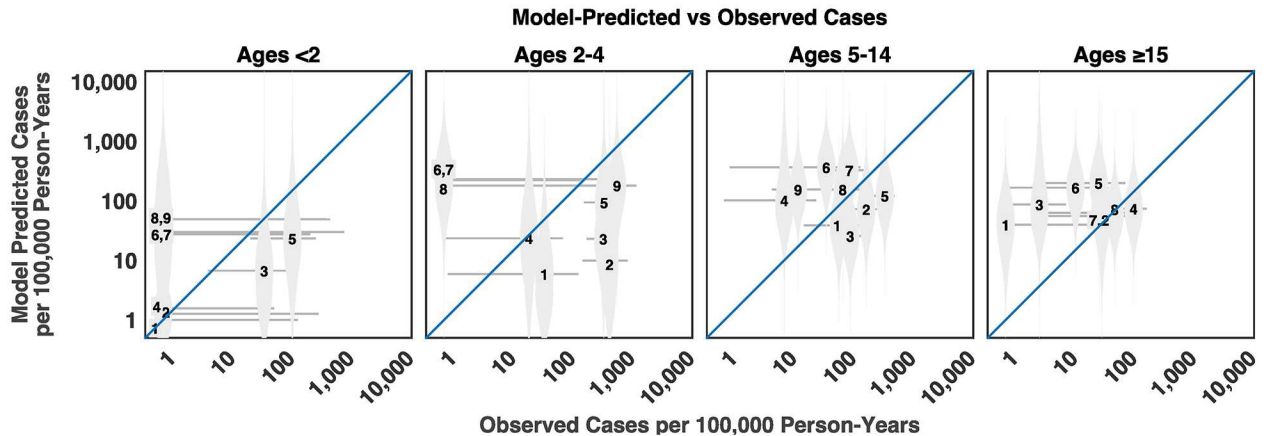


Fig 6. Out-of-sample validation. The observed versus predicted incidence of typhoid fever is plotted for studies in the Typhoid Fever Surveillance in Africa Program (TSAP) using a model estimated from previously published data identified in our literature review. The numbers represent the median posterior predicted incidence for each TSAP site: 1- Nioko II, Burkina Faso. 2 –Polesgo, Burkina Faso. 3 – Ashanti Akim North, Ghana. 4 –Bandim, Guinea Bissau. 5 –Kibera, Kenya. 6 –Antananarivo, Madagascar. 7 –Imerintsiasosika, Madagascar. 8 –Moshi rural, Tanzania. 9 –Moshi urban, Tanzania. The gray markers represent the density of model-predicted posterior distributions of incidence. The gray horizontal lines represent 95% confidence intervals for the observed incidence.

doi:10.1371/journal.pntd.0005376.g006

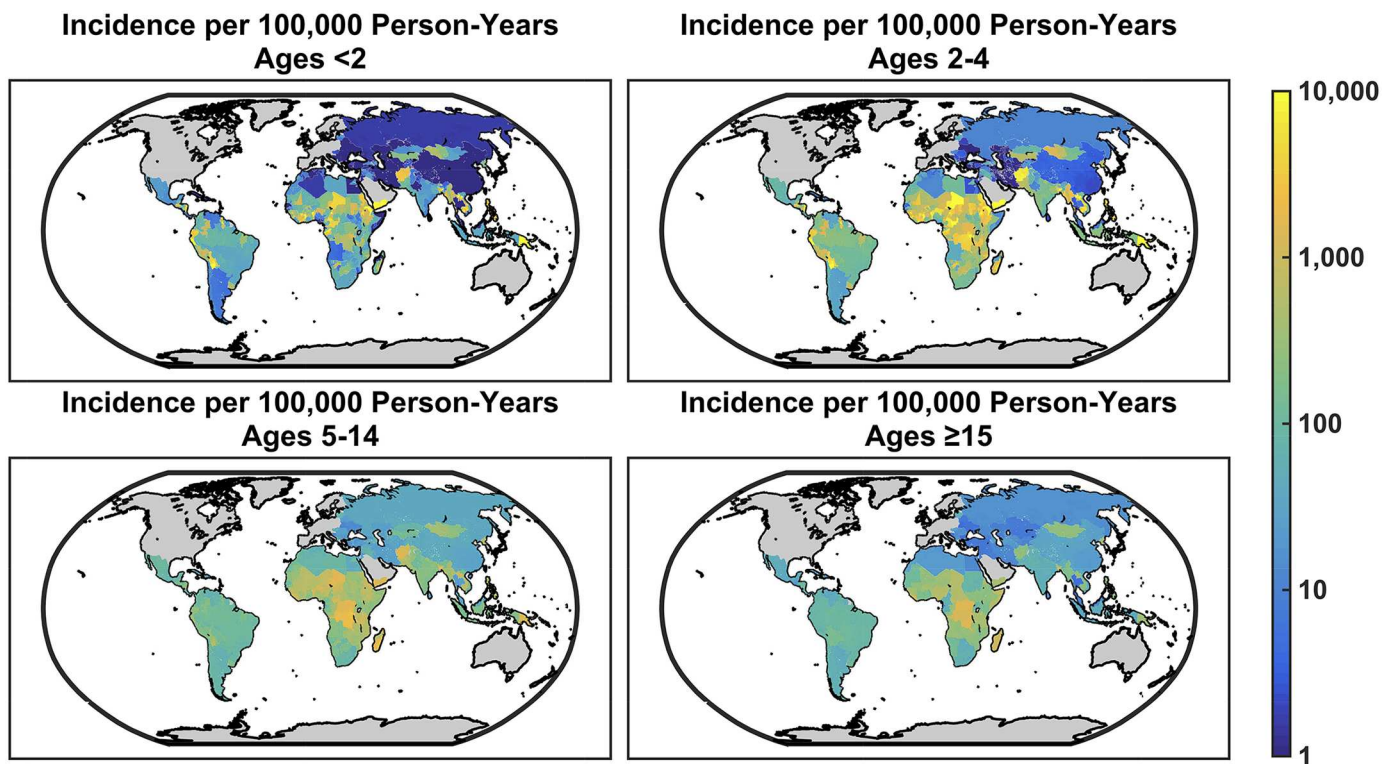


Fig 7. Model-predicted age-specific incidence per 100,000 person-years. The median posterior predicted incidence per 100,000 person-years in each of the age groups (<2 years, 2–4 years, 5–14 years, and ≥15 years) is mapped for all low- and middle-income countries (LMICs) with a resolution of 0.1 degrees.

