View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Apollo

Received: 19 October 2020 | Revised: 23 July 2021 | Accepted: 27 July 2021

DOI: 10.1002/sim.9159

RESEARCH ARTICLE

Statistics in Medicine WILEY

A Bayesian approach for estimating typhoid fever incidence from large-scale facility-based passive surveillance data

Maile T. Phillips[1] | James E. Meiring[2,3,4] | Merryn Voysey[2,3] | Joshua L. Warren[5] | Stephen Baker[6] | Buddha Basnyat[7] | John D. Clemens[8] | Christiane Dolecek[9,10] | Sarah J. Dunstan[11] | Gordon Dougan[6] | Melita A. Gordon[4,12] | Deus Thindwa[4,13] | Robert S. Heyderman[4,14] | Kathryn E. Holt[15,16] | Firdausi Qadri[8] | Andrew J. Pollard[2,3] | Virginia E. Pitzer[1] | the STRATAA Study Group

[1]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, Connecticut, USA

[2]Oxford Vaccine Group, Department of Paediatrics, University of Oxford, Oxford, UK

[3]National Institute for Health Research Oxford Biomedical Research Centre, Oxford, UK

[4]Malawi Liverpool Wellcome Programme, Kamuzu University of Health Sciences, Blantyre, Malawi

[5]Department of Biostatistics, Yale School of Public Health, New Haven, Connecticut, USA

[6]Department of Medicine, University of Cambridge, Cambridge, UK

[7]Oxford University Clinical Research Unit, Patan Academy of Health Sciences, Kathmandu, Nepal

[8]International Centre for Diarrhoeal Diseases Research, Bangladesh (icddr,b), Dhaka, Bangladesh

[9]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK

[10]Mahidol Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand

[11]The Peter Doherty Institute for Infection and Immunity, The University of Melbourne, Melbourne, Victoria, Australia

Abstract

Decisions about typhoid fever prevention and control are based on estimates of typhoid incidence and their uncertainty. Lack of specific clinical diagnostic criteria, poorly sensitive diagnostic tests, and scarcity of accurate and complete datasets contribute to difficulties in calculating age-specific population-level typhoid incidence. Using data from the Strategic Typhoid Alliance across Africa and Asia program, we integrated demographic censuses, healthcare utilization surveys, facility-based surveillance, and serological surveillance from Malawi, Nepal, and Bangladesh to account for under-detection of cases. We developed a Bayesian approach that adjusts the count of reported blood-culture-positive cases for blood culture detection, blood culture collection, and healthcare seeking—and how these factors vary by age—while combining information from prior published studies. We validated the model using simulated data. The ratio of observed to adjusted incidence rates was 7.7 (95% credible interval [CrI]: 6.0-12.4) in Malawi, 14.4 (95% CrI: 9.3-24.9) in Nepal, and 7.0 (95% CrI: 5.6-9.2) in Bangladesh. The probability of blood culture collection led to the largest adjustment in Malawi, while the probability of seeking healthcare contributed the most in Nepal and Bangladesh; adjustment factors varied by age. Adjusted incidence rates were within or below the seroincidence rate limits of typhoid infection. Estimates of blood-culture-confirmed typhoid fever without these adjustments results in considerable underestimation of the true incidence of typhoid fever. Our approach allows each phase of the reporting process to

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Statistics in Medicine* published by John Wiley & Sons Ltd.

[12]Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK

[13]Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK

[14]NIHR Global Health Research Unit on Mucosal Pathogens, Division of Infection and Immunity, University College London, London, UK

[15]Department of Infectious Diseases, Central Clinical School, Monash University, Melbourne, Victoria, Australia

[16]Department of Infection Biology, London School of Hygiene and Tropical Medicine, London

Correspondence
Maile T. Phillips, Department of Epidemiology of Microbial Diseases, Yale School of Public Health, P.O. Box 208034, 60 College St., New Haven, CT 6520-8034, USA.
Email: maile.phillips@yale.edu

Funding information
Bill and Melinda Gates Foundation, Grant/Award Number: OPP1141321; Public Health Research Programme; Wellcome Trust, Grant/Award Numbers: 106158/Z/14/Z, 206545/Z/17/Z; UK National Institute for Health Research (NIHR) Research Professorship in GlobalHealth, Grant/Award Number: NIHR300039

be synthesized to estimate the adjusted incidence of typhoid fever while correctly characterizing uncertainty, which can inform decision-making for typhoid prevention and control.

KEYWORDS
incidence estimation, passive surveillance, reporting pyramid, typhoid fever

1 | BACKGROUND

Current estimates of typhoid fever incidence serve as a basis for decision-making around typhoid control. However, facility-based cases of blood-culture-confirmed typhoid fever are considerably lower than the true number of those with the disease[1] because the reported numbers do not account for individuals with typhoid fever who do not seek healthcare, fail to receive a diagnostic test, or falsely test negative for typhoid (Figure 1). Annually, typhoid fever is estimated to cause 11 to 18 million infections and 100 000 to 200 000 deaths,[2-4] but there is considerable uncertainty in these estimates.

Studies suggest that somewhere between 60% and 90% individuals with typhoid fever do not receive adequate medical attention, in part because they do not to seek formal treatment.[5,6] Previous studies have found that healthcare utilization is correlated with the number of household members, distance to the healthcare facility, financial affordability, and trust in formal healthcare.[7,8] Furthermore, typhoid fever is often misdiagnosed based on physical examinations alone.[1] Inconsistent clinical diagnoses arise because symptoms of typhoid, particularly prolonged fever, are also the main characteristics of other common infectious diseases in typhoid endemic settings.[1,9] Even if a blood culture test is recommended and laboratory facilities are available, not all patients will consent. Diagnostic tests can be invasive, and parents or guardians of young children sometimes do not want their children to have large amounts of blood drawn when they are already ill. In resource-poor countries in particular, lack of supplies and personnel lead to long wait times for receiving healthcare, further adding to lower rates of confirmatory testing. Clinical opinion on the cause of fever can also affect the likelihood of blood being drawn for culture.[10]

Suboptimal diagnostic tests further contribute to underestimation of cases. Blood culture collection is the mainstay diagnostic test for typhoid fever,[11] but it fails to capture approximately half of the true cases. The test sensitivity depends

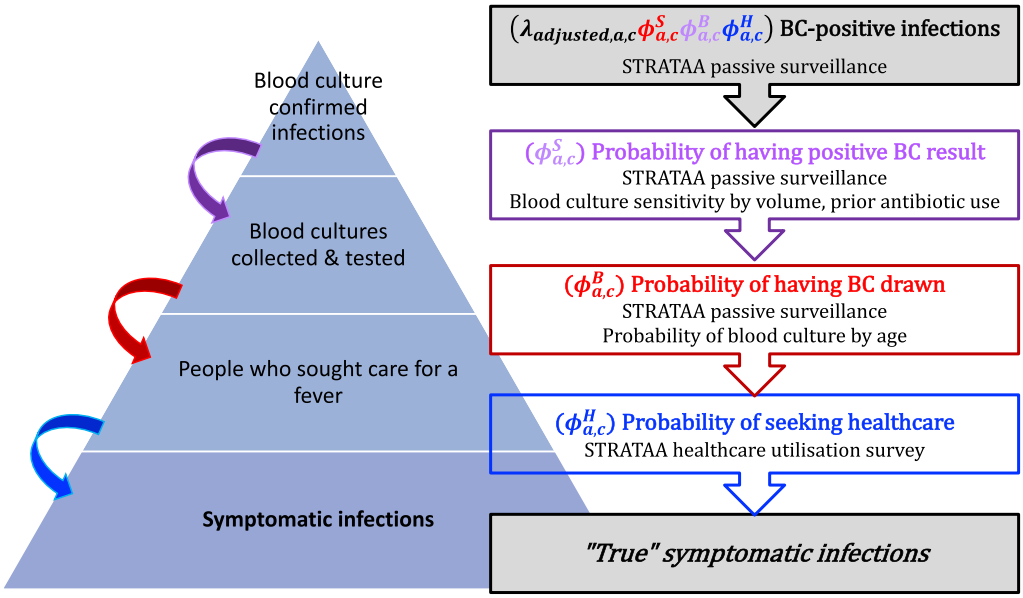

FIGURE 1 Flowchart of typhoid disease and observation process, and adjustment method to estimate the true number of cases. The pyramid (left) illustrates the different steps in the observation process for reporting typhoid incidence, with details on how parameters are estimated at each step. The flowchart (right) illustrates the corresponding Bayesian framework for each step of the observation process and which datasets and variables are used for adjustment. Adjustments for blood culture sensitivity are shown in purple, the probability of receiving a blood culture test is shown in red, and the probability of seeking healthcare is shown in blue. Variable definitions: λ, typhoid incidence rate; ϕ, a probability estimated in the model; S, sensitivity of blood culture; B, blood culture collection; H, healthcare seeking; a, age category; c, site. Abbreviation: BC, blood culture [Colour figure can be viewed at wileyonlinelibrary.com]

on the volume of blood drawn and whether a patient has recently received antibiotics.[12] Thus, even if an individual with typhoid receives a blood culture test, he or she may falsely test negative and not be included in the reported number of confirmed cases.

