

# Misclassification of yellow fever vaccination status revealed through hierarchical Bayesian modeling

Quan Minh Tran and T. Alex Perkins

Department of Biological Sciences and Eck Institute for Global Health  
University of Notre Dame

Correspondence: [taperkins@nd.edu](mailto:taperkins@nd.edu)

## Abstract

Vaccination coverage estimates are crucial inputs to decisions about investments in vaccination, yet they can be prone to inaccuracies. At the individual level, inaccuracies can be described in terms of the sensitivity and specificity of vaccination status. We estimated these quantities using a hierarchical Bayesian analysis of data from a test-negative study design with reported yellow fever vaccination status as the exposure. Our analysis accounted for the possibility of misclassification of both the exposure and the test at the country level. Across all countries, our median estimates of the sensitivity and specificity of vaccination status were 0.69 (95% credible interval [CrI]: 0.21-0.98) and 0.70 (95% CrI: 0.21-0.98), respectively. Median estimates at the country level ranged from 0.06 (95% CrI: 0.04-0.09) to 0.96 (95% CrI: 0.94-0.98) for sensitivity, and from 0.15 (95% CrI: 0.09-0.23) to 0.98 (95% CrI: 0.90-1.00) for specificity. This suggests that there is substantial misclassification of yellow fever vaccination status in general and extensive variation in misclassification across countries. Taking into account misclassification in vaccination status, we made adjustments to reported vaccination coverage and showed that reported coverage may be significantly underestimated in 10 out of 20 countries and significantly overestimated in 5 out of 20.

## Introduction

Vaccination coverage—defined as the proportion of a population that is vaccinated—is a critically important piece of information for public health. This information can be used to monitor vaccination efforts, identify areas where coverage is lower than desired, and determine potential barriers to vaccination. For example, knowing that certain age groups or regions have lower coverage can be used as a basis for prioritizing those populations for vaccination [1]. Additionally, vaccination coverage is useful for assessing the effectiveness of vaccination programs at reducing the burden of vaccine-preventable diseases [2]. It is also used as a standard to assess the strength of vaccination programs, to evaluate the accessibility of healthcare systems [3, 4], and as a criterion for financial support for vaccination programs in underresourced settings [5].

Despite its importance, there is considerable potential for error in vaccination coverage estimates. In some countries, these estimates are derived from computer-based systems [4, 6]. Though considered to be more advanced, errors in these systems can arise from mistakes in recall, recording, compiling records, or not taking into account human migration. In most low- and middle-income countries (LMICs), vaccination coverage is usually reported as an “administrative coverage estimate” [4], which is defined as the number of vaccine doses that were administered divided by the size of the target population. Errors may arise from this approach as a result of duplicated vaccination records, doses being administered but not recorded, or inaccurate estimates of the target population size [7]. Household surveys can be conducted to address these issues but may suffer from selection bias (i.e., unrepresentative households) and information bias (i.e., inaccurate data transcription or recall) [4, 8, 9].

One disease for which vaccination coverage estimates are particularly important is yellow fever. Although most contemporary cases are thought to result from zoonotic spillover [10], yellow fever is

capable of large outbreaks with high mortality [11, 12]. As a result, the World Health Organization (WHO) has dedicated an initiative to the elimination of yellow fever epidemics [13]. Because there is a highly efficacious vaccine for this disease and few other effective tools for prevention [14], achieving high vaccination coverage in at-risk populations is essential. Understanding of vaccination coverage for yellow fever may be impaired by the fact that it poses the greatest risk to rural populations in Africa and South America, where accurate estimation of vaccination coverage is challenging [7, 15]. This emphasizes the need for accurate estimates of yellow fever vaccination coverage.

In this study, we estimated the sensitivity and specificity of vaccination status using a test-negative, case-control study design involving surveillance data on yellow fever in Africa. Similar approaches have been taken previously to estimate vaccine effectiveness in light of diagnostic inaccuracy [16, 17, 18, 19], exposure inaccuracy [20], or both [21]. Our analysis allowed for the possibility of inaccuracy in both the exposure (vaccination status) and the test (infection status), not to adjust estimates of vaccine effectiveness but to obtain estimates of vaccination misclassification in a real-world setting. Additionally, we assessed the extent of geographic variability in misclassification by comparing alternative models with differing assumptions about geographic variability in the sensitivity and specificity of vaccination status, the sensitivity and specificity of diagnostic testing, the proportion of infections that were reported, and the force of infection. Finally, we showed how estimates of sensitivity and specificity of vaccination status can be used to adjust vaccination coverage estimates to account for vaccination status misclassification.

## Methods

### Data

We used a database of yellow fever surveillance records from Africa collated by the World Health Organization. This database consists of suspected cases of yellow fever (based primarily on a case definition of fever and jaundice [22]), of which there were approximately 15,000 in the database from 20 African countries between 2005 and 2011. Suspected cases who tested positive for yellow fever virus immunoglobulin M were classified as confirmed cases of yellow fever. Each individual’s age and reported history of vaccination against yellow fever were also recorded. We used spatial information about each case at the first administrative level (adm1). Use of this database for this analysis was deemed not human subjects research by the University of Notre Dame Institutional Review Board (Protocol 22-04-7193). For our main analysis, we used yellow fever vaccination coverage estimates by age and year [23]. These estimates span 1950 to 2050 and provide information on vaccination coverage at the adm1 level.

### Model

Each suspected case of yellow fever in the database was classified according to two observed states:  $C$  (1 = test-positive, 0 = test-negative) and  $R$  (1 = reported vaccinated, 0 = reported non-vaccinated). Our model also considered two latent states for each suspected case:  $I$  (1 = truly infected, 0 = truly not infected) and  $T$  (1 = truly vaccinated, 0 = truly not vaccinated). These states are related in our model according to causal dependencies that describe how the data were generated (Fig. 1).

First,  $T \rightarrow R$  is determined by the sensitivity and specificity of reported vaccination status in each country  $k$ , which we represent with  $Se_{vac,k}$  and  $Sp_{vac,k}$ , respectively. Specifically,

$$\Pr(R = 1|T = 1) = Se_{vac,k} \quad (1)$$

$$\Pr(R = 0|T = 1) = 1 - Se_{vac,k} \quad (2)$$

$$\Pr(R = 0|T = 0) = Sp_{vac,k} \quad (3)$$

$$\Pr(R = 1|T = 0) = 1 - Sp_{vac,k} \quad (4)$$

describe the relationships among all combinations of  $T$  and  $R$ . We enforced a constraint on  $Se_{vac,k}$  and  $Sp_{vac,k}$  to ensure that  $Se_{vac,k} + Sp_{vac,k} > 1$ , because the alternative would imply that vaccination status is, on average, misclassified more often than not. To implement this constraint, we modeled  $1 - Se_{vac,k}$ ,  $1 - Sp_{vac,k}$ , and  $Se_{vac,k} + Sp_{vac,k} - 1$  as random effects across countries  $k$  following a Dirichlet distribution. In other words, we assumed that the country-specific values of these parameters follow a global Dirichlet distribution with concentration parameters  $\alpha_{vac,1}$ ,  $\alpha_{vac,2}$ , and  $\alpha_{vac,3}$ .