doi:10.1371/journal.pntd.0005376.g007

Table 2. Total cases and incidence for the Global Burden of Disease regions and subregions made up of low- and middle-income countries. Total cases are shown in millions and incidence is per 100,000 person-years.

	Cases	Incidence
All LMICs	17.8 (6.9, 48.4)	293 (111, 794)
Central Europe, Eastern Europe, and Central Asia	0.1 (0.02, 0.6)	28 (7, 166)
Central Asia	0.05 (0.01, 0.5)	55 (12, 541)
Central Europe	0.01 (0.003, 0.06)	21 (4, 100)
Eastern Europe	0.03 (0.01, 0.13)	16 (4, 65)
Latin America and Caribbean	1.0 (0.2, 3.9)	169 (32, 642)
Andean Latin America	0.4 (0.04, 2.1)	704(80, 3751)
Caribbean	0.02 (0.004, 0.05)	47(12, 166)
Central Latin America	0.3 (0.07, 1.3)	120 (30, 512)
Southern Latin America	0.04 (0.01, 0.2)	61 (15, 276)
Tropical Latin America	0.2 (0.04, 1.1)	89 (18, 517)
North Africa and Middle East	2.6 (0.5, 5.7)	557 (100, 1208)
Sub-Saharan Africa	7.2 (2.2, 30.2)	762 (230, 3208)
Central Sub-Saharan Africa	1.7 (0.4, 8.4)	1459 (371, 6984)
Eastern Sub-Saharan Africa	2.4 (0.8, 11.3)	620(213, 2921)
Southern Sub-Saharan Africa	0.1 (0.04, 0.4)	149 (57, 571)
Western Sub-Saharan Africa	2.8 (0.7, 11.2)	753 (198, 3075)
Southeast Asia, East Asia, and Oceania	2.21 (0.7, 6.8)	108 (36, 334)
Southeast Asia	1.3 (0.4, 5.3)	217 (88, 571)
East Asia	0.5 (0.1, 1.7)	33 (9, 122)
Oceania	0.4 (0.03, 0.5)	5454 (397, 6576)
South Asia	3.6 (1.5, 9.4)	204 (64, 851)

doi:10.1371/journal.pntd.0005376.t002

Quantifying uncertainty

Our model allowed us to quantify the probability that each continent and its sub-regions had a total incidence that fell into one of four incidence categories: <10 cases per 100,000 person-years (“low incidence”), 10-<100 cases per 100,000 person-years (“medium incidence”), 100-<500 cases per 100,000 person-years (“high incidence”) and ≥500 cases per 100,000 person-years (“very high incidence”) (Table 3). All sub-regions of sub-Saharan Africa except Southern Sub-Saharan Africa have a high probability of being in the highest incidence category. North Africa and the Middle East have the second highest probability of being in the very high incidence category, while South Asia falls into the high incidence category. Latin America and the region of Southeast Asia, East Asia and Oceania have a 67% and 56% probability, respectively, of belonging in the high-incidence category, but geographic heterogeneity within both regions is considerable. While most parts of Latin America are likely to experience medium incidence, Central Latin America is most likely to experience high incidence, and Andean Latin America is most likely to experience very high incidence. Similarly, while East Asia is most likely to experience medium incidence, Southeast Asia is most likely to experience high incidence, and Oceania is most likely to experience very high incidence (Table 3 and Fig 8).

Discussion

We developed a meta-regression framework incorporating widely available indicators of economic and social development and the environment to estimate the incidence of typhoid fever

Table 3. Uncertainty in incidence estimates for the Global Burden of Disease regions and subregions made up of low- and middle-income countries.

	Percent of posterior sample in each incidence category			
	Low (<10 per 100,000 person-years)	Medium (10-<100 per 100,000 person-years)	High (100-<500 per 100,000 person-years)	Very high (≥500 per 100,000 person-years)
Central Europe, Eastern Europe, and Central Asia	7	87	6	<1
Central Asia	<1	77	20	3
Central Europe	14	83	2	<1
Eastern Europe	23	76	1	<1
Latin America and Caribbean	<1	27	67	7
Andean Latin America	<1	5	37	59
Caribbean	1	90	9	<1
Central Latin America	<1	40	57	3
Southern Latin America	<1	75	25	<1
Tropical Latin America	<1	57	40	3
North Africa and Middle East	<1	2	41	57
Sub-Saharan Africa	<1	<1	26	74
Central Sub-Saharan Africa	<1	<1	7	93
Eastern Sub-Saharan Africa	<1	<1	38	62
Southern Sub-Saharan Africa	<1	24	72	4
Western Sub-Saharan Africa	<1	<1	26	74
Southeast Asia, East Asia, and Oceania	<1	44	56	<1
Southeast Asia	<1	12	79	10
East Asia	3	92	4	<1
Oceania	<1	<1	4	96
South Asia	<1	5	91	4

doi:10.1371/journal.pntd.0005376.t003

across LMICs, as well as the concomitant uncertainty. We identified predictors that explain a substantial amount of heterogeneity in the incidence of typhoid fever, which significantly improved predictions of incidence across all age groups and for school-aged children in particular. This analysis represents substantial progress over existing models for the incidence of typhoid fever by allowing for estimation of variation in typhoid incidence both within and between countries. Given the limited data available, the credible intervals around the model predictions are appropriately large in parts of the world where typhoid surveillance is weak or non-existent. Although considerable uncertainty remains, an additional strength of our analysis is that we calculate the probability that incidence surpasses the criteria for low, medium, high, and very high incidence in each country, which could help guide policy in the face of uncertainty.