The true incidence of typhoid fever cannot be directly assessed but can be estimated by accounting for steps in the reporting process. Methods to combine data from several sources to adjust for underestimation while accurately quantifying the uncertainty have been previously applied to estimate the incidence of HIV and influenza.[13-17] Bayesian methods are conducive to integrating multiple data sources in this way. In this work, we developed a Bayesian multiplier framework to estimate population-based incidence of typhoid fever based on data collected from study sites in Africa and Asia.

2 | METHODS

2.1 | Study design and data

We developed a framework within the Bayesian setting to integrate data from multiple sources to estimate the population-based incidence of typhoid fever based on passive surveillance in Malawi, Nepal, and Bangladesh, three typhoid-endemic countries with different demographics, healthcare systems, and access to diagnostics.[18] Using this model, we sought to estimate the adjustment factors needed to calculate the "true" incidence of typhoid fever occurring in the population under surveillance, and to examine how these values varied by age across the three study sites, by combining information collected from the study population with estimates from prior published studies. We also compared our final estimates to serosurveillance data collected from the same population catchment areas.

Data came from the Strategic Typhoid Alliance across Africa and Asia (STRATAA) program, a prospective observational, population-based epidemiological study of typhoid incidence, transmission, and antibiotic resistance. From 2016 to 2018, the STRATAA investigators conducted demographic censuses, healthcare utilization surveys (HUS), passive surveillance, and serosurveys at each of three sites (Blantyre, Malawi; Kathmandu, Nepal; and Dhaka, Bangladesh). STRATAA's study design and methods have been detailed elsewhere,[18] and are briefly described below.

2.1.1 | Demographic census data

The demographic census was used to estimate the overall person-time contribution for incidence rate calculations. The survey documented household locations and individual characteristics for each geographically demarcated study area. The census provided information on each individual's birthdate, sex, position in the household, marital status (if applicable), education level, and employment status (if applicable). Participants were surveyed and consented as households. Approximately 100 000 individuals were enrolled at each site, and census updates were carried out one to three times depending on the site. Over the 2-year study period, this population amounted to 200 018 person-years of observation (pyo) in Malawi, 203 444 in Nepal, and 222 636 in Bangladesh.

2.1.2 | Passive surveillance

Clinical cases of culture-confirmed typhoid fever were identified through passive surveillance. Individuals living in the study areas who presented at partner facilities with a documented temperature of ≥38.0°C or a history of fever lasting at least 2 days upon presentation were eligible for enrollment. Healthcare workers collected clinical and demographic information from enrolled febrile individuals, as well as microbiological samples (urine, feces, and blood).

2.1.3 | Healthcare utilization survey

To estimate the proportion of cases captured by study facilities, we collected data from the head of household regarding actual and hypothetical usage of healthcare facilities for febrile episodes among household members from each of three age groups (<5 years, 5-14 years, >14 years), if available. Approximately 735 households were randomly chosen at each site, with the requirement that all households have at least one child (14 years or younger). The HUS contained questions regarding household and individual health behavior; house and household characteristics; water, sanitation and hygiene practices; and healthcare utilization (actual and hypothetical).

2.1.4 | Serosurveillance

Serosurveys were conducted in the census population to assess the underlying rate of seroconversion to typhoid, and to identify potential chronic carriers, initially based on anti-Vi immunoglobulin G (IgG). Approximately 8500 participants from each site were randomly selected in an age-stratified manner. Healthcare workers collected serum samples from each individual upon enrollment and again 3 months later. Seroconversion was defined as a ≥2-fold rise in anti-Vi IgG titre between the first and second sample drawn and an absolute titre ≥50 EU/mL at the second time point to account for small variations above the lower limit of detection for the assay. We estimated the seroincidence by dividing the number of seroconversions by the person-time contribution between serum samples in each age group; 95% binomial confidence intervals (CI) were estimated.

2.2 | Approach and data analysis

We adjusted the reported number of blood-culture-positive cases of typhoid fever for each of the three phases of the reporting process: blood culture detection, blood culture collection, and healthcare seeking. We used information on healthcare seeking for fever and the proportion of fever cases who were enrolled and had blood taken for culture and combined this information with published data on risk factors for typhoid fever to reach our final adjusted numbers. We then examined whether our adjusted estimates were within the maximum range expected based on seroincidence estimates, which capture all clinical and sub-clinical cases.

The estimation of typhoid fever incidence is complicated by the relationship between fever and typhoid fever. In each phase of the reporting process, we can observe whether a person has a fever, but not necessarily whether he/she has typhoid fever. Thus, the symptomatic typhoid fever pyramid is nested within a larger fever pyramid (Figure S1).

We assumed that the observed number of blood-culture-positive individuals ($n_{\text{observed},a,c}$) follows a Poisson distribution given as

$$n_{\text{observed},a,c} \sim \text{Poisson}\left(\frac{\lambda_{\text{observed},a,c}}{\text{person_time}_{a,c}}\right) \tag{1}$$

where $\lambda_{\text{observed},a,c}$ is the observed incidence rate of typhoid fever, adjusted for pyo from the demographic census ($\text{person_time}_{a,c}$) (Figure 1, Table 1). Subscript a represents age category, where

$$a = \begin{cases} 1 \text{ for children} < 5 \text{ years old,} \\ 2 \text{ for children 5 to 9 years old,} \\ 3 \text{ for children 10 to 14 years old,} \\ 4 \text{ for individuals 15 to 29 years old,} \\ 5 \text{ for individuals} \geq 30 \text{ years old,} \end{cases}$$

and subscript c represents the site, where

$$c = \begin{cases} 1 \text{ for Malawi,} \\ 2 \text{ for Nepal,} \\ 3 \text{ for Bangladesh.} \end{cases}$$

The incidence rate of reported cases can be rewritten as the product of the healthcare-seeking typhoid incidence rate and the probability of a case being captured at each of the three steps of the reporting process,

$$\lambda_{\text{observed},a,c} = \lambda_{\text{adjusted},a,c} * \phi^S_{a,c} * \phi^B_{a,c} * \phi^H_{a,c}, \tag{2}$$

where $\lambda_{\text{adjusted},a,c}$ is the final adjusted incidence of typhoid fever incidence, $\phi^S_{a,c}$ is blood culture sensitivity, $\phi^B_{a,c}$ is the probability of receiving a blood culture test, and $\phi^H_{a,c}$ is the probability of seeking healthcare for typhoid fever. Superscript S denotes that the parameter refers to the blood culture detection (ie, sensitivity) phase of reporting, superscript B denotes that the parameter is referring to the blood culture collection (ie, testing) phase of reporting, and superscript H denotes that the parameter is referring to the healthcare-seeking phase of reporting.

All model input parameters, with their prior distributions and corresponding data sources, are listed in Table 1. The three steps in the adjustment process (blood culture detection, blood culture collection, and healthcare seeking), are detailed below.