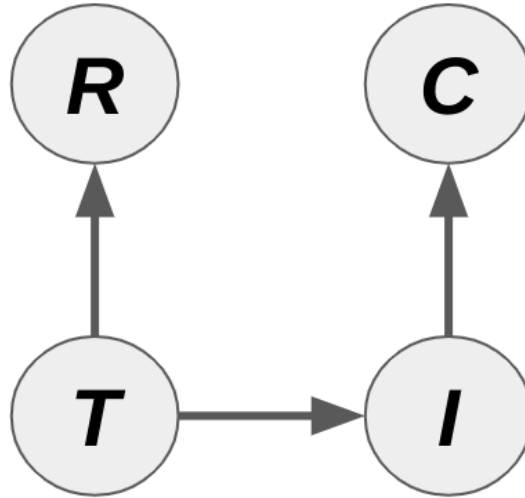


Figure 1: **Vaccination status misclassification model.** The diagram shows the causal relationships among four states for each individual in the database:  $R$ , reported vaccination status;  $T$ , true vaccination status;  $I$ , true infection status; and  $C$ , reported disease status. The core of the model’s likelihood,  $\mathcal{L}_{data}$ , depends on probabilities of each of these three causal relationships.  $\Pr(R|T)$  depends on parameters describing the sensitivity ( $Se_{vac}$ ) and specificity ( $Sp_{vac}$ ) of vaccination classification,  $\Pr(I|T)$  depends on the force of infection ( $\lambda$ ) and reporting probability ( $\rho$ ), and  $\Pr(C|I)$  depends on parameters describing the sensitivity ( $Se_{test}$ ) and specificity ( $Sp_{test}$ ) of case confirmation.

Second,  $T \rightarrow I$  is determined by the probability of contracting yellow fever given one’s true vaccination status, which we represent with  $\Pr(I|T)$ . We start with the case where  $T = 0$  and  $I = 1$ , the probability of which is

$$\Pr(I = 1|T = 0) = e^{-(a-1)\lambda_k} (1 - e^{-\lambda_k}) \rho_k, \quad (5)$$

where  $a$  is age at the time that an individual is recorded as a suspected case,  $\lambda_k$  is the force of infection of yellow fever virus in the individual’s country  $k$ , and  $\rho_k$  is the probability that an infection is reported in country  $k$ . The term  $e^{-(a-1)\lambda_k}$  represents the probability that an individual avoids infection before age  $a$ , whereas  $1 - e^{-\lambda_k}$  represents the probability that the individual becomes infected at age  $a$ . This formulation is widely known within epidemiology as a “catalytic” model [24] and has been used in prior models of yellow fever disease burden [22, 25, 26]. Eqn. (5) implies that

$$\Pr(I = 0|T = 0) = 1 - \Pr(I = 1|T = 0), \quad (6)$$

which encompasses all the ways in which an individual would not become a suspected case with an etiology of yellow fever virus at age  $a$ . Individuals who are truly vaccinated become suspected cases with an etiology of yellow fever virus with probability

$$\Pr(I = 1|T = 1) = (1 - VE) \Pr(I = 1|T = 0), \quad (7)$$

where  $VE$  is vaccine efficacy. Similar to eqn. (6),

$$\Pr(I = 0|T = 1) = 1 - \Pr(I = 1|T = 1) \quad (8)$$

is the probability that a truly vaccinated person does not become a suspected case with an etiology of yellow fever virus at age  $a$ .

Third,  $I \rightarrow C$  is determined by the sensitivity and specificity of laboratory testing in each country  $k$ , which we represent with  $Se_{test,k}$  and  $Sp_{test,k}$ , respectively. Like eqns. (1)-(4),

$$\Pr(C = 1|I = 1) = Se_{test,k} \quad (9)$$

$$\Pr(C = 0|I = 1) = 1 - Se_{test,k} \quad (10)$$

$$\Pr(C = 0|I = 0) = Sp_{test,k} \quad (11)$$

$$\Pr(C = 1|I = 0) = 1 - Sp_{test,k} \quad (12)$$

describe the probabilities that confirmed cases result from suspected cases with differing true infection statuses in country  $k$ . We note that individuals with a status of  $I = 1$  represent reported infections only, given that the reporting probability  $\rho_k$  filters out infections that went unreported in country  $k$ . We treat  $Se_{test,k}$  and  $Sp_{test,k}$  as fixed effects.

Based on these three causal dependencies, we can define the joint distribution of all four states  $(c, i, r, t)$  for each individual to be

$$\Pr(C = c, I = i, R = r, T = t) = \Pr(R = r|T = t)\Pr(I = i|T = t)\Pr(C = c|I = i)\Pr(T = t). \quad (13)$$

All of the terms on the right-hand side of eqn. (13) have already been defined except for  $\Pr(T = t)$ , which pertains to the latent state of true vaccination status. To define this, we begin by specifying the relationship between  $\Pr(T = 1)$  and  $\Pr(R = 1)$ , which is

$$\Pr(R = 1) = \Pr(R = 1|T = 1)\Pr(T = 1) + \Pr(R = 1|T = 0)\Pr(T = 0). \quad (14)$$

Solving for  $\Pr(T = 1)$ , we obtain

$$\Pr(T = 1) = \frac{\Pr(R = 1) - \Pr(R = 1|T = 0)}{\Pr(R = 1|T = 1) - \Pr(R = 1|T = 0)}. \quad (15)$$

As  $\Pr(R = 1|T = 1)$  and  $\Pr(R = 1|T = 0)$  have already been defined, that leaves only  $\Pr(R = 1)$  to be defined. For this quantity, we used vaccination coverage estimates appropriate to each suspected case's home adm1, year, and age. Finally, to obtain a probability of the observed states only, we marginalize over states  $T$  and  $I$  in eqn. (13) to obtain

$$\Pr(C = c, R = r) = \sum_{t=0,1} \sum_{i=0,1} \Pr(R = r|T = t)\Pr(I = i|T = t)\Pr(C = c|I = i)\Pr(T = t), \quad (16)$$

which is the joint probability of the observed states  $(c, r)$  for each individual.