Our results provide insights into the likely predictive power of widely available risk factors for typhoid incidence; however, these should not be interpreted as providing inference on the causes of typhoid transmission. First, most of the covariate data is available at the national or sub-national administrative unit, potentially failing to capture the epidemiological dynamics that operate at smaller scales. Second, these indices might be adequate proxies for the causal factors that drive disease incidence in some locations, but not across all LMICs. For instance, although contaminated drinking water has been established as the vehicle of transmission in numerous outbreak investigations, the proportion of the population using piped water is not a particularly helpful metric for predicting typhoid incidence in LMICs compared to other

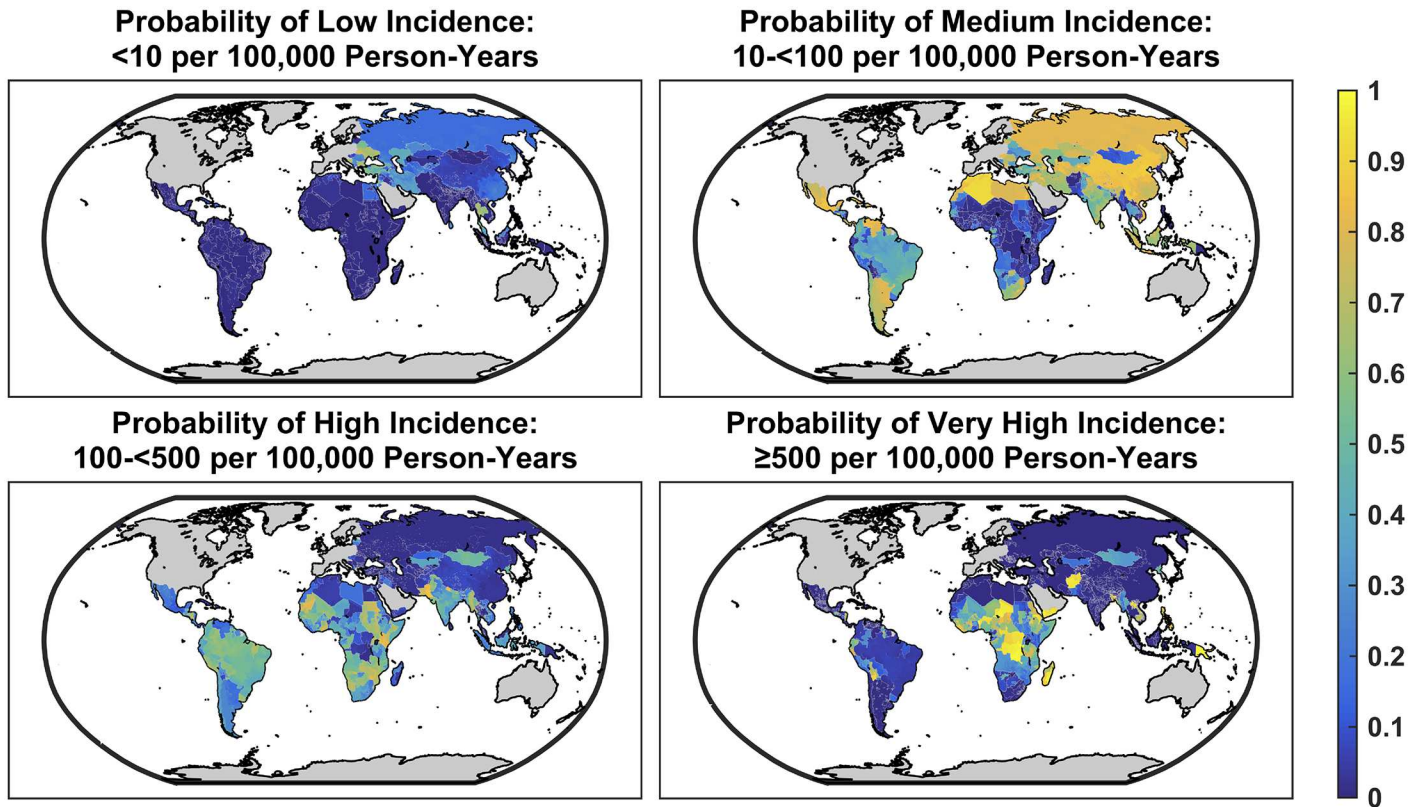


Fig 8. Probability that each location falls into one of four incidence categories: <10, 10-100, 100-500, and \geq 500 cases per 100,000 person-years, designated as low, medium, high, and very high incidence, respectively.

doi:10.1371/journal.pntd.0005376.g008

indicators of development and health, e.g. percent of roads paved, the population living in extreme poverty, or the prevalence of stunting and HIV [33–36]. The relatively weak association between typhoid incidence and the percent of the population with access to flush toilets or piped water may reflect that these indicators do not capture the microbial quality of water in the home [37]. Other estimates of typhoid incidence have relied on indicators of improved sources of water and sanitation in order to make assumptions about country-to-country and sub-national variation in typhoid fever incidence, but these studies performed limited or no assessment of the validity of these predictors for typhoid incidence [2,3]. Notably, no other study has examined the comparative utility of such a wide variety of indicators for predicting typhoid fever incidence.