2.2.1 | Adjustment for blood culture sensitivity ($\phi^S_{a,c}$)

For each individual who received a blood culture test, we inferred whether or not they were a "true" typhoid fever case by adjusting for the specificity and sensitivity of blood culture for typhoid diagnosis (Table S1). First, we assumed that the specificity of blood culture is 100%; thus, all individuals who tested positive for typhoid were assumed to be true cases of the disease. Second, we assumed that among those who tested negative, the probability of being an actual typhoid fever case depended on the volume of blood drawn and prior antimicrobial use, both of which were recorded in the passive surveillance data. Previous studies have shown that the sensitivity of blood culture for typhoid diagnosis is on average 59% (95% CI: 54%-64%) but increases by 3% for each additional mL of blood drawn, and decreases by 34% with antibiotic use in the past 2 weeks.[12] Thus, each individual who tested negative for typhoid had a probability $1 - w_{i,v(i),u(i)}$ of being a false negative and thus a true case of typhoid fever, where $w_{i,v(i),u(i)}$ is the blood culture sensitivity for individual i, $v(i)$ is the volume of blood collected from individual i, and $u(i) = 1$ if individual i reported prior antibiotic use in the previous 2 weeks and $u(i) = 0$ otherwise. Blood culture sensitivity is defined as

$$w_{i,v(i),u(i)} = (e^{-0.73+0.03v_{(i)}}) * (1 - 0.34u_{(i)}). \tag{3}$$

TABLE 1 Model input parameters

Symbol	Description	Uncertainty distribution	Data source
Main model (Equations 1 and 2)			
$\lambda_{\text{observed}}$	Observed incidence rate of typhoid fever among febrile individuals who sought care	$n_{\text{observed}} \sim \text{Poisson}\left(\frac{\lambda_{\text{observed}}}{\text{person_time}}\right)$	See rest of table
n_{observed}	Number of blood-culture-positive individuals	Observed directly	STRATAA PS
person_time	Person-years of observation over the 2-year study period	Observed directly	STRATAA Demographic Census
$\lambda_{\text{adjusted}}$	The incidence rate of typhoid fever after adjusting for all three phases of reporting (blood culture sensitivity, the probability of receiving a blood culture test, and the probability of seeking healthcare)	$\lambda_{\text{observed}} = \lambda_{\text{adjusted}} * \phi^S * \phi^B * \phi^H$	See rest of table
Adjustment for blood culture sensitivity (Equations 3 and 4)			
ϕ^S	Blood culture sensitivity	$\phi^S \mid U \sim N(m_u, d_u)$	See rest of table
U	An indicator variable for whether or not an individual took antibiotics in the past 2 weeks	$U \sim \text{Bernoulli}(r)$	See rest of table
r	The proportion of individuals who took antibiotics in the past 2 weeks	Observed directly	STRATAA PS
m_u	The mean blood culture sensitivity for those who did and did not take antibiotics the past 2 weeks	Calculated using observed data with adjustment from Antillon et al	STRATAA PS, Antillon et al[12]
d_u	The SD of blood culture sensitivity for those who did and did not take antibiotics the past 2 weeks	Calculated using observed data with adjustment from Antillon et al	STRATAA PS, Antillon et al[12]
Adjustment for the probability of receiving a blood culture test (Equations 5a and 5b)			
ϕ^B	*Malawi:* Probability of receiving a blood culture test	$\phi^B \sim \text{Beta}(g^B, j^B)$	See rest of table
	Nepal and Bangladesh: Probability of receiving a blood culture test after adjusting for variation in risk of typhoid fever among those who are and are not tested	$\phi^B = 1 + \frac{(1-k)}{R_S}$	See rest of table
g^B	*Malawi:* Number of people who presented to a STRATAA facility with fever, were enrolled, and received a blood culture test	Observed directly	STRATAA PS
	Nepal and Bangladesh: Number of people who presented to a TyVAC facility with fever, were enrolled, and received a blood culture test	Observed directly	TyVAC[19]
j^B	*Malawi:* Number of people who presented to a STRATAA facility with fever but were not enrolled	Observed directly	STRATAA PS
	Nepal and Bangladesh: number of people who presented to a TyVAC facility with fever but were not enrolled	Observed directly	TyVAC[19]

TABLE 1 Continued

Symbol	Description	Uncertainty distribution	Data source
k	*Nepal and Bangladesh:* probability of receiving a blood culture test prior to adjusting for the relative risk of blood culture positivity	$k \sim \text{Beta}(g^B, j^B)$	See rest of table
R_S	*Nepal and Bangladesh*: Relative risk of blood culture positivity among those who received a blood culture test compared to those who did not	$\log(R_S) \sim N(\log(1.87), 0.37)$	10
Adjustment for the probability of healthcare seeking (Equations 6-8)			
ϕ^H	Probability of seeking healthcare	$\phi^H = \frac{\lambda^{S,B}}{\lambda_{\text{adjusted}}}$	See rest of table
$\lambda^{S,B}$	The incidence rate of typhoid fever after adjusting for blood culture sensitivity and the probability of receiving a blood culture test	$\lambda^{S,B} = \frac{\lambda_{\text{observed}}}{\phi^S * \phi^B}$	See rest of table
λ_0	Incidence of typhoid fever among those without the risk factor	$\lambda_0 = \frac{\lambda^{S,B}}{(h_1 R_{TF} p + h_0(1-p))}$	See rest of table
R_{TF}	Relative risk for typhoid fever among those with the risk factor compared to those without it	$\log(R_{TF}) \sim N(m_{TF}, d_{TF})$ *Malawi:* mean $m_{TF} = \log(0.5)$; standard deviation $d_{TF} = 0.34$ *Nepal:* $m_{TF} = \log(5.71)$; $d_{TF} = 0.47$ *Bangladesh:* $m_{TF} = \log(7.6)$; $d_{TF} = 0.64$	20-22
h_0	Probability of self-reported healthcare seeking for a fever among those without the risk factor	$h_0 \sim \text{Beta}(l_{h_0}, q_{h_0})$	See rest of table
l_{h_0}	Number of individuals who sought care out of those without the risk factor	Observed directly	STRATAA HUS
q_{h_0}	Number of individuals who did not seek care out of those without the risk factor	Observed directly	STRATAA HUS
h_1	Probability of self-reported healthcare seeking for a fever among those with the risk factor	$S_1 \sim \text{Beta}(l^{S_1}, q^{S_1})$	See rest of table
l_{h_1}	Number of individuals who sought care out of those with the risk factor	Observed directly	STRATAA HUS
q_{h_1}	Number of individuals who did not seek care out of those with the risk factor	Observed directly	STRATAA HUS
p	Probability of having the self-reported risk factor. In Malawi this was soap available after defecation, in Nepal this was unshared toilets, and in Bangladesh this was boiled drinking water	$p \sim \text{Beta}(l_p, q_p)$	See rest of table
l_p	Number of individuals with the identified typhoid fever risk factor	Observed directly	STRATAA HUS
q_p	Number of individuals without the identified typhoid fever risk factor	Observed directly	STRATAA HUS

Note: Parameters used in the incidence adjustment model are described below, with their corresponding uncertainty distributions and data sources. The symbols appear in the table in the order that they appear in the text, organized by the steps in the reporting process (main model, adjustment for blood culture sensitivity, adjustment for the probability of receiving a blood culture test, and adjustment for healthcare-seeking). All parameters are both age- and country-specific, except for p and R_{TF}, for which only a country-level estimate was available. Note that some of parameters in the adjustment for the probability of receiving a blood culture test are specific to only some countries (ϕ^B, g^B, j^B, k^B, and R_S).
Abbreviations: HUS, Healthcare Utilization Survey; PS, passive surveillance.

The use or non-use of antibiotics created a bimodal distribution for the sensitivity based on the observed individual-level data calculated using Equation (3) (Figure S2). Thus, we chose to model the distribution of sensitivity among the full population using a normal mixture model with a separate mean and SD for the distribution of blood culture sensitivity with and without the use of antibiotics, varied over blood culture volume. This mixture model is denoted

$$\begin{aligned} &\phi^S_{a,c} \mid U_{a,c} \sim \mathrm{N}(m_{u,a,c}, d_{u,a,c}), \\ &U_{a,c} \sim \mathrm{Bernoulli}(r_{a,c}), \end{aligned} \tag{4}$$

where $\phi^S_{a,c}$ is the blood culture sensitivity adjustment from Equation (2), $r_{a,c}$ is the proportion of individuals who took antibiotics in the past 2 weeks, and $m_{u,a,c}$ and $d_{u,a,c}$ represent the mean and SD of the distribution of blood culture sensitivity after adjusting for blood culture volume among those who did ($U = 1$) and did not ($U = 0$) take antibiotics (from the distribution created using Equation 3), again estimated separately for each age category and site.

2.2.2 | Adjustment for the probability of receiving a blood culture test ($\phi^B_{a,c}$)

The probability of receiving a blood culture test was estimated differently for Malawi vs Nepal and Bangladesh due to data availability and differences in the primary reasons why individuals were not tested. In Malawi, the main reason why individuals meeting the fever criteria for enrollment did not receive a blood culture test was due to limited capacity and long waiting times at the primary health facility. Individuals often arrived at the clinic early in the morning for clinical review. Once they had seen the government clinician, they were referred for study enrollment and blood-culture collection. If there was a delay in enrollment activities, individuals often left prior to blood cultures being collected. Thus, we assumed that data for those who did not receive a blood culture test were missing completely at random. We used the passive surveillance screening data to estimate the probability of receiving a blood culture test in Malawi, and assumed that the probability followed a beta distribution,

$$\phi^B_{a,1} \sim \mathrm{Beta}(g^B_{a,1i}, j^B_{a,1}), \tag{5a}$$

where $g^B_{a,1}$ was the number of eligible patients enrolled and $j^B_{a,1}$ was the number of eligible patients who were not enrolled.