The likelihood of the model and its parameters consists of two parts. First, the contribution to the likelihood from data on the confirmed case ( $c_j$ ) and reported vaccination ( $r_j$ ) status across all individuals  $j$  is

$$\mathcal{L}_{data,k}(Se_{vac,k}, Sp_{vac,k}, Se_{test,k}, Sp_{test,k}, \lambda_k, \rho_k, VE|\{c_j\}, \{r_j\}) = \prod_j \Pr(C = c_j, R = r_j) \quad (17)$$

for country  $k$ . The variables  $\{c_j\}$  and  $\{r_j\}$  refer to the status of  $c$  and  $r$  across the set of individuals  $j$  within country  $k$ . The data in this analysis ultimately consist of counts of the number of individuals within each country with each possible combination of values of  $c$  (0 or 1) and  $r$  (0 or 1), because that determines how many times each  $\Pr(C = c, R = r)$  from the four possible combinations of  $c$  and  $r$  is multiplied in the likelihood in eqn. (17). Our analysis constitutes a test-negative design in the sense that it utilizes information on the proportions of individuals with a given vaccination status ( $r$ ) in groups distinguished by their case confirmation status ( $c$ ). Differences in the proportions between these groups provide information about the parameters, with  $Se_{vac}$  and  $Sp_{vac}$  being of primary interest.

Second, the contribution to the likelihood from the distribution of country-specific sensitivities and specificities of reported vaccination status is

$$\mathcal{L}_{vac}(\vec{\alpha}_{vac}|\{Se_{vac,k}\}, \{Sp_{vac,k}\}) = \prod_k \text{Dirichlet}(\{1 - Se_{vac,k}, 1 - Sp_{vac,k}, Se_{vac,k} + Sp_{vac,k} - 1\}|\vec{\alpha}_{vac}), \quad (18)$$

where  $\vec{\alpha}_{vac}$  contains three concentration parameters,  $\alpha_{vac,1}$ ,  $\alpha_{vac,2}$ , and  $\alpha_{vac,3}$ . This component of the likelihood allows for a balance to be achieved between the freedom of the  $Se_{vac,k}$  and  $Sp_{vac,k}$  parameters to vary among countries and the constraint that similar combinations of those parameters may apply to multiple countries. Mathematically, this balance is achieved by allowing the country-specific parameter values to follow a common distribution. The choice of a Dirichlet distribution was motivated by a desire to model the  $Se_{vac,k}$  and  $Sp_{vac,k}$  parameters jointly and to ensure that they were restricted to values between 0 and 1. The third implied parameter of this Dirichlet distribution,  $Se + Sp - 1$ , allowed us to specify a prior for the overall accuracy of case confirmation.

Together, these component likelihoods lead to

$$\mathcal{L}_{overall}(\{Se_{vac,k}\}, \{Sp_{vac,k}\}, \{Se_{test,k}\}, \{Sp_{test,k}\}, \{\lambda_k\}, \{\rho_k\}, VE, \vec{\alpha}_{vac}) = \prod_k \mathcal{L}_{data,k} \mathcal{L}_{vac} \quad (19)$$

as the overall likelihood of the model and its parameters. Notation such as  $\{Se_{vac,k}\}$  refers to the collection of  $Se_{vac,k}$  parameters across the set of countries  $k$ , and similar for other parameters specified for each country. In contrast to  $Se_{vac,k}$  and  $Sp_{vac,k}$ , the country-specific parameters  $Se_{test,k}$ ,  $Sp_{test,k}$ ,  $\lambda_k$ , and  $\rho_k$  were not modeled as belonging to common distributions and were allowed to vary freely among countries. Posterior estimates of these parameters were found to closely follow their prior distributions, suggesting that modeling them as belonging to a common distribution would have had only a negligible impact on our results.

## Priors

We specified informative priors for each parameter to the extent possible based on independently available information. For  $\lambda_k$ , we used posterior estimates of force of infection at the country level from a previous study [26]. Because that study produced separate posterior estimates under eight alternative scenarios, we fitted a normal distribution on the  $\log_{10}$  scale to posterior samples of force of infection for each country pooled across the eight scenarios. For  $\rho_k$ , we used posterior estimates from the same study [26] to inform our priors, albeit with the same prior for each country  $k$  since that study only obtained a single estimate of reporting probability. In this case, we fitted a normal distribution to logit-transformed values of reporting proportion estimates pooled across the eight scenarios. For  $VE$ , we used a posterior estimate from a previous study [27] to inform a beta-distributed prior, the parameters of which were obtained by matching to the previous study's median estimate of 0.975 and 95% credible interval of 0.83 - 1.0. For  $Se_{vac}$  and  $Sp_{vac}$ , we restricted the intensity parameters  $\alpha_{vac}$  to be larger than 1 to ensure a unimodal shape to the inferred distribution of  $Se_{vac}$  and  $Sp_{vac}$  across countries. Specifically, the priors of these parameters were all exponential(1) + 1. For  $Se_{test,k}$  and  $Sp_{test,k}$ , we selected values of Dirichlet intensity parameters that resulted in a median of 0.95 and 95% intervals of the marginals of  $Se_{test,k}$  and  $Sp_{test,k}$  equal to 0.8 - 1.0 and 0.6 - 1.0, respectively [28, 29, 30]. This was obtained with Dirichlet intensity parameters of 0.145, 0.569, and 5.359 for  $1 - Se_{test,k}$ ,  $1 - Sp_{test,k}$ , and  $Se_{test,k} + Sp_{test,k} - 1$ .

## Hamiltonian Monte Carlo

We used the likelihood from eqn. (19) and the priors described above as the basis for obtaining Bayesian estimates of the model's parameters using Hamiltonian Monte Carlo implemented through the rstan package [31] in R 4.2.1 [32]. We ran the models with  $10^5$  iterations and a burn-in period of  $5 \times 10^4$  iterations. Convergence was assessed using  $\hat{R}$  estimates from the rstan package [31] and trace plots for all parameters. We specified the threshold for convergence to be  $\hat{R} < 1.05$ .

## Model comparison

To account for uncertainty in some of the choices we made in the design of our model, we compared our full model described above to 15 other models that represent simpler cases of it. Specifically, we constructed alternative models that simplified the parameters that we allowed to vary geographically, with the alternative models treating different combinations of geographically variable parameters as constants. For example, instead of considering country-specific values of  $Se_{vac,k}$  and  $Sp_{vac,k}$ , we considered an alternative with the same  $Se_{vac}$  and  $Sp_{vac}$  applying to all countries. We did the same

for  $Se_{test}$  and  $Sp_{test}$ ,  $\lambda$ , and  $\rho$ . This resulted in a total of 16 different models that explored all combinations of the ways in which geographic variability could be toggled on or off across these four sets of parameters. We compared these 16 models on the basis of the Watanabe-Akaike information criterion (WAIC). We were also interested in assessing how parameter estimates differed across models.