Our estimate of the overall incidence of typhoid fever is similar to those published previously, but it reflects considerably greater uncertainty [1,2,12]. Previous studies assumed that each country experienced the mean incidence in its UN region [2,12] or median incidence in its GBD super region [1]. When we ran the model assuming regional hyperpriors for the location-specific random effects of the intercept and the slopes (rather than a single global hyperprior for the random effects), we found that regional hyperpriors did not differ significantly from each other or from the global hyperprior for the regional-level effects (S4 Table). Furthermore, none of the previous studies attempted to perform within-sample or out-of-sample validation [1,3,12,38]. Finally, it should also be noted that previous studies did not adjust for the apparent difference in incidence reported in passive versus “augmented passive” or active surveillance studies.

Past estimates of the incidence of typhoid fever in different age groups did not account for differential diagnostic procedures amongst age groups (in particular blood volume used for culture confirmation). It is important to consider the degree to which the observed age distribution of disease, particularly the lower incidence often observed in <5 year olds, could be attributed to the relationship between test sensitivity, the amount of blood drawn for diagnosis, and age. By simultaneously estimating the overall incidence, the incidence rate ratios between different age groups, and the observation process, we have allowed the data to drive our estimates of age-specific incidence, rather than relying on assumptions derived from a subset of studies.

However, we did not adjust for all differences in diagnostic procedures across studies. For example, while most studies limited eligible participants to those with a fever of three days or more, at least two of the studies enrolled all febrile children <5 years old regardless of the duration of fever [14,39]. Furthermore, blood cultures were carried out manually in older studies, whereas most recent studies have used an automated blood culture apparatus such as BacTec or BacT/ALERT that enhances the sensitivity of the culture to detect *S. Typhi*. The random-effect terms should account for at least some of these differences, as well as unmeasured differences in healthcare-seeking behavior, but were not factored into our model predictions, since such information is not available for all LMIC settings.

After carrying out a rigorous, statistically sound analysis of the available data on typhoid incidence, sizable uncertainty around our incidence estimates remains. The uncertainty in the incidence rate ratio (IRR) among <5 year olds and the covariance in the IRR between the <2 and 2–4 year olds outpaces the between and within-group variance in the other groups (S3 Table). Although numerous studies did not sample adults, the variance around the IRR for adults was the smallest of any age group. This indicates that future studies should focus on the incidence in children <5 years of age in order to refine the uncertainty around typhoid incidence estimates.

We strongly emphasize the need to consider the uncertainty in estimated incidence, in addition to the point estimates. Whereas past estimates designated countries into low, medium, and high incidence categories without any discussion on the potential for misclassification, our Bayesian framework allows us to present estimates of the probability that the incidence falls into each of four categories of incidence (Table 3, Fig 8). In addition, our stochastic search variable selection approach allowed us to incorporate both parameter and structural uncertainty in our estimates, which reflects our uncertainty not just in model coefficients but also in the combination of covariates that best explains incidence. We believe that this is a more honest appraisal of the extant data on typhoid incidence and its predictors, and that it could serve as a useful guide for policymakers on two matters: 1) what is the potential value of bolstering surveillance in order to refine our understanding of typhoid incidence, and 2) under the current level of uncertainty, what are the possible outcomes of different interventions?

Past estimates indicated that a large number of settings are predicted to fall into the high typhoid incidence category, previously defined as ≥ 100 cases per 100,000 person-years [2,40]. Our analysis suggests that this masks a formidable amount of the heterogeneity present in higher-incidence contexts; for instance, within Southeast Asia, East Asia, and Oceania, Southeast Asia is most likely to belong to the 100–<500 cases per 100,000 person-years incidence category, but Oceania is likely to have a higher incidence. We recommend further delineation of a “very high typhoid incidence” category to describe settings with an incidence of ≥ 500 per 100,000 person-years, which may motivate different or additional strategies for the control of typhoid fever in these settings. Our predictions also highlight potential within-country

variation in the incidence of typhoid, which can be useful to policymakers in developing control strategies targeted at particular regions of a country.

It is clear that the incidence of typhoid fever in Africa is not yet well understood. Out-of-sample validation of the model against data from nine TSAP sites showed that the model has mixed success in predicting incidence for locations outside the estimation sample. The incidence in the <5 year age group rural site in Moshi, Tanzania, and in both sites in Madagascar were all over-estimated. Model predictions for Kibera, Kenya and Ashanti Akim North, Ghana were more consistent with the observed data; however, there were previous studies from these two sites that were used to estimate the model. This suggests that there was some consistency in the incidence of typhoid over time, which our model was able to capture. Moreover, we draw attention to the potential levels of typhoid fever in regions where typhoid has received limited or no attention in the literature, such as Central Africa, Central Asia, and Oceania, and Latin America.

The current study modeled the long-term incidence (technically speaking, the mean annual incidence) of typhoid fever instead of predicting the number of cases in any given year. As we lack high-quality data in almost all LMIC settings, predicting year-to-year variation in the incidence of typhoid fever would be considerably more difficult, but would likely lead to even greater uncertainty in the incidence of typhoid in any given year. Due to limited surveillance capacity for typhoid fever over prolonged periods of time, our estimates of the spatial distribution of typhoid incidence reflect only the spatial distribution of risk factors, rather than temporal processes of spatial contagion suggested by phylogeographic studies [35,41].