In Nepal and Bangladesh, the primary reason febrile individuals did not receive a blood culture test likely depended on factors associated with their probability of testing positive (eg, age, number of days of fever, and clinical suspicion of the disease); furthermore, screening data for the passive surveillance were not available. Instead, we relied on published estimates of the relative risk of typhoid fever among those who did not have blood taken for culture and screening data from the Typhoid Vaccine Acceleration Consortium (TyVAC).[19] As part of TyVAC, typhoid conjugate vaccine trials are being conducted in the same populations as STRATAA utilizing the same passive surveillance facilities. Baseline information on eligible patients presenting to fever surveillance facilities was recorded both for those who did and did not have blood drawn for culture. Based on the analysis of these data in the TyVAC population in Nepal, the relative risk of blood culture positivity (R_S) was 1.87 times higher (95% CI: 0.9-3.9) among those who received a blood culture test compared to those who did not.[10] Thus, the overall probability of receiving a blood culture test, after taking into account variations in the risk of typhoid fever among those who are and are not tested, is

$$\phi^B_{a,c} = 1 - \frac{(1 - k_{a,c})}{R_S} \quad \text{for } c \in \{2, 3\}, \tag{5b}$$

where $k_{a,c}$ is the probability of receiving a blood culture test prior to adjusting for the relative risk of blood culture positivity. Prior to adjusting for the relative risk, the probability of receiving a blood culture test was assumed to follow a beta distribution, $k_{a,c} \sim \mathrm{Beta}(g^B_{a,c}, j^B_{a,c})$, where $g^B_{a,c}$ was the number of eligible patients in age category a in country c who were enrolled and $j^B_{a,c}$ was the number of eligible patients who were not enrolled during TyVAC (for $c = 2$ or 3). While the previous analysis focused only on Nepal, we used the same adjustment for the relative risk of blood culture positivity in Bangladesh, since the reasons for having or not having blood drawn were similar.

2.2.3 | Adjustment for healthcare-seeking behavior ($\phi^H_{a,c}$)

Previous multiplier methods assume that reported healthcare-seeking for a fever is the same as that for typhoid fever; however, this is not necessarily the case. Individuals with typhoid fever may be more or less likely to seek healthcare. In preliminary analyses, we found no difference in reported healthcare seeking by severity of a person's reported fever, but other factors may explain differential healthcare seeking among those with typhoid fever vs fever due to other causes. To correct for this difference, we measured the probability of seeking care for a fever adjusted for a specified typhoid risk factor to estimate the probability of seeking care for typhoid fever (Figure S1), as described below.

For this phase of reporting, we assumed that everyone in the population either had or did not have a typhoid risk factor, identified from the literature. For each site, we used a different risk factor, based on studies carried out in that specific site and variables for which data was collected as part of STRATAA. The risk factor identified in Malawi was soap available after defecation[20]; in Nepal, it was unshared toilets,[21] and in Bangladesh it was boiled drinking water.[22]

In Malawi, the odds ratio for having typhoid fever was 2.0 (95% CI: 1.3-2.5) comparing those who did not use soap after defecation to those who did.[20] In Nepal, the odds ratio for having typhoid fever was 5.7 (2.3-14.4) comparing those who used a household latrine to those who shared a community latrine.[21] In Bangladesh, the odds ratio for having typhoid was 7.6 (2.2-26.5) comparing those who did not boil drinking water to those who did.[22] Since the overall prevalence of typhoid in the population is low, we used these odds ratio estimates from the literature to approximate relative risks for typhoid. We used the same relative risk estimates for all age groups. All other values for the healthcare-seeking adjustment were estimated separately by age and site, as noted below.

We can calculate the marginal probability of seeking care for a fever among those with typhoid fever (alternatively, the incidence of typhoid after adjusting for blood culture sensitivity and the probability of receiving a blood culture test) as:

$$\lambda^{S,B}_{a,c} = \lambda_{0,a,c}(h_{1,a,c}R_{TF,c}p_{a,c} + h_{0,a,c}(1 - p_{a,c})), \tag{6}$$

where $p_{a,c}$ is the probability of having the risk factor for typhoid fever among individuals in age group a in site c, $\lambda_{0,a,c}$ is the incidence of typhoid fever among those without the risk factor, $R_{TF,c}$ is the relative risk for typhoid among those with the risk factor (and hence $R_{TF,c}\lambda_{0,a,c}$ is the incidence among those with the risk factor), and $h_{1,a,c}$ and $h_{0,a,c}$ are the probability of self-reported healthcare seeking for a fever among those with and without the risk factor, respectively.

We can also estimate the overall adjusted incidence of typhoid as the weighted average of the population-based typhoid fever incidence (after adjusting for blood culture sensitivity, the probability of receiving a blood culture test, and healthcare-seeking) among those with or without the typhoid fever risk factor:

$$\lambda_{\text{adjusted},a,c} = R_{TF,c}\lambda_{0,a,c}p_{a,c} + \lambda_{0,a,c,}(1 - p_{a,c}). \tag{7}$$

Combining Equations (6) and (7), we can estimate the standardized risk of seeking healthcare for typhoid fever ($\phi^H_{a,c}$) as:

$$\phi^H_{a,c} = \frac{\lambda^{S,B}_{a,c}}{\lambda_{\text{adjusted},a,c}} = \frac{h_{1,a,c}R_{TF,c}p_{a,c} + h_{0,a,c}(1 - p_{a,c})}{R_{TF,c}p_{a,c} + (1 - p_{a,c})}. \tag{8}$$

The numbers of individuals with the risk factor and who sought healthcare for a fever were observed directly in a sample of the population in the HUS. We assumed that the probabilities for the occurrence of these numbers (p, h_0, h_1) followed an underlying beta distribution based on the observed values (Table 1).

2.2.4 | Model validation and sensitivity analyses

The final adjusted incidence of symptomatic typhoid fever should be less than or equal to the seroincidence of typhoid infections captured in the serosurveillance data. In this study, seroconversion to typhoid was defined as a ≥2-fold rise in anti-Vi IgG titre between the first and second sample drawn and an absolute titre ≥50 EU/mL at the second time point. The seroincidence was estimated as the quotient of the number of people who seroconverted between the first and second blood sample and person-time in years (the number of people sampled multiplied by the mean time between serological samples in age group a in country c). Since seroincidence includes asymptomatic as well as symptomatic cases,

it is expected to be higher than the adjusted (symptomatic) incidence. As a result, the final adjusted incidence estimates should not exceed the upper bound of the estimated seroincidence rates.

To ensure the model was correctly formulated, we simulated data with known probabilities and incidence rates and compared model estimates to the true values used to generate the data. We simulated high and low values for all of the parameters estimated in the model, starting with a "true" typhoid fever incidence of 1000 infections per 100 000 pyo. We began by assuming there were 1250 cases of typhoid fever among 25 000 individuals with a typhoid fever risk factor ($X_i = 1$) and 750 cases among 75 000 individuals without the risk factor ($X_i = 0$) over 2 years of surveillance, such that $R_{TF} = 5$. We then sampled from independent Bernoulli random variables, $H_i \sim \text{Bernoulli}(h_{X_i})$, $B_i \sim \text{Bernoulli}(b_{TF_i})$, and $U_i \sim \text{Bernoulli}(u)$, for the probability of seeking healthcare ($\boldsymbol{H}$), the probability of having blood drawn for culturing ($\boldsymbol{B}$), and the probability of antibiotic usage ($\boldsymbol{U}$). We considered high and low values for each probability, which also varied depending on the risk-factor and typhoid fever status of individual i for h and b, respectively:

$$h_{X_i=1} = \begin{cases} 0.1 \text{ for low scenario,} \\ 0.5 \text{ for high scenario,} \end{cases} \quad h_{X_i=0} = \begin{cases} 0.2 \text{ for low scenario,} \\ 0.7 \text{ for high scenario,} \end{cases}$$

$$b_{TF_i=1} = \begin{cases} 0.68 \text{ for low scenario,} \\ 0.95 \text{ for high scenario,} \end{cases} \quad b_{TF_i=0} = \begin{cases} 0.4 \text{ for low scenario,} \\ 0.9 \text{ for high scenario,} \end{cases}$$

$$u = \begin{cases} 0.2 \text{ for low scenario,} \\ 0.8 \text{ for high scenario.} \end{cases}$$