## Results

### Model comparison

We fitted all 16 models to the data and determined that all models converged, as evidenced by  $\hat{R} < 1.05$  and visual inspection of trace plots. We found that the default model, which is the most complex model with four sets of geographically variable parameters, performed among the best on the basis of WAIC (Fig. 2). We observed minimal differences in WAIC among models that did or did not allow for geographic variability in force of infection or reporting proportion. Notably, geographic variability in  $Se_{vac}$  and  $Sp_{vac}$  was the most important model feature for enabling a low WAIC. This was followed by geographic variability in  $Se_{test}$  and  $Sp_{test}$  as the second most important model feature, but to a lesser extent (Fig. 2).

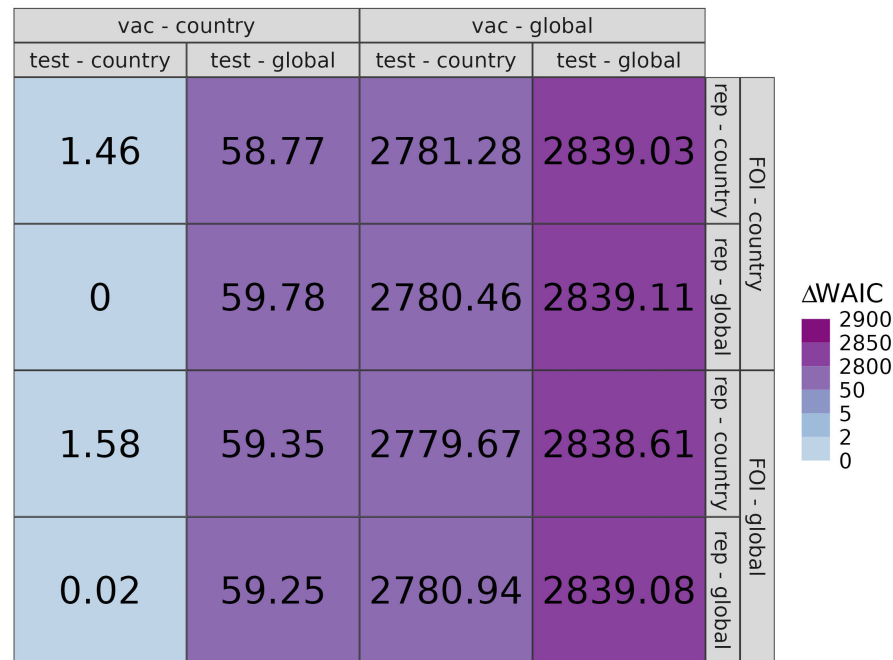


Figure 2:  $\Delta$ WAIC values of the 16 models.  $\Delta$ WAIC is defined as the difference between the WAIC value of a given model and that of the model with the lowest WAIC value. The default model (top left) includes geographic variability in  $Se_{vac}$  and  $Sp_{vac}$  (vac),  $Se_{test}$  and  $Sp_{test}$  (test), force of infection (FOI), and reporting proportion (rep). In the columns and rows, geographic variability is indicated by whether the parameters vary by country or do not vary and are thus global constants.

### Estimation of $Se_{vac}$ and $Sp_{vac}$

The default model showed considerable variability in estimates of the sensitivity and specificity of vaccination status across countries, with median estimates of 0.69 (95% CrI: 0.21 - 0.98) and 0.70 (95% CrI: 0.21 - 0.98), respectively. Across the 20 countries that we considered, country-specific estimates ranged from medians of 0.15 (95% CrI: 0.09 - 0.23) to 0.98 (95% CrI: 0.90 - 1.0) for sensitivity, and from medians of 0.06 (95% CrI: 0.04 - 0.09) to 0.96 (95% CrI: 0.94 - 0.98) for specificity (Fig. 3). The Pearson correlation between  $Se_{vac,k}$  and  $Sp_{vac,k}$  across countries had a median value of  $-0.85$

(95% CrI:  $-0.924 - -0.711$ ), indicative of a trade-off between the sensitivity and specificity of differing practices of vaccination status classification among countries. Notably, estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$  for several countries fell near the diagonal line, suggesting that classification of vaccination status was only marginally better than chance.

We observed that the amount of uncertainty associated with estimates of country-specific values of  $Se_{vac,k}$  and  $Sp_{vac,k}$  was influenced by the number of data points from each country. Specifically, countries with very few samples (i.e.,  $\leq 6$ ) exhibited markedly higher levels of uncertainty in their estimates (Fig. S1). To assess the robustness of our results to influences from particular countries, we performed a sensitivity analysis in which we excluded the top one to five countries with the most data points. The results of this analysis indicated that overall estimates of the distribution of  $Se_{vac}$  and  $Sp_{vac}$  across countries remained largely unchanged (Fig. S2), supporting the robustness of our primary results.

Among the eight alternative models that allowed for geographic variability in  $Se_{vac,k}$  and  $Sp_{vac,k}$ , we found minimal differences in estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$  (Fig. S3). The negative correlation between  $Se_{vac,k}$  and  $Sp_{vac,k}$  across countries that we observed for the default model was also apparent for the other seven models (Table S1, Fig. S4). For models that lacked geographic variability in  $Se_{vac}$  and  $Sp_{vac}$ , estimates of  $Se_{vac}$  and  $Sp_{vac}$  were all similar. For example, the simplest model (Fig. S3, bottom right), yielded median estimates of  $Se_{vac}$  and  $Sp_{vac}$  of 0.49 (95% CrI: 0.48 - 0.50) and 0.71 (95% CrI: 0.70 - 0.72), respectively.

To validate our inferences, we repeated our inference procedure with the default model on ten data sets simulated with parameters set to the medians of their marginal posterior distributions. In doing so, we observed broad consistency between the country-level estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$  and their simulated values (Fig. S5). In addition to alignment in the central tendencies of the posteriors obtained from simulated and empirical data sets, the spans of their uncertainty intervals were also similar.

## Estimation of other parameters

Unlike  $Se_{vac,k}$  and  $Sp_{vac,k}$ , estimates of  $Se_{test,k}$  and  $Sp_{test,k}$  varied negligibly across countries in the default model (Fig. S6). Specifically, the country-specific  $Se_{test,k}$  estimates ranged from medians of 0.99 (95% CrI: 0.82 - 1.00) to 1.00 (95% CrI: 0.81 - 1.00), and  $Sp_{test,k}$  estimates ranged from medians of 0.96 (95% CrI: 0.66 - 1.00) to 1.00 (95% CrI: 0.99 - 1.00). We consistently observed a similar pattern across all seven country-specific models, where the estimates for each country were similar across different models (Fig. S7, eight left panels). For the simplest model (Fig. S7, bottom right), which assumes no geographic variation in any parameter, the median estimates of  $Se_{test}$  and  $Sp_{test}$  were 0.99 (95% CrI: 0.82 - 1.00) and 0.99 (95% CrI: 0.98 - 0.99), respectively. In general, we observed that all estimates of  $Se_{test}$  and  $Sp_{test}$  overlapped significantly with their prior distributions, which had medians of 0.98 (95% CrI: 0.77 - 1.00) and 0.97 (95% CrI: 0.58 - 1.00), respectively. Posterior estimates of other parameters—i.e.,  $\rho_k$ ,  $\lambda_k$ , and  $VE$ —also closely resembled their prior distributions (Figs. S8-S10).