Furthermore, we do not account for temporal variations in incidence associated to emergent properties of the pathogen or its transmission dynamics, such as the recent outbreaks of typhoid reported across eastern and southern Africa [42–44], possibly related to the emergence of the H58 haplotype [45]. A few studies in our sample were carried out in the same place in different years using similar surveillance protocols, and in these places, observed incidence was rarely significantly different from one year to another, with the exception of Dong Thap, Vietnam, where there was a small downward trend in incidence (Fig 4). While some hospital-based time series attest to the stability of typhoid incidence over periods >10 years [46–48], others highlight the potential for abrupt changes in incidence [48]. Our model assumes typhoid is endemic, and therefore may not perform as well in Africa, which appears to be more prone to epidemics of typhoid (as well as cholera) [49]. Rather, these incidence estimates should be interpreted as the potential endemic burden of typhoid under current conditions. Nevertheless, it is important to consider how variation in typhoid incidence over time and space may impact the design and implementation of optimal control strategies for typhoid fever.

Our analysis achieved two main objectives: 1) to identify widely available predictors of typhoid fever incidence; 2) to point out places in the world that have the most uncertainty in typhoid incidence, thereby motivating future studies into the scale and spatiotemporal distribution of typhoid fever in these areas. However, many LMICs have limited capacity for typhoid surveillance. The model we present provides a validated means of predicting typhoid incidence in countries with limited or no typhoid surveillance data based on widely available indicators. Understanding and predicting the burden of typhoid fever is an essential first step in motivating the need for better control measures, including typhoid conjugate vaccines.

Supporting information

S1 Text. Detailed methods.

(DOCX)

S1 Table. Incidence studies used to estimate the parameters of the prediction model. Incidence rates (and 95% confidence intervals) per 100,000 person-years. Incidence rates shown are not adjusted for participation rate, surveillance type, or blood culture sensitivity.
(DOCX)

S2 Table. Incidence studies used to validate our model. Incidence rates (and 95% confidence intervals) per 100,000 person-years. Incidence rates shown are not adjusted for participation rate, surveillance type, or blood culture sensitivity.
(DOCX)

S3 Table. Random effect variance-covariance matrix.
(DOCX)

S4 Table. Posterior distributions for a model with 2-level hyperpriors to estimate random effects. We ran a null model assuming regional hyperpriors for the location-specific random effects (in addition to a global hyperprior) using two schemes to group countries: continents and the Global Burden of Disease (GBD) Regions for 2015. We discovered that regional hyperpriors would not differ significantly from global hyperpriors. Further, we note that there are only data for two locations in the Americas, making a separate hyperprior for this region unnecessary.
(DOCX)

S1 Fig. Blood culture sensitivity. A) Studies that estimate blood culture sensitivity in bone-marrow culture-confirmed patients for typhoid fever [50–55]. B) Our predictions of sensitivity of blood cultures performed with a range of sample volume 2–15mL.
(EPS)

S2 Fig. Posterior distributions for intercept and slope coefficients. Distribution of the covariate coefficients for the intercept and each of the slopes are shown. The symbol size corresponds to the proportion of the models that included that covariate. The intercept describes the incidence rate for children 5–14 years of age, slope 1 describes the incidence rate ratio (IRR) between children <2 and 5–14, slope 2 describes IRR between children 2–4 and 5–14, and slope 3 describes the IRR between adults (≥ 15) and children 5–14 years of age.
(EPS)

S3 Fig. Observed versus model-predicted incidence using location-specific random effects. Posterior predictions from the model using fixed effects for the predictors and location- and age-specific random effects. The gray markers represent the density of model-predicted posterior distributions of incidence, while the red dots represent the median posterior predicted incidence. The size of the circular markers represents the number of person-years of observation in each study.
(EPS)

S4 Fig. Predictor selection under leave-k-out validation. To validate the model, we re-estimated the model seven times, each time leaving out the data from three locations to assess how well the model would estimate the incidence in those locations. Leave-3-out validation allowed us to evaluate the consistency with which spike-and-slab priors (our stochastic search variable selection algorithm) would include each variables as predictors of the intercept or of the intercept and the slopes.
(EPS)

S5 Fig. Model size under leave-k-out validation. Validation also allowed us to evaluate the posterior number of predictors selected for the model. Chain 1 was initiated using a model

that included all the covariates as predictors of the main effect, while chain 2 was initiated as the null model. The null model was never sampled, implying that the models including at least one predictor better described the data than the null model.

(EPS)

Acknowledgments

The authors would like to thank Prof. Jeroen Smits and Dr. Rutger Schilpzand of the Global Data Lab at Nijmegen Center for Economics (Radboud Universiteit Nijmegen) for providing assistance with extracting data on subnational socioeconomic and demographic indices.

Author Contributions

Conceptualization: MA VEP.

Data curation: MA GDP FM.

Formal analysis: MA JLW EK.

Funding acquisition: VEP.

Investigation: GDP FM.

Methodology: MA FWC DMW VEP.

Resources: VEP.

Software: MA.

Supervision: VEP.

Validation: MA GDP FM.

Visualization: MA VEP.

Writing – original draft: MA VEP.

Writing – review & editing: MA JLW FWC DMW EK GDP FM VEP.