Among individuals with $TF_i = 1$ and antibiotic usage U_i, the probability of testing positive for typhoid fever (ie, test sensitivity) was $S_i \sim \text{Bernoulli}(s_{U_i})$, where

$$s_{U_i=1} = \begin{cases} 0.4 \text{ for low scenario,} \\ 0.5 \text{ for high scenario,} \end{cases} \quad s_{U_i=0} = \begin{cases} 0.6 \text{ for low scenario,} \\ 0.75 \text{ for high scenario.} \end{cases}$$

Under both scenarios, the mean test sensitivity was ~55%. We assumed perfect test specificity, such that $s_i = 0$ if $TF_i = 0$. The incidence of fever due to other causes was assumed to be 5% per year; we did not allow for multiple episodes of fever over the 2-year period, such that

$$F_{\text{other},i} \sim \text{Bernoulli}(0.10).$$

Moreover, we assumed that the incidence of fever due to any cause was

$$F_i = \max(TF_i, F_{\text{other},i}),$$

that is, all individuals could have at most one episode of fever, either due to typhoid fever or another cause. We then generated vectors for whether or not individuals who sought healthcare ($\boldsymbol{HC}$), had blood drawn for culturing ($\boldsymbol{BC}$), and tested positive for typhoid fever ($\boldsymbol{Y}$) as:

$$HC_i = F_i * H_i,$$

$$BC_i = \begin{cases} HC_i * B_i & \text{if } HC_i = 1, \\ NA & \text{otherwise,} \end{cases}$$

$$Y_i = \begin{cases} BC_i * P_i & \text{if } BC_i = 1, \\ NA & \text{otherwise.} \end{cases}$$

We simulated four scenarios using these probabilities: (1) low probability of seeking care, high probability of being tested, and low prior antibiotic usage; (2) low probability of seeking care, high probability of being tested, and high antibiotic usage; (3) high probability of seeking care, low probability of being tested, and low prior antibiotic usage; and (4) high probability of seeking care, low probability of being tested, and high prior antibiotic usage.

To evaluate whether the final estimates were sensitive to the number of individuals sampled in the HUS, we compared estimates for models that sampled approximately the same number of individuals in each age category as the HUS (735) to models that sampled more individuals (1000 and 2000 individuals).

We additionally compared the adjusted incidence estimates from our model to those from a simpler approach that assumed there was no variation in blood culture sensitivity due to prior antibiotic use (ie, we used a normal distribution for $\phi^S_{a,c}$ instead of a normal mixture model) and no variation in typhoid incidence among those who were or were not tested (ie, using the simpler approach to estimate $\phi^B_{a,c}$ given in Equation 5a) and those who did or did not seek care (ie, assuming $\phi^H_{a,c}$ is equal to the probability of seeking care for fever). This simpler approach is more commonly used in multiplier methods for adjustments; however, this approach often faces criticism since it inherently assumes that data are missing completely at random and that reported healthcare-seeking behavior for a fever is the same as that for typhoid fever.

2.2.5 | Model fitting

To estimate the posterior distributions of the adjusted incidence rates, we collected 100 000 posterior samples from the adjustment factors described above following a burn-in period of 10 000 iterations prior to convergence. Convergence was assessed using the Gelman-Rubin diagnostic[23] for individual parameters. To ensure the model validation was done without knowledge of the true values, one person simulated the data and another person fit the model to the simulated data. Code for generating simulated data was written in MATLAB version 9.3.0.[24] All other analyses were performed using JAGS version 4.3.0 [25] in R version 3.4.0.[26] Model code, including code for generating the simulated data, is provided at https://github.com/mailephillips/adjusted-typhoid-incidence.[27]

3 | RESULTS

The magnitude of the adjustment factors used to estimate the incidence of typhoid fever varied among the three sites (Table 2). In Nepal and Bangladesh, the probability of seeking healthcare was low ($\phi^H_{c=2} = 0.15$, 95% credible interval (CrI): 0.09-0.22; and $\phi^H_{c=3} = 0.27$, 95% CrI: 0.22-0.33, for all ages) and thus contributed the most to the adjustments, while the probability of receiving a blood culture test when eligible was high ($\phi^B_{c=2} = 0.84$, 95% CrI: 0.67-0.92; and $\phi^B_{c=3} = 0.96$, 95% CrI: 0.92-0.98) and contributed the least to adjustments. However, in Malawi, the probability of seeking healthcare was relatively high ($\phi^H_{c=1} = 0.71$, 95% CrI: 0.64-0.77), while the probability of receiving a blood culture test was low ($\phi^B_{c=1} = 0.35$, 95% CrI: 0.34-0.36). Blood culture sensitivity was fairly consistent across sites, with median estimates of $\phi^S_c = 0.54$ to 0.55.

The different adjustment factors also varied by age. Blood culture sensitivity was slightly higher in older age groups compared to younger age groups (median estimates of individuals >14 years, $\phi^S_{a\in(4,5)} = 0.56-0.58$; compared to children ≤14 years, $\phi^S_{a\in(1,2,3)} = 0.53-0.54$). While there was no consistent pattern in prior antibiotic use by age, the amount of blood drawn for a blood culture test generally increased with age. As a result, blood culture sensitivity slightly increased with age. Had we not adjusted for blood culture volume or prior antibiotic use, the estimate for blood culture sensitivity would have been higher for all ages and countries (Figure S5, in blue).

The probability of receiving a blood culture test had different patterns across age groups depending on the country. In Malawi, the probability of receiving a blood culture test decreased with age ($\phi^B_{1,1} = 0.40$, 95% CrI: 0.38−0.41 for children <5 years vs $\phi^B_{5,1} = 0.20$, 95% CrI: 0.17−0.23 for adults 30+ years), while in Nepal and Bangladesh, the probability increased ($\phi^B_{1,2} = 0.81$, 95% CrI: 0.61−0.91 and $\phi^B_{1,3} = 0.94$, 95% CrI: 0.87−0.97 for children <5 years vs $\phi^B_{5,2} = 0.91$, 95% CrI: 0.71−0.98 and $\phi^B_{5,3} = 0.98$, 95% CrI: 0.96−0.99 for adults 30+ years). If we had used a simpler approach (not adjusting for the variation in the risk of typhoid fever among those who were and were not tested) to adjust for the probability of receiving a blood culture test in Nepal and Bangladesh, we would have underestimated the probabilities and thus overestimated the contribution of this adjustment (Figure S5). In Nepal, in particular, the simpler approach would have substantially underestimated the probability of receiving a blood culture test among younger age groups. In Bangladesh, the unadjusted proportion of individuals receiving a blood culture test was already close to one, so the adjusted value did not increase the estimate considerably.

The probability of seeking healthcare did not have a consistent pattern across age groups, likely due to the different components contributing to the final estimate. Malawi overall had the highest rates of healthcare seeking, with the lowest rates among children under 5 and the highest rates among children 5 to 14 (Table 2). Nepal had the lowest rates of

TABLE 2 Posterior probability estimates for each adjustment factor by age and site

		Age category (a), in years					
Probability	**Country (c)**	**<5**	**5 to 9**	**10 to 14**	**15 to 29**	**30+**	**All**
Sensitivity of BC, given HC sought, BC taken ($\phi^S_{a,c}$)	Malawi	0.53 (0.48-0.57)	0.53 (0.51-0.55)	0.53 (0.50-0.56)	0.58 (0.55-0.62)	0.58 (0.54-0.62)	0.54 (0.35-0.60)
	Nepal	0.54 (0.51-0.57)	0.54 (0.51-0.57)	0.54 (0.51-0.58)	0.56 (0.54-0.58)	0.56 (0.55-0.57)	0.55 (0.51-0.58)
	Bangladesh	0.53 (0.52-0.53)	0.53 (0.53-0.53)	0.53 (0.53-0.53)	0.56 (0.54-0.58)	0.56 (0.56-0.56)	0.54 (0.51-0.57)
Probability BC was drawn, given HC sought ($\phi^B_{a,c}$)	Malawi	0.40 (0.38-0.41)	0.38 (0.36-0.41)	0.33 (0.29-0.36)	0.21 (0.19-0.24)	0.20 (0.17-0.23)	0.35 (0.34-0.36)
	Nepal	0.81 (0.61-0.91)	0.87 (0.72-0.94)	0.87 (0.73-0.94)	0.91 (0.71-0.98)	0.91 (0.71-0.98)	0.84 (0.67-0.92)
	Bangladesh	0.94 (0.87-0.97)	0.97 (0.94-0.99)	0.97 (0.94-0.99)	0.98 (0.96-0.99)	0.98 (0.96-0.99)	0.96 (0.92-0.98)
Probability of seeking HC ($\phi^H_{a,c}$)	Malawi	0.62 (0.52-0.72)	0.83 (0.74-0.90)	0.83 (0.74-0.90)	0.71 (0.60-0.81)	0.71 (0.60-0.81)	0.71 (0.64-0.77)
	Nepal	0.21 (0.11-0.34)	0.11 (0.04-0.22)	0.11 (0.04-0.22)	0.13 (0.05-0.25)	0.13 (0.05-0.25)	0.15 (0.09-0.22)
	Bangladesh	0.32 (0.21-0.45)	0.34 (0.24-0.45)	0.33 (0.24-0.45)	0.19 (0.11-0.27)	0.19 (0.11-0.27)	0.27 (0.22-0.33)

Note: Each estimate (posterior mean) is shown with its 95% credible interval for the sensitivity of the blood culture (BC) given that healthcare (HC) was sought and a blood culture was taken ($\phi^S_{a,c}$), the probability of receiving a blood culture test given that healthcare was sought ($\phi^B_{a,c}$), and the probability of seeking healthcare ($\phi^H_{a,c}$). Estimated adjustment factors are shown by age category (A) and country (C).