Given that our model allows for multiple sources of error in observed data, there is potential for some parameter combinations to not be identifiable. To assess the ability of the model to obtain separate, uncorrelated estimates of the sensitivities and specificities of vaccination status and testing, we first examined the correlations of their posterior samples within each of the 20 countries in our analysis. Correlations between  $Se_{vac,k}$  and  $Se_{test,k}$  ranged from  $-0.10$  to  $0.09$ , as did correlations between  $Sp_{vac,k}$  and  $Sp_{test,k}$ . As a further test of the identifiability of these parameters, we generated another ten simulated data sets, but with lower values of  $Se_{test,k}$  and  $Sp_{test,k}$  (0.70 for both) than their median posterior estimates. Even under these lower values of  $Se_{test,k}$  and  $Sp_{test,k}$ , our inferences of  $Se_{vac,k}$  and  $Se_{test,k}$  remained similar (Fig. S11).

We also explored the extent to which misestimation of other parameters—namely, the force of infection,  $\lambda_k$ —could bias estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$ . To test this, we simulated ten data sets with values of  $\lambda_k$  varying randomly from  $10^{-5}$  to  $10^{-1}$  by factors of 10 across the 20 countries in our analysis. When we applied our inference method to those data, we found that estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$  were similar to those from our primary analysis (Fig. S11). Moreover, estimates of  $\lambda_k$  also resembled those from the primary analysis (Fig. S12). In both sets of analyses, posterior estimates of  $\lambda_k$  were driven strongly by their priors.

## Vaccination coverage adjusted for misclassification

By using the default model estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$ , we were also able to adjust for bias in estimates of vaccination coverage. We did this by calculating adjusted vaccination coverage using eqn. (15), where reported vaccination coverage comes from population-weighted vaccination coverage estimates from POLICI between 1980 and 2023 (see Fig. 4).

Our findings revealed that adjusted vaccination coverage could either increase or decrease relative to reported coverage, depending on country-specific estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$  (Fig. 4). Specifically, our analysis revealed that in ten countries, the 2.5% quantile of adjusted vaccination coverage exceeded reported coverage, implying underestimation of reported vaccination coverage. These countries exhibited a pattern of low  $Se_{vac,k}$  and high  $Sp_{vac,k}$ . In five countries, the 97.5% quantile of adjusted vaccination coverage was lower than reported coverage, implying overestimation of reported vaccination coverage. These countries exhibited a pattern of high  $Se_{vac,k}$  but low  $Sp_{vac,k}$ . In some countries, we noted that bias of reported vaccination coverage may be negligible, such as when reported coverage is very low. Uncertainty around adjusted vaccination coverage was driven by the degree of uncertainty about  $Se_{vac,k}$  and  $Sp_{vac,k}$ , which was in turn influenced by the number of data points from each country (Fig. 4).



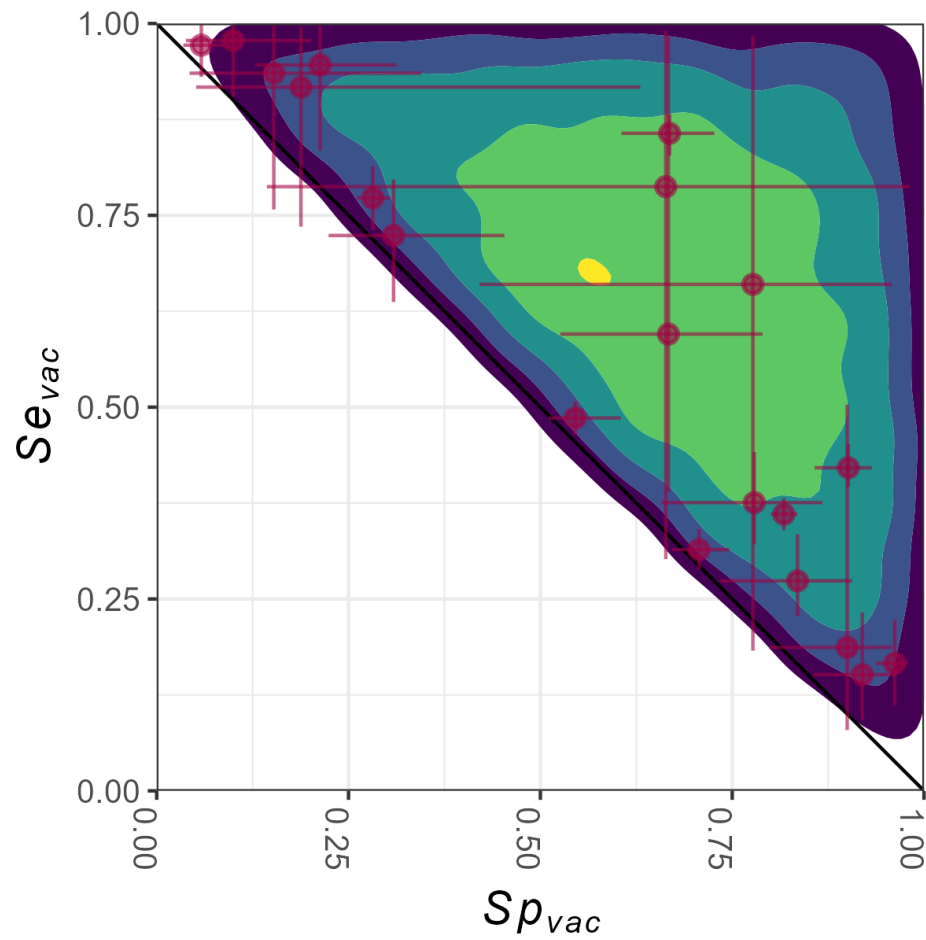


Figure 3: **Estimates of  $Se_{vac}$  and  $Sp_{vac}$  from the default model.** The points with their lines represent the median estimates and the 95% CrI of country-specific estimates  $Se_{vac,k}$  and  $Sp_{vac,k}$ . The shades indicate the density of  $Se_{vac}$  and  $Sp_{vac}$  at a global level, where lighter shades indicate higher densities.