References

1. Buckle GC, Walker CLF, Black RE. Typhoid fever and paratyphoid fever: Systematic review to estimate global morbidity and mortality for 2010. *J Glob Health*. 2012; 2: 10401.
2. Mogasale V, Maskery B, Ochiai RL, Lee JS, Mogasale V V, Ramani E, et al. Burden of typhoid fever in low-income and middle-income countries: a systematic, literature-based update with risk-factor adjustment. *Lancet Glob Heal*; 2014; 2: e570–80.
3. Lee J-S, Mogasale V V., Mogasale V, Lee K. Geographical distribution of typhoid risk factors in low and middle income countries. *BMC Infect Dis*. *BMC Infectious Diseases*; 2016; 16: 732. doi: [10.1186/s12879-016-2074-1](https://doi.org/10.1186/s12879-016-2074-1) PMID: [27919235](https://pubmed.ncbi.nlm.nih.gov/27919235/)
4. Parry CM, Hien TT, Dougan G, White NJ, Farrar JJ. Typhoid fever. *N Engl J Med*. 2002; 347: 1770–82. doi: [10.1056/NEJMr020201](https://doi.org/10.1056/NEJMr020201) PMID: [12456854](https://pubmed.ncbi.nlm.nih.gov/12456854/)
5. Bhan MK, Bahl R, Bhatnagar S. Typhoid and paratyphoid fever. *Lancet*. 2005; 366: 749–62. doi: [10.1016/S0140-6736\(05\)67181-4](https://doi.org/10.1016/S0140-6736(05)67181-4) PMID: [16125594](https://pubmed.ncbi.nlm.nih.gov/16125594/)
6. Dougan G, Baker S. Salmonella enterica serovar Typhi and the pathogenesis of typhoid fever. *Annu Rev Microbiol*. 2014; 68: 317–36. doi: [10.1146/annurev-micro-091313-103739](https://doi.org/10.1146/annurev-micro-091313-103739) PMID: [25208300](https://pubmed.ncbi.nlm.nih.gov/25208300/)
7. Bhutta ZA. Impact of age and drug resistance on mortality in typhoid fever. *Arch Dis Child*. 1996; 75: 214–217. PMID: [8976660](https://pubmed.ncbi.nlm.nih.gov/8976660/)
8. Parry CM, Beeching NJ. Epidemiology, diagnosis and treatment of enteric fever. *Curr Opin Infect Dis*. 1998; 11: 583–590. PMID: [17033425](https://pubmed.ncbi.nlm.nih.gov/17033425/)

9. Khan MI, Soofi SB, Ochiai RL, Khan MJ, Sahito SM, Habib MA, et al. Epidemiology, clinical presentation, and patterns of drug resistance of Salmonella Typhi in Karachi, Pakistan. *J Infect Dev Ctries. Italy*; 2012; 6: 704–714.
10. Sur D, von Seidlein L, Manna B, Dutta S, Deb AK, Sarkar BL, et al. The malaria and typhoid fever burden in the slums of Kolkata, India: data from a prospective community-based study. *Trans R Soc Trop Med Hyg*. 2006; 100: 725–33. doi: [10.1016/j.trstmh.2005.10.019](https://doi.org/10.1016/j.trstmh.2005.10.019) PMID: [16455118](https://pubmed.ncbi.nlm.nih.gov/16455118/)
11. Crump JA. Typhoid Fever and the challenge of nonmalaria febrile illness in sub-saharan Africa. *Clin Infect Dis*. 2012; 54: 1107–9. doi: [10.1093/cid/cis024](https://doi.org/10.1093/cid/cis024) PMID: [22357701](https://pubmed.ncbi.nlm.nih.gov/22357701/)
12. Crump JA, Luby SP, Mintz ED. The global burden of typhoid fever. *Bull World Health Organ*. 2004; 82: 346–53. PMID: [15298225](https://pubmed.ncbi.nlm.nih.gov/15298225/)
13. Lin F-YC, Ho VA, P Van Bay, Nguyen TTT, Bryla D, Thanh TC, et al. The epidemiology of typhoid fever in the Dong Thap Province, Mekong Delta region of Vietnam. *Am J Trop Med Hyg*. 2000; 62: 644–8. PMID: [11289678](https://pubmed.ncbi.nlm.nih.gov/11289678/)
14. Brooks WA, Hossain A, Goswami D, Sharmeen AT, Nahar K, Alam K, et al. Bacteremic typhoid fever in children in an urban slum, Bangladesh. *Emerg Infect Dis*. 2005; 11: 326–329. doi: [10.3201/eid1102.040422](https://doi.org/10.3201/eid1102.040422) PMID: [15752457](https://pubmed.ncbi.nlm.nih.gov/15752457/)