TABLE 3 Estimated adjustment factors from final models

Age (years)	Malawi	Nepal	Bangladesh
0 to 4	7.6 (4.8-11.6)	10.7 (4.3-26.8)	6.3 (4.2-10.2)
5 to 9	5.9 (4.1-8.3)	19.7 (9.0-54.9)	5.8 (4.1-8.6)
10 to 14	6.9 (4.3-10.4)	19.6 (8.6-55.2)	5.8 (3.9-8.9)
15 to 29	11.4 (6.9-18.0)	15.8 (7.4-42.4)	9.8 (6.2-16.7)
30+	12.0 (6.0-21.7)	15.0 (4.6-48.8)	9.7 (5.5-17.9)
All ages	7.7 (6.0-12.4)	14.4 (9.3-24.9)	7.0 (5.6-9.2)

Note: The ratio of the median estimate (95% credible interval) of adjusted-to-observed incidence rates is shown for each country and age category.

healthcare seeking overall, with slightly higher rates among children under 5. Bangladesh also had low healthcare seeking rates, with the lowest rates among adults 15+ years of age (Table 2). In Malawi and Bangladesh, the proportion of those with the relevant typhoid risk factor did not differ by age group (Table S2). Healthcare seeking for a fever was higher among those with the typhoid risk factor in Malawi, but lower among those with the risk factor in Bangladesh and Nepal. When compared to the simpler approach to estimate the probability of healthcare seeking (using the unadjusted proportion of those who sought care for fever), the estimates were similar but slightly higher in most age groups in Malawi but slightly lower in Nepal and Bangladesh (Figure S5).

The magnitude of the overall adjustment to estimate typhoid fever incidence ($\lambda_{\text{adjusted},a,c}/\lambda_{\text{observed},a,c} = (\phi^S_{a,c}\phi^B_{a,c}\phi^H_{a,c})^{-1}$) varied between countries and age groups. Nepal had the highest adjustment factors in every age group, with an overall adjustment factor of 14.4 (95% CrI: 9.3–24.9). Malawi and Bangladesh were similar, with adjustment factors of 7.7 (95% CrI: 6.0–12.4) and 7.0 (95% CrI: 5.6–9.2), respectively (Table 3). The highest adjustment factor was for the 5 to 9-year age group in Nepal (19.7, 95% CrI: 9.0-54.9), while the lowest was for the 5 to 9- and 10 to 14-year age groups in Bangladesh (5.8, 95% CrI: 4.1–8.6; and 5.8, 95% CrI: 3.9–8.9, respectively).

Most of the final adjusted incidence estimates fell below the upper bound of the estimated seroincidence of typhoid infection (Table 4). In Malawi and Bangladesh, all of the upper bounds of the estimates were well within or below the upper bounds of the seroincidence rate confidence intervals. In Nepal among children aged 5 to 9 years, the upper bound of the adjusted incidence rate exceeded that of the estimated seroincidence; however, the confidence/credible intervals overlapped. Similarly, children 10 to 14 years of age in Bangladesh had a higher median adjusted incidence rate than seroincidence (which was the lowest across all age groups and sites), but the upper bound of the adjusted incidence credible interval was still lower than the upper bound of the confidence interval for seroincidence.

When we evaluated the model against simulated data, the full model was able to reliably estimate both the "true" incidence of typhoid fever and the probabilities used to generate the simulated data for a range of values, while the simpler approach over- or under-estimated the true incidence in some scenarios (Figures 2 and S3). Estimates of blood culture sensitivity were similar to the true value, but incorporated additional uncertainty compared to the simpler approach, consistent with the different sensitivity of blood culture in those who reported prior antibiotic use compared to those that did not (Figure S3). The probability of receiving a blood culture test contributed most to the difference in accuracy between the two approaches. In every scenario, the full model reliably estimated the true value, while the simpler approach underestimated the true value (Figure S3). The adjustment for healthcare seeking in the full model again consistently captured the true value across levels of the probability as compared to the simpler approach, which generally had a narrower 95% CrI that did not always contain the true value (Figure S3). As expected, in both models, the uncertainty in the probability of seeking healthcare decreased as the sampling fraction for the hypothetical HUS increased. As a result, the 95% CrIs in the overall incidence estimates also narrowed as the sampling fraction increased (Figure 2).

4 | DISCUSSION

In order to make informed decisions regarding typhoid control and prevention, it is important to have accurate estimates of population-based typhoid fever incidence. Unreliable reports, inconsistent healthcare utilization, inconsistent clinical diagnoses, suboptimal diagnostic tests, and scarcity of accurate or full data contribute to difficulties in calculating the

TABLE 4 Adjusted typhoid incidence estimates compared to seroincidence

	Malawi			Nepal			Bangladesh		
Age	**Crude rates**	**Adjusted rates**	**Seroincidence**	**Crude rates**	**Adjusted rates**	**Seroincidence**	**Crude rates**	**Adjusted rates**	**Seroincidence**
0 to 4 years	83 (53-124)	632 (398-965)	2868 (1153-5911)	72 (33-136)	764 (307-1921)	7813 (2537-18 232)	417 (337-511)	2625 (1764-4244)	3401 (1904-5610)
5 to 9 years	146 (103-201)	861 (599-1203)	1205 (146-4352)	341 (250-455)	6713 (3085-18 730)	5217 (1915-11 356)	554 (456-666)	3228 (2276-4757)	3435 (1571-6521)
10 to 14 years	88 (56-132)	602 (377-915)	3061 (631-8946)	191 (128-275)	3750 (1653-10 559)	8910 (4075-16 916)	268 (203-348)	1564 (1050-2384)	599 (15-3336)
15 to 29 years	32 (20-48)	361 (219-567)	3774 (1384-8213)	92 (71-119)	1457 (684-3918)	10 169 (5255-17 764)	98 (76-124)	956 (603-1635)	5310 (2744-9275)
30+ years	21 (10-37)	248 (124-447)	2076 (762-4518)	6 (2-13)	92 (29-301)	7322 (5100-10 183)	29 (19-42)	279 (157-514)	2988 (1672-4928)
All ages	58 (48-70)	444 (347-717)	2505 (1605-3728)	74 (62-87)	1062 (683-1839)	7631 (5914-9691)	161 (145-179)	1135 (898-1480)	3256 (2432-4270)

Note: The final adjusted typhoid incidence estimates from the models are shown with 95% credible intervals, as well as the seroincidence estimates with their 95% confidence intervals, by age and site.