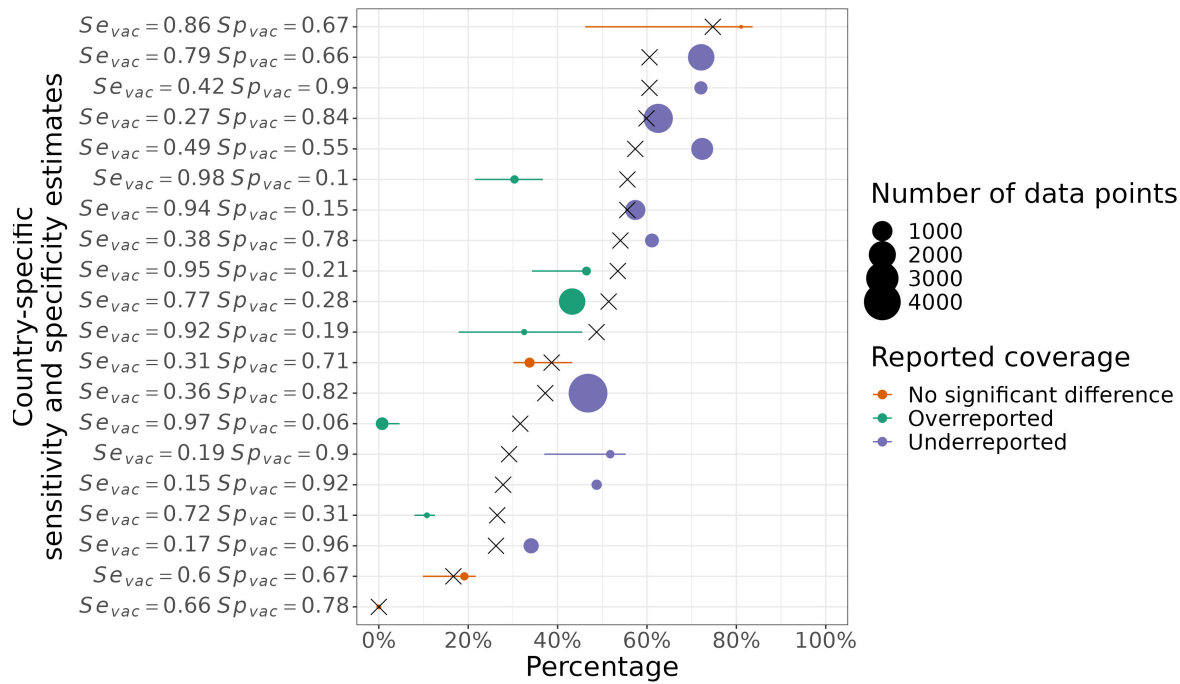


Figure 4: **Reported versus adjusted vaccination coverage.** Reported vaccination coverages from POLICI are shown with an “X”. The horizontal points and lines indicate the medians and 95% posterior predictive intervals of adjusted vaccination coverage, colored by the direction of bias (underestimate, overestimate, or no significant difference). The size of each point corresponds to the number of data points from each country. The countries are arranged from bottom to top based on their reported vaccination coverage. The sensitivity and specificity of vaccination status for each country is shown on the left.

## Discussion

Our study aimed to quantify the degree of misclassification associated with yellow fever vaccination status using a test-negative case-control study design with surveillance data on yellow fever in Africa. We found that vaccination status misclassification appears to be prevalent, with country-level estimates of sensitivity and specificity of vaccination status being highly negatively correlated and with the average of sensitivity and specificity generally not exceeding 0.70. By quantifying these inaccuracies, we were able to obtain adjusted estimates of vaccination coverage that account for vaccination status misclassification. Our estimates of adjusted vaccination coverage indicated that reported vaccination coverage may be an underestimate in ten out of twenty countries, an overestimate in five, and approximately correct in five.

We observed considerable variability in our estimates of the sensitivity and specificity of vaccination status across the 20 countries that we considered, which could be attributable to differences in vaccination recording systems across those countries. To remedy the unique issues with each country's vaccination recording system and improve the quality of vaccination coverage data, the WHO has developed guidelines to help identify weaknesses in vaccination recording systems and improve them [33, 34, 35]. Whatever the specific causes of this country-level variation are, this variation is unlikely to be explainable based on any readily available or predictable factors. As a result, it would be difficult to extrapolate our estimates to countries not included in our analysis, and efforts should be made to extend our analysis with additional data.

Analyses of simulated data suggest that inferences of the sensitivity and specificity of vaccination status are valid but that inferences of other parameters closely mirror prior assumptions. A primary driver of this asymmetry in parameter identifiability is that our data—which ultimately come down to the proportion vaccinated among test-negative versus test-positive individuals—are influenced more by the sensitivity and specificity of vaccination status than other parameters. Given that the data are limited to suspected cases who underwent confirmatory testing, they simply do not contain much information about other parameters, such as the force of infection or reporting probability. Reassuringly, the prior distributions for those parameters were well-informed by a previous study [26], for which estimation of those parameters was of primary interest.

Using our estimates of the sensitivity and specificity of vaccination status, we quantified discrepancies between reported and adjusted vaccination coverage. Estimates of vaccination coverage are important for estimating the burden of yellow fever, because of their implications for the use of serological data to infer the force of infection [22]. Because seropositivity as a result of natural infection or vaccination cannot be differentiated empirically, estimates of vaccination coverage are important for determining how much of measured seropositivity is attributable to natural infection [26]. If vaccination coverage is higher than reported, then force of infection may be overestimated. Conversely, if vaccination coverage is lower than reported, then force of infection may be underestimated. Such biases in estimates of the force of infection are expected to lead to corresponding biases in estimates of disease burden, which could be consequential for prioritizing investments in different vaccine-preventable diseases [2].

There are important limitations of our analysis to note, especially in relation to the state of currently available data. Of the more than 15,000 suspected cases in the database that we analyzed, only 206 were confirmed yellow fever cases. Within the same timeframe of 2008 to 2013 in the 20 countries included in our database, a total of 3,590 confirmed cases were reported by another source [36], suggesting that the data we analyzed may be very incomplete. In addition, there is substantial imbalance in the sample sizes from different countries, which could affect our global estimates of the sensitivity and specificity of vaccination status. At the same time, our sensitivity analysis suggested that omission of data from countries with the most data resulted in minimal changes to our global estimates. There was, nevertheless, high uncertainty in parameters for countries with low sample sizes. Moreover, the database that we used includes only 20 of the more than 30 countries in Africa where there is a risk of yellow fever. There are also limitations of the vaccination coverage estimates that we used to inform  $\Pr(R = 1)$ . Reported vaccination coverage may vary at finer temporal and spatial scales than the estimates available to us, meaning that the values of  $\Pr(R = 1)$  that we used may be imperfect approximations and could potentially lead to bias in our estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$ .

In the future, the analytical framework we used could be applied to evaluate country-level vaccination reporting systems for other diseases. The requirements for doing so would be the same type of data on vaccination and infection statuses for suspected cases of whatever disease is of interest.