15. Siddiqui FJ, Rabbani F, Hasan R, Nizami SQ, Bhutta ZA. Typhoid fever in children: some epidemiological considerations from Karachi, Pakistan. *Int J Infect Dis*. 2006; 10: 215–22. doi: [10.1016/j.ijid.2005.03.010](https://doi.org/10.1016/j.ijid.2005.03.010) PMID: [16431148](https://pubmed.ncbi.nlm.nih.gov/16431148/)
16. Gridded Population of the World, Version 3 (GPWv3): National Identifier Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC); 2005.
17. Nordhaus WD. Geography and macroeconomics: new data and new findings. *Proc Natl Acad Sci U S A*. 2006; 103: 3510–7. doi: [10.1073/pnas.0509842103](https://doi.org/10.1073/pnas.0509842103) PMID: [16473945](https://pubmed.ncbi.nlm.nih.gov/16473945/)
18. World Bank. World Development Indicators. Washington, DC.: World Bank (producer and distributor); 2015.
19. Smits J. GDL Area Database. Nijmegen, The Netherlands; 2016. Report No.: 16–101.
20. WHO/UNICEF Joint Monitoring Program for Water and Sanitation [Internet]. 2016 [cited 1 Nov 2016]. <http://www.wssinfo.org/data-estimates/>
21. Adhikari P, Hong Y, Douglas KR, Kirschbaum DB, Gourley J, Adler R, et al. A digitized global flood inventory (1998–2008): Compilation and preliminary results. *Nat Hazards*. 2010; 55: 405–422.
22. UNAIDS AIDSInfo [Internet]. [cited 1 Nov 2016]. <http://aidsinfo.unaids.org>
23. Mekonnen MM, Hoekstra AY. Four billion people facing severe water scarcity. *Sci Adv*. 2016; 2: e1500323. doi: [10.1126/sciadv.1500323](https://doi.org/10.1126/sciadv.1500323) PMID: [26933676](https://pubmed.ncbi.nlm.nih.gov/26933676/)
24. Nichols C, Cruz Espinoza LM, Von Kalckreuth V, Aaby P, Ahmed El Tayeb M, Ali M, et al. Bloodstream infections and frequency of pretreatment associated with age and hospitalization status in Sub-Saharan Africa. *Clin Infect Dis*. 2015; 61: S372–S379. doi: [10.1093/cid/civ730](https://doi.org/10.1093/cid/civ730) PMID: [26449954](https://pubmed.ncbi.nlm.nih.gov/26449954/)
25. Marks F, Von Kalckreuth V, Aaby P, Adu-sarkodie Y, Ahmed M, Tayeb E, et al. Articles Incidence of invasive salmonella disease in sub-Saharan Africa: a multicentre population-based surveillance study. *Lancet Glob Health*. 2017; 5: e310–e323. doi: [10.1016/S2214-109X\(17\)30022-0](https://doi.org/10.1016/S2214-109X(17)30022-0). PMID: [28193398](https://pubmed.ncbi.nlm.nih.gov/28193398/).
26. Date KA, Bentsi-Enchill AD, Fox KK, Abeyasinghe N, Mintz ED, Khan MI, et al. Typhoid fever surveillance and vaccine use, South-East Asia and Western Pacific Regions, 2009–2013. *Morb Mortal Wkly Rep*. 2014; 63: 855–61.
27. Date KA, Bentsi-Enchill A, Marks F, Fox K. Typhoid fever vaccination strategies. *Vaccine*. Elsevier Ltd; 2015; 33: C55–C61.
28. George EI, McCulloch RE. Variable Selection via Gibbs Sampling. *Journal of the American Statistical Association*. 1993; 88: 881–889.
29. Ishwaran H, Rao JS. Spike and slab variable selection: Frequentist and Bayesian strategies. *Ann Stat*. 2005; 33: 730–773.
30. Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proc 3rd Int Work Distrib Stat Comput*. 2003; 1–10.
31. Steyvers M, Kalish M. MATJAGS, a Matlab interface for JAGS. 2014.
32. von Kalckreuth V, Konings F, Aaby P, Adu-Sarkodie Y, Ali M, Aseffa A, et al. The Typhoid Fever Surveillance in Africa Program (TSAP): Clinical, Diagnostic, and Epidemiological Methodologies. *Clin Infect Dis*. 2016; 62: S9–S16. doi: [10.1093/cid/civ693](https://doi.org/10.1093/cid/civ693) PMID: [26933028](https://pubmed.ncbi.nlm.nih.gov/26933028/)
33. Srikantiah P, Vafokulov S, Luby SP, Ishmail T, Earhart K, Khodjaev N, et al. Epidemiology and risk factors for endemic typhoid fever in Uzbekistan. *Trop Med Int Heal*. 2007; 12: 838–47.