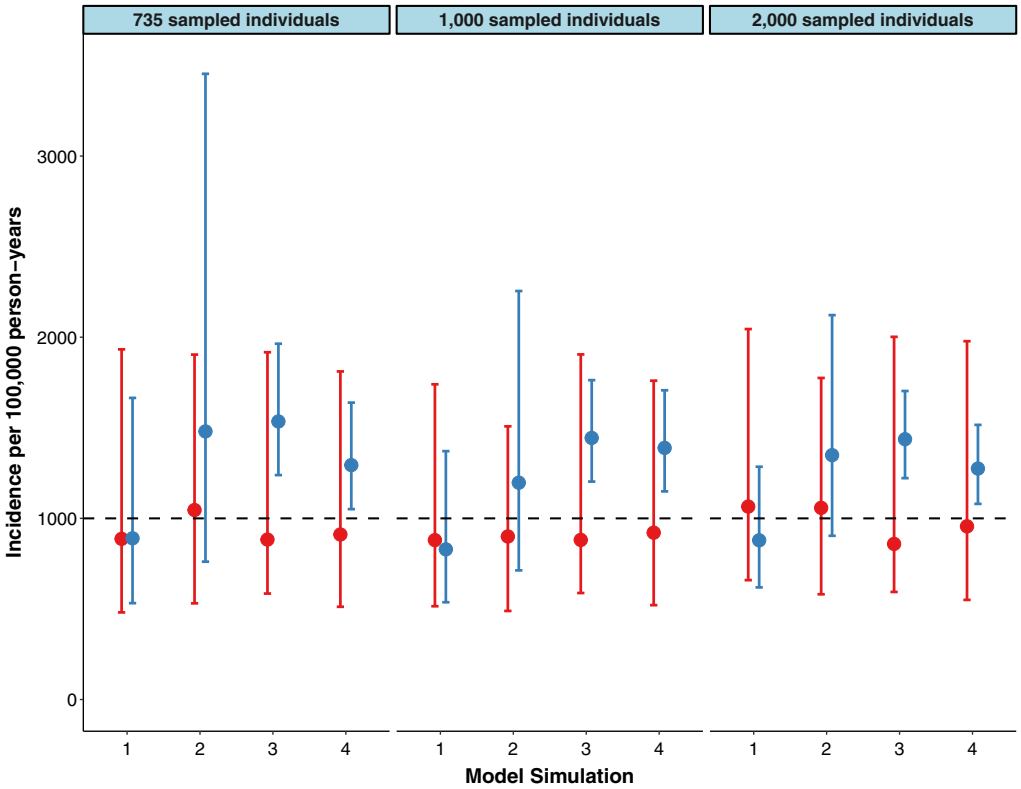

FIGURE 2 Estimated typhoid incidence based on simulated data: Full model vs simplified approach. The typhoid incidence per 100 000 person-years of observation was estimated from simulated data based on a true incidence of 1000 typhoid infections per 100 000 person-years (dashed horizontal black line). Data were simulated for low and high probabilities of seeking healthcare, receiving a blood culture diagnostic test, and antibiotic use. Scenarios were as follows: (1) low probability of seeking care, high probability of being tested, and low prior antibiotic usage; (2) low probability of seeking care, high probability of being tested, and high prior antibiotic usage; (3) high probability of seeking care, high probability of being tested, and low prior antibiotic usage; and (4) high probability of seeking care, high probability of being tested, and high prior antibiotic usage. Each simulation was performed sampling 735; 1000; and 2000 individuals from the population for the hypothetical healthcare utilization survey. Estimated "true" values are shown for models that did (red) and did not (blue) account for variation in blood culture sensitivity and variation in typhoid incidence among those who did or did not seek care and were or were not tested [Colour figure can be viewed at wileyonlinelibrary.com]

population-based incidence of typhoid fever. We developed a new methodology within the Bayesian setting to estimate population-based incidence in a context where cases often go undetected and under-reported. Through this approach, we were able to calculate the adjustment factors that can be applied to estimate the "true" incidence of typhoid fever in the STRATAA surveillance sites. These estimates suggested that the adjusted incidence of typhoid fever in Malawi, Nepal, and Bangladesh is 7- to 14-fold higher than the reported blood-culture-confirmed numbers.

It is commonly accepted that cases of typhoid fever go unrecognized at each phase of the reporting process, but the degree to which each step contributes to the underestimation of and uncertainty in the population-based incidence is often not fully quantified. Each of the three intervening processes contributed differently to the underestimation in each country and age group. The probability of seeking healthcare was responsible for the largest portion of underestimation in Nepal and Bangladesh, while the probability of receiving a blood culture test was the biggest factor in Malawi. These results reflect differences in the healthcare systems and fever surveillance processes at the different sites. In Nepal and Bangladesh, antibiotics are widely available and individuals tend to seek care first at a pharmacy instead of a healthcare facility.[28] In Nepal, a considerable proportion of people with fever neither seek healthcare nor visit a pharmacy possibly due to lack of funds. In Malawi, healthcare seeking is high, but the resources at healthcare facilities are limited. As a result, many people report to healthcare facilities with a fever, receive antibiotics, but do not remain in the facility long enough to receive an additional blood culture test due to prolonged wait times.

Other studies use different approaches to estimate the true incidence of typhoid fever. In many cases, studies simply double the reported cases to account for blood culture sensitivity.[29] Numerous studies make use of a simple multiplier

method,[30-33] which often do not accurately reflect the uncertainty associated with the reporting process. Previous studies have not attempted to integrate data sources to account for potential differences between the observed healthcare seeking and testing probabilities for fever vs the corresponding unobserved probabilities specific to typhoid fever, which our analysis suggests can impact the final adjustment factors. Our approach can be used to estimate typhoid fever incidence in other study populations. Moreover, some of the issues we encountered (eg, under-detection due to poor test sensitivity that varies depending patient characteristics, preferential testing of individuals more likely to have the disease of interest, and potential differences in healthcare seeking for those who have the syndrome vs disease of interest) are common to other diseases as well.

By utilizing a Bayesian approach, we were able to measure the contribution to underestimation at each phase of the reporting process while also properly quantifying the uncertainty for each of our estimates. When comparing our model to simulated data, we showed that having more data available (due to higher probabilities of seeking care and receiving a blood culture test) reduced uncertainty in the estimates. Furthermore, we showed that if more people had been sampled in the HUS, uncertainty would also have been reduced.

Our analysis and approach had some limitations. We assumed that healthcare-seeking behavior for fever in households without children is the same as households with children, because the HUS only sampled households with children. Less than a third of households did not have children and were not included in this survey across sites. Studies suggest that households with children are more likely to seek healthcare,[34] which means that if our estimates are biased, the adjusted incidence estimates are likely conservative. We also were not able to differentiate between febrile illnesses at different parts of the reporting process. We addressed this issue by adjusting for possible differences in healthcare seeking among those with typhoid fever compared to other febrile etiologies using weighted averages based on known risk factors for typhoid fever, but other factors may also lead to differences in healthcare seeking for fever vs for typhoid fever. By comparing our adjusted incidence estimates to estimates of seroincidence, we are able to provide some assurance that the adjusted incidence is within the range of plausible values. However, methods and immunological markers for estimating the seroincidence of typhoid fever are not well established, and the cut-off we used (a $\geq$2-fold and absolute value of $\geq$50 EU/mL in anti-Vi IgG) may not be a reliable indicator of acute typhoid infection across all individuals and immunological backgrounds. Another limitation of the approach is that it can very labor-intensive and time-consuming to collect the necessary data.

Calculation of incidence based on data from passive surveillance of blood-culture-confirmed typhoid fever results in considerable underestimation of the true incidence of typhoid fever in the population. Our model provides an approach for estimating typhoid fever incidence while accounting for different sources of information from the reporting process. Typhoid fever remains a major cause of morbidity and mortality in developing countries, so control and prevention are needed. To effectively prioritize, implement and evaluate interventions, estimates of the number of cases should accurately reflect the uncertainty in the reporting process. This analysis provides a platform that can be updated with new or additional data as they become available and can be adapted to other contexts. This model framework could also be used to adjust for underreporting in other diseases.

ACKNOWLEDGEMENTS

The STRATAA study is supported by a Wellcome Trust Strategic Award (106 158/Z/14/Z) and the Bill and Melinda Gates Foundation (OPP1141321). The Malawi-Liverpool-Wellcome Unit is core-funded by a grant from the Wellcome Trust (206545/Z/17/Z). RSH is supported by the National Institute for Health Research (NIHR) Global Health Research Unit on Mucosal Pathogens using UK aid from the UK Government. MAG is supported by a UK National Institute for Health Research (NIHR) Research Professorship in Global Health (NIHR300039). The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

CONFLICT OF INTEREST

Virginia E. Pitzer is a member of the World Health Organization's (WHO) Immunization and Vaccine-related Implementation Research Advisory Committee. Andrew J. Pollard chairs the UK Department of Health's (DoH) Joint Committee on Vaccination and Immunisation (JCVI) and the European Medicines Agency Scientific Advisory Group on Vaccines and is a member of the World Health Organization's (WHO) Strategic Advisory Group of Experts. The views expressed in this manuscript are those of the authors and do not necessarily reflect the views of the JCVI, the DoH, or the WHO.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Maile T. Phillips https://orcid.org/0000-0001-6952-0440