Replication of our analysis for other diseases could shed light on the consistency (or not) of the sensitivity and specificity of vaccination status across multiple diseases, leading to a unique approach for comprehensive assessment of a given country's vaccination recording system. Accurate understanding of vaccination coverage is important for achieving public health goals for a variety of other infectious diseases, such as measles [37] and polio [38].

## Acknowledgements

We acknowledge the World Health Organization for sharing its yellow fever surveillance database for use in this analysis. This work was carried out as part of the Vaccine Impact Modelling Consortium ([www.vaccineimpact.org](http://www.vaccineimpact.org)), but the views expressed are those of the authors and not necessarily those of the Consortium or its funders. The funders were given the opportunity to review this paper prior to publication, but the final decision on the content of the publication was taken by the authors. This work was supported, in whole or in part, by the Bill & Melinda Gates Foundation, via the Vaccine Impact Modelling Consortium [Grant Number INV-009125]. Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript version that might arise from this submission.

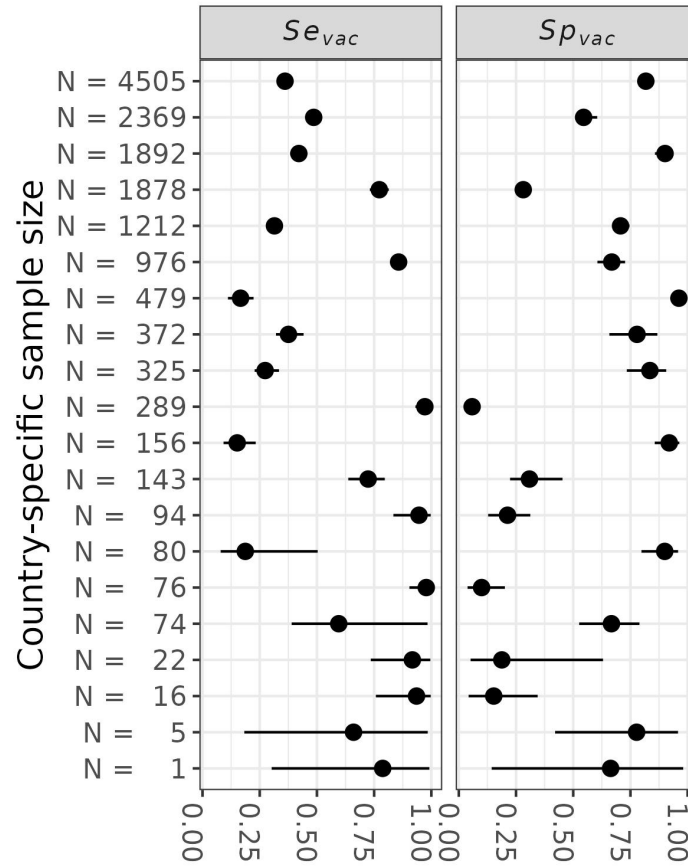


Figure S1: Relationship between the number of data points from each country (y-axis) and uncertainty in the estimates of  $Se_{vac}$  (left) and  $Sp_{vac}$  (right) for the default model. Points and lines represent the posterior medians and 95% credible intervals of  $Se_{vac}$  and  $Sp_{vac}$  for each country.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

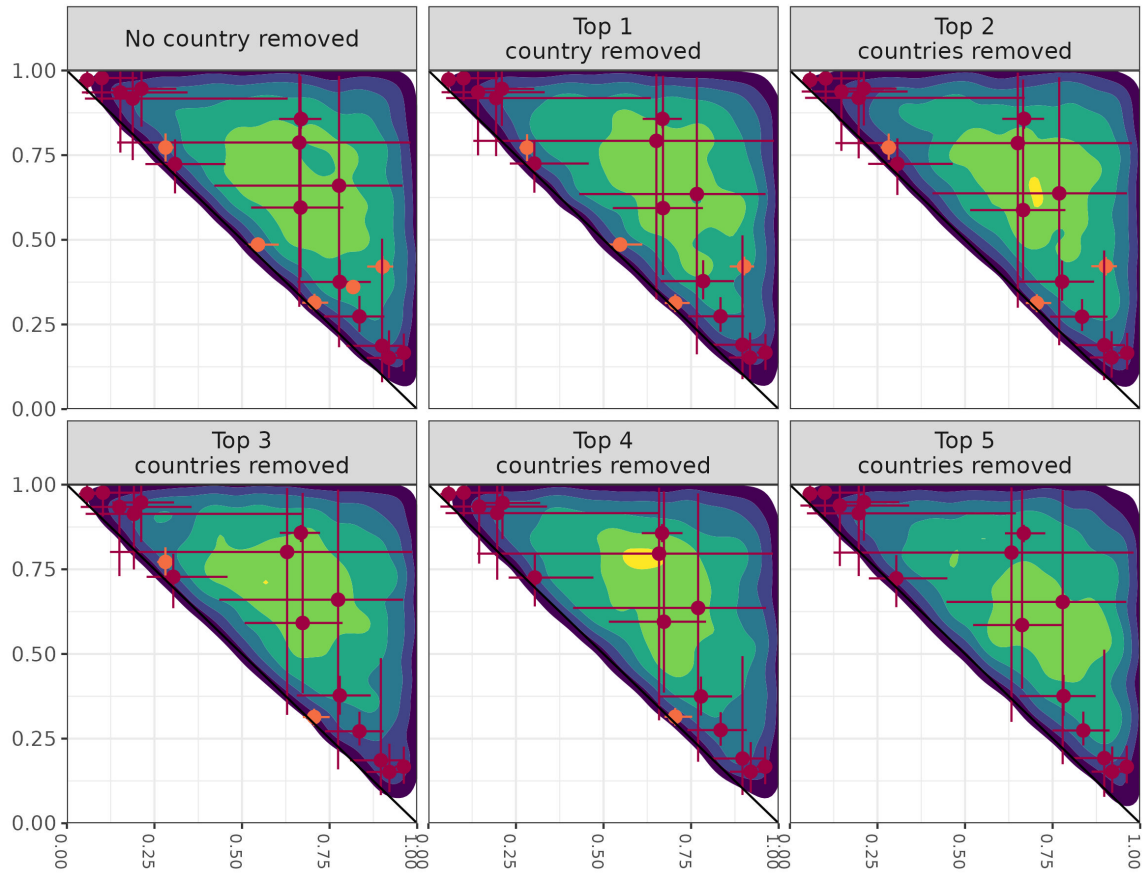


Figure S2: **Estimates of  $Se_{vac}$  and  $Sp_{vac}$  from the default model after excluding the top five countries with the most data points.** The dots and lines represent posterior medians and 95% credible intervals of country-specific estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$ . The orange dots and lines indicate the top five countries with the most data points. Lighter shading indicates higher probability densities.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