34. Sur D, Ali M, von Seidlein L, Manna B, Deen JL, Acosta CJ, et al. Comparisons of predictors for typhoid and paratyphoid fever in Kolkata, India. *BMC Public Health*. 2007; 7: 289. doi: [10.1186/1471-2458-7-289](https://doi.org/10.1186/1471-2458-7-289) PMID: [17935611](https://pubmed.ncbi.nlm.nih.gov/17935611/)
35. Baker S, Holt KE, Clements ACA, Karkey A, Arjyal A, Boni MF, et al. Combined high-resolution genotyping and geospatial analysis reveals modes of endemic urban typhoid fever transmission. *Open Biol*. 2011; 1: 110008. doi: [10.1098/rsob.110008](https://doi.org/10.1098/rsob.110008) PMID: [22645647](https://pubmed.ncbi.nlm.nih.gov/22645647/)
36. Mermin JH, Villar R, Carpenter J, Roberts L, Samariddin a, Gasanova L, et al. A massive epidemic of multidrug-resistant typhoid fever in Tajikistan associated with consumption of municipal water. *J Infect Dis*. 1999; 179: 1416–1422. doi: [10.1086/314766](https://doi.org/10.1086/314766) PMID: [10228063](https://pubmed.ncbi.nlm.nih.gov/10228063/)
37. Bain R, Cronk R, Hossain R, Bonjour S, Onda K, Wright J, et al. Global assessment of exposure to faecal contamination through drinking water based on a systematic review. *Trop Med Int Heal*. 2014; 19: 917–927.
38. Mogasale V, Maskery B, Ochiai RL, Lee JS, Mogasale VV, Ramani E, et al. Revisiting the burden of typhoid fever in low and middle-income countries to inform policy decisions. *Am J Trop Med Hyg*. 2014; 1): 571.
39. Sinha A, Sazawal S, Kumar R, Sood S, Reddaiah VP, Singh B, et al. Typhoid fever in children aged less than 5 years. *Lancet*. 1999; 354: 734–7. doi: [10.1016/S0140-6736\(98\)09001-1](https://doi.org/10.1016/S0140-6736(98)09001-1) PMID: [10475185](https://pubmed.ncbi.nlm.nih.gov/10475185/)
40. Crump JA, Mintz ED. Global Trends in Typhoid and Paratyphoid Fever. *Clin Infect Dis*. 2010; 50: 241–6. doi: [10.1086/649541](https://doi.org/10.1086/649541) PMID: [20014951](https://pubmed.ncbi.nlm.nih.gov/20014951/)
41. Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella* Typhi identifies inter- and intracontinental transmission events. *Nat Genet*. 2015; 47: 632–639. doi: [10.1038/ng.3281](https://doi.org/10.1038/ng.3281) PMID: [25961941](https://pubmed.ncbi.nlm.nih.gov/25961941/)
42. Neil KP, Sodha S V, Lukwago L, O-tipo S, Mikoleit M, Simington SD, et al. A Large Outbreak of Typhoid Fever Associated With a High Rate of Intestinal Perforation in Kasese District, Uganda, 2008–2009. *Clin Infect Dis*. 2012; 54: 1091–9. doi: [10.1093/cid/cis025](https://doi.org/10.1093/cid/cis025) PMID: [22357703](https://pubmed.ncbi.nlm.nih.gov/22357703/)
43. Sejvar J, Lutterloh E, Naiena J, Likaka A, Manda R, Nygren B, et al. Neurologic manifestations associated with an outbreak of typhoid fever, Malawi -Mozambique, 2009: an epidemiologic investigation. *PLoS One*. 2012; 7: e46099. doi: [10.1371/journal.pone.0046099](https://doi.org/10.1371/journal.pone.0046099) PMID: [23226492](https://pubmed.ncbi.nlm.nih.gov/23226492/)
44. Gordon MA, Graham SM, Walsh AL, Wilson L, Phiri A, Molyneux E, et al. Epidemics of invasive *Salmonella enterica* serovar enteritidis and *S. enterica* Serovar typhimurium infection associated with multi-drug resistance among adults and children in Malawi. *Clin Infect Dis*. 2008; 46: 963–9. doi: [10.1086/529146](https://doi.org/10.1086/529146) PMID: [18444810](https://pubmed.ncbi.nlm.nih.gov/18444810/)
45. Pitzer VE, Feasey NA, Msefula C, Mallewa J, Kennedy N, Dube Q, et al. Mathematical Modeling to Assess the Drivers of the Recent Emergence of Typhoid Fever in Blantyre, Malawi. *Clin Infect Dis*. 2015; 61: S251–S258. doi: [10.1093/cid/civ710](https://doi.org/10.1093/cid/civ710) PMID: [26449939](https://pubmed.ncbi.nlm.nih.gov/26449939/)
46. Maskey AP, Basnyat B, Thwaites GE, Campbell JI, Farrar JJ, Zimmerman MD. Emerging trends in enteric fever in Nepal: 9124 cases confirmed by blood culture 1993–2003. *Trans R Soc Trop Med Hyg*. 2008; 102: 91–5. doi: [10.1016/j.trstmh.2007.10.003](https://doi.org/10.1016/j.trstmh.2007.10.003) PMID: [18023462](https://pubmed.ncbi.nlm.nih.gov/18023462/)
47. Pitzer VE, Bowles CC, Baker S, Kang G, Balaji V, Farrar JJ, et al. Predicting the impact of vaccination on the transmission dynamics of typhoid in South Asia: a mathematical modeling study. *PLoS Negl Trop Dis*. 2014; 8: e2642. doi: [10.1371/journal.pntd.0002642](https://doi.org/10.1371/journal.pntd.0002642) PMID: [24416466](https://pubmed.ncbi.nlm.nih.gov/24416466/)
48. Feasey NA, Gaskell K, Wong V, Msefula C, Selemeni G, Kumwenda S, et al. Rapid emergence of multi-drug resistant, H58-lineage *Salmonella typhi* in Blantyre, Malawi. *PLoS Negl Trop Dis*. United States; 2015; 9: e0003748.
49. Sauvageot D, Njanpop-Lafourcade BM, Akilimali L, Anne JC, Bidjada P, Bompangue D, et al. Cholera Incidence and Mortality in Sub-Saharan African Sites during Multi-country Surveillance. *PLoS Negl Trop Dis*. 2016; 10: 1–16.
50. Gilman RH, Termini M, Levine MM, Hernandez-Mendoza P, Hornick RB. Relative efficacy of blood, urine, rectal swab, bone-marrow, and rose-spot cultures for recovery of *Salmonella typhi* in typhoid fever. *Lancet*. 1975; 1: 1211–3. PMID: [48834](https://pubmed.ncbi.nlm.nih.gov/48834/)
51. Hoffman SL, Edman DC, Punjabi NH, Lesmana M, Cholid A, Sundah S, et al. Bone marrow aspirate culture superior to streptokinase clot culture and 8 ml 1:10 blood-to-broth ratio blood culture for diagnosis of typhoid fever. *Am J Trop Med Hyg*. 1986; 35: 836–9. PMID: [3089041](https://pubmed.ncbi.nlm.nih.gov/3089041/)
52. Gasem MH, Dolmans WM, Isbandrio BB, Wahyono H, Keuter M, Djokomoeljanto R, et al. Culture of *Salmonella typhi* and *Salmonella paratyphi* from blood and bone marrow in suspected typhoid fever. *Trop Geogr Med*. Amsterdam, Netherlands; 1995; 47: 164–167.
53. Guerra-Caceres JG, Gotuzzo-Herencia E, Crosby-Dagnino E, Miro-Quesada M, Carrillo-Parodi C. Diagnostic-Value of Bone-Marrow Culture in Typhoid-Fever. *Trans R Soc Trop Med Hyg*. 1979; 73: 680–683. PMID: [538809](https://pubmed.ncbi.nlm.nih.gov/538809/)

54. Wain J, Hosoglu S. The laboratory diagnosis of enteric fever. *J Infect Dev Ctries*. 2008; 2: 421–425. <http://dx.doi.org/10.3855/jidc.164> PMID: [19745517](https://pubmed.ncbi.nlm.nih.gov/19745517/)
55. Gasem MH, Smits HL, Goris MGA, Dolmans WM V. Evaluation of a simple and rapid dipstick assay for the diagnosis of typhoid fever in Indonesia. *J Med Microbiol*. 2002; 51: 173–177. doi: [10.1099/0022-1317-51-2-173](https://doi.org/10.1099/0022-1317-51-2-173) PMID: [11865843](https://pubmed.ncbi.nlm.nih.gov/11865843/)