REFERENCES

1. Crump JA, Youssef FG, Luby SP, et al. Estimating the incidence of typhoid fever and other febrile illnesses in developing countries. *Emerg Infect Dis*. 2003;9(5):539-544. https://doi.org/10.3201/eid0905.020428
2. Antillon M, Warren J, Crawford F, Weinberger D, Kurum E, Pitzer V. The burden of typhoid fever in low- and middle-income countries: a meta-regression approach. *PLoS Negl Trop Dis*. 2017;11(2):e0005376. https://doi.org/10.1371/journal.pntd.0005376
3. Mogasale V, Maskery B, Ochiai RL, et al. Burden of typhoid fever in low-income and middle-income countries: a systematic, literature-based update with risk-factor adjustment—the lancet global health. *Lancet*. 2014;2(10):e570-e580. https://doi.org/10.1016/S2214-109X(14)70301-8
4. Global Burden of Disease Collaborative Network (IHME) IfHMaE. *Global Burden of Disease Study 2016 (GBD 2016) Results*, Seattle, WA: Global Burden of Disease Collaborative Network, 2017. http://ghdx.healthdata.org/gbd-results-tool. Accessed July 23, 2018.
5. Parry CM, Hien TT, Dougan G, White NJ, Farrar JJ. Typhoid fever. *New Engl J Med*. 2002;347:1770-1782. https://doi.org/10.1056/NEJMra020201
6. Panzner U, Pak GD, Aaby P, et al. Utilization of healthcare in the typhoid fever surveillance in Africa program. *Clin Infect Dis*. 2016;62(Suppl 1):S56-S68. https://doi.org/10.1093/cid/civ891
7. Kassile T, Lokina R, Mujinja P, Mmbando BP. Determinants of delay in care seeking among children under five with fever in Dodoma region, central Tanzania: a cross-sectional study. *Malaria J*. 2014;13(1):348. https://doi.org/10.1186/1475-2875-13-348
8. Dillip A, Hetzel MW, Gosoniu D, et al. Socio-cultural factors explaining timely and appropriate use of health facilities for degedege in South-Eastern Tanzania. *Malaria J*. 2009;8:144. https://doi.org/10.1186/1475-2875-8-144
9. Elven J, Dahal P, Ashley EA, et al. Non-malarial febrile illness: a systematic review of published aetiological studies and case reports from Africa, 1980–2015. *BMC Med*. 2020;18:279. https://doi.org/10.1186/s12916-020-01744-1
10. Voysey M, Pant D, Shakya M, et al. Under-detection of blood culture-positive enteric fever cases: the impact of missing data and methods for adjusting incidence estimates. *PLoS Negl Trop Dis*. 2020;14(1):e0007805. https://doi.org/10.1371/journal.pntd.0007805
11. World Health Organization (WHO). Typhoid vaccines position paper. *Wkly Epidemiol Rec*. 2018;13(93):153-172.
12. Antillon M, Saad NJ, Baker S, Pollard AJ, Pitzer VE. The relationship between blood sample volume and diagnostic sensitivity of blood culture for typhoid and paratyphoid fever: a systematic review and meta-analysis. *J Infect Dis*. 2018;218(suppl_4):S255-S267. https://doi.org/10.1093/infdis/jiy471
13. Reed C, Angulo FJ, Swerdlow DL, et al. Estimates of the prevalence of pandemic (H1N1) 2009, United States, April–July 2009. *Emerg Infect Dis*. 2009;15(12):2004-2007.
14. Presanis AM, De Angelis D, Hagy A, et al. The severity of pandemic H1N1 influenza in the United States, from April to July 2009: a Bayesian analysis. *PLoS Med*. 2009;6(12):e1000207. https://doi.org/10.1371/journal.pmed.1000207
15. Health Protection Agency. HIV in the United Kingdom: 2008 Report; 2008. http://data.unaids.org/pub/report/2008/20081125_unitedkingdom_hivestimate2008_en.pdf
16. Goubar A, Ades AE, De Angelis D, et al. Estimates of human immunodeficiency virus prevalence and proportion diagnosed based on Bayesian multiparameter synthesis of surveillance data [with discussion]. *J R Stat Soc Ser A (Stat Soc)*. 2008;171(3):541-580.
17. Presanis AM, De Angelis D, Spiegelhalter DJ, Seaman S, Goubar A, Ades AE. Conflicting evidence in a Bayesian synthesis of surveillance data to estimate human immunodeficiency virus prevalence. *J R Stat Soc Ser A (Stat Soc)*. 2008;171(4):915-937.
18. Darton TC, Meiring JE, Tonks S, et al. The STRATAA study protocol: a programme to assess the burden of enteric fever in Bangladesh, Malawi and Nepal using prospective population census, passive surveillance, serological studies and healthcare utilisation surveys. *BMJ Open*. 2017;7(6):e016283. https://doi.org/10.1136/bmjopen-2017-016283
19. Meiring J, Gibani M, The TyVAC Consortium Meeting Group. The typhoid vaccine acceleration consortium (TyVAC): vaccine effectiveness study designs: accelerating the introduction of typhoid conjugate vaccines and reducing the global burden of enteric fever. report from a meeting held on 26–27 October 2016, Oxford, UK. *Vaccine*. 2017;35(38), 5081-5088doi:https://doi.org/10.1016/j.vaccine.2017.08.001
20. Gauld JS, Olgemoeller F, Nkhata R, et al. Domestic river water use and risk of typhoid fever: results from a case-control study in Blantyre, Malawi. *Clin Infect Dis*. 2020;70(7):1278-1284. https://doi.org/10.1093/cid/ciz405
21. Karkey A, Thompson CN, Tran Vu Thieu N, et al. Differential epidemiology of *Salmonella* Typhi and Paratyphi a in Kathmandu, Nepal: a matched case control investigation in a highly endemic enteric fever setting. *PLoS Negl Trop Dis*. 2013;7(8):e2391. https://doi.org/10.1371/journal.pntd.0002391
22. Ram PK, Naheed A, Brooks WA, et al. Risk factors for typhoid fever in a slum in Dhaka, Bangladesh. *Epidemiol Infect*. 2007;135(3):458-465. https://doi.org/10.1017/s0950268806007114
23. Gelman A, Rubin D. Inference from iterative simulation using multiple sequences. *Stat Sci*. 1992;7:457-472.
24. *MATLAB*. Version 9.3.0 (R2017b). The MathWorks Inc; 2017.
25. *rjags: Bayesian Graphical Models using MCMC*. Version R package version 4-6; 2016. https://CRAN.R-project.org/package=rjags

26. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2017. https://www.R-project.org/
27. Phillips MT. Adjusted Typhoid Incidence; 2020. Accessed January 13, 2020. https://github.com/mailephillips/adjusted-typhoid-incidence
28. Ocan M, Obuku EA, Bwanga F, et al. Household antimicrobial self-medication: a systematic review and meta-analysis of the burden, risk factors and outcomes in developing countries. *BMC Publ Health*. 2015;15:742. https://doi.org/10.1186/s12889-015-2109-3
29. Srikantiah P, Girgis F, Luby SP, et al. Population-based surveillance of typhoid fever in Egypt. *Am J Trop Med Hyg*. 2006;74(1):114-119.
30. Andrews JR, Barkume C, Yu AT, et al. Integrating facility-based surveillance with healthcare utilization surveys to estimate enteric fever incidence: methods and challenges. *J Infect Dis*. 2018;218(suppl_4):S268-S276. https://doi.org/10.1093/infdis/jiy494
31. Carey ME, MacWright WR, Im J, et al. The Surveillance for Enteric Fever in Asia Project (SEAP), Severe Typhoid Fever Surveillance in Africa (SETA), Surveillance of Enteric Fever in India (SEFI), and Strategic Typhoid Alliance across Africa and Asia (STRATAA) population-based enteric fever studies: a review of methodological similarities and differences. *Clin Infect Dis*. 2020;71(S2):S102-S110. https://doi.org/10.1093/cid/ciaa367
32. Marks F, von Kalckreuth V, Aaby P, et al. Incidence of invasive salmonella disease in sub-Saharan Africa: a multicentre population-based surveillance study. *Lancet Glob Health*. 2019;5(3):e310-e323. https://doi.org/10.1016/S2214-109X(17)30022-0
33. Breiman RF, Cosmas L, Njuguna H, et al. Population-based incidence of typhoid fever in an urban informal settlement and a rural area in Kenya: implications for typhoid vaccine use in Africa. *PLoS One*. 2012;7(1):e29119. https://doi.org/10.1371/journal.pone.0029119
34. Soni A, Fahey N, Phatak AG, et al. Differential in healthcare-seeking behavior of mothers for themselves versus their children in rural India: results of a cross sectional survey. *Int Publ Health J*. 2014;6(1):57-66.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Phillips MT, Meiring JE, Voysey M, et al. A Bayesian approach for estimating typhoid fever incidence from large-scale facility-based passive surveillance data. *Statistics in Medicine*. 2021;1–18. https://doi.org/10.1002/sim.9159