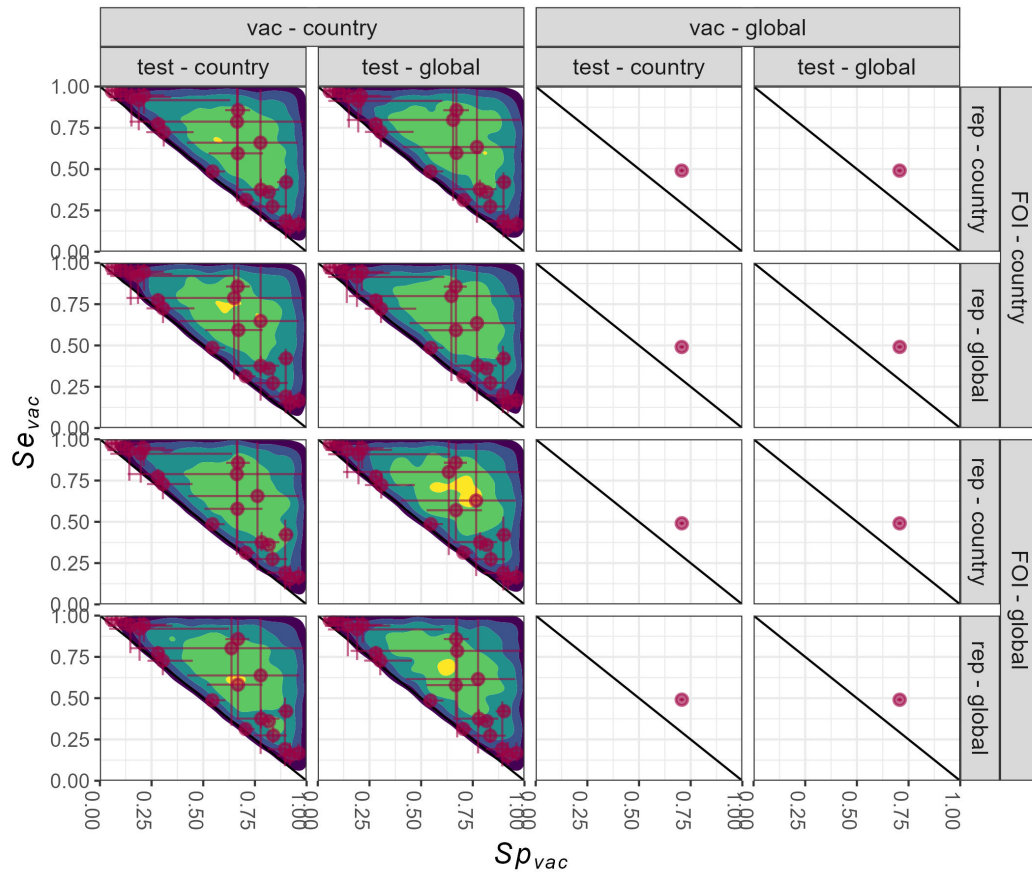


Figure S3: **Estimates of  $Se_{vac}$  (y-axis) and  $Sp_{vac}$  (x-axis) from the 16 models (panels).** In the columns and rows, geographic variability is indicated by whether parameters vary by country or do not vary and are thus global constants. The default model (top left) includes geographic variability in  $Se_{vac}$  and  $Sp_{vac}$  (vac),  $Se_{test}$  and  $Sp_{test}$  (test), force of infection (FOI), and reporting proportion (rep). The dots and lines represent posterior medians and 95% credible-intervals of country-specific estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$ . Lighter shading indicates higher probability densities.

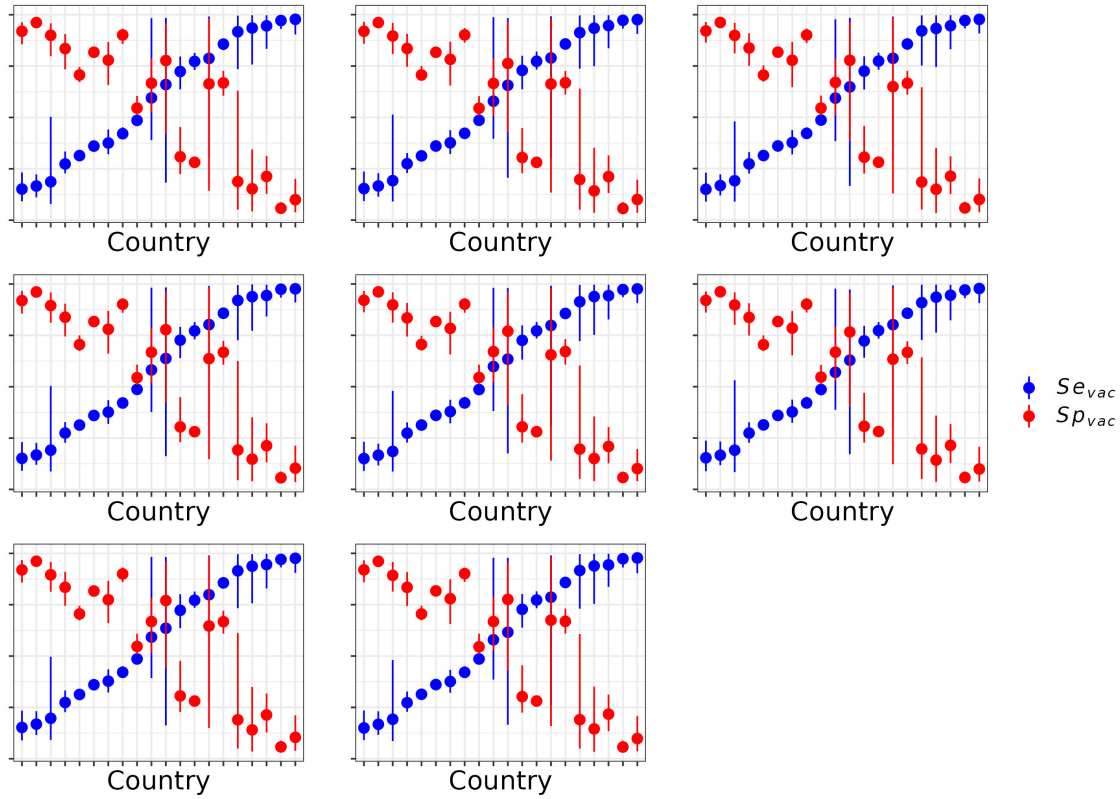


Figure S4: Marginal posterior distributions of  $Se_{vac}$  (blue) and  $Sp_{vac}$  (red) from the eight models (panels) that accounted for geographic variability in those parameters. The dots and lines represent medians and 95% credible intervals of country-specific estimates of  $Se_{vac,k}$  and  $Sp_{vac,k}$ . The country-specific estimates are sorted in ascending order based on  $Se_{vac,k}$ . A noticeable trend of negative correlation between  $Se_{vac,k}$  and  $Sp_{vac,k}$  is observed for all models.



It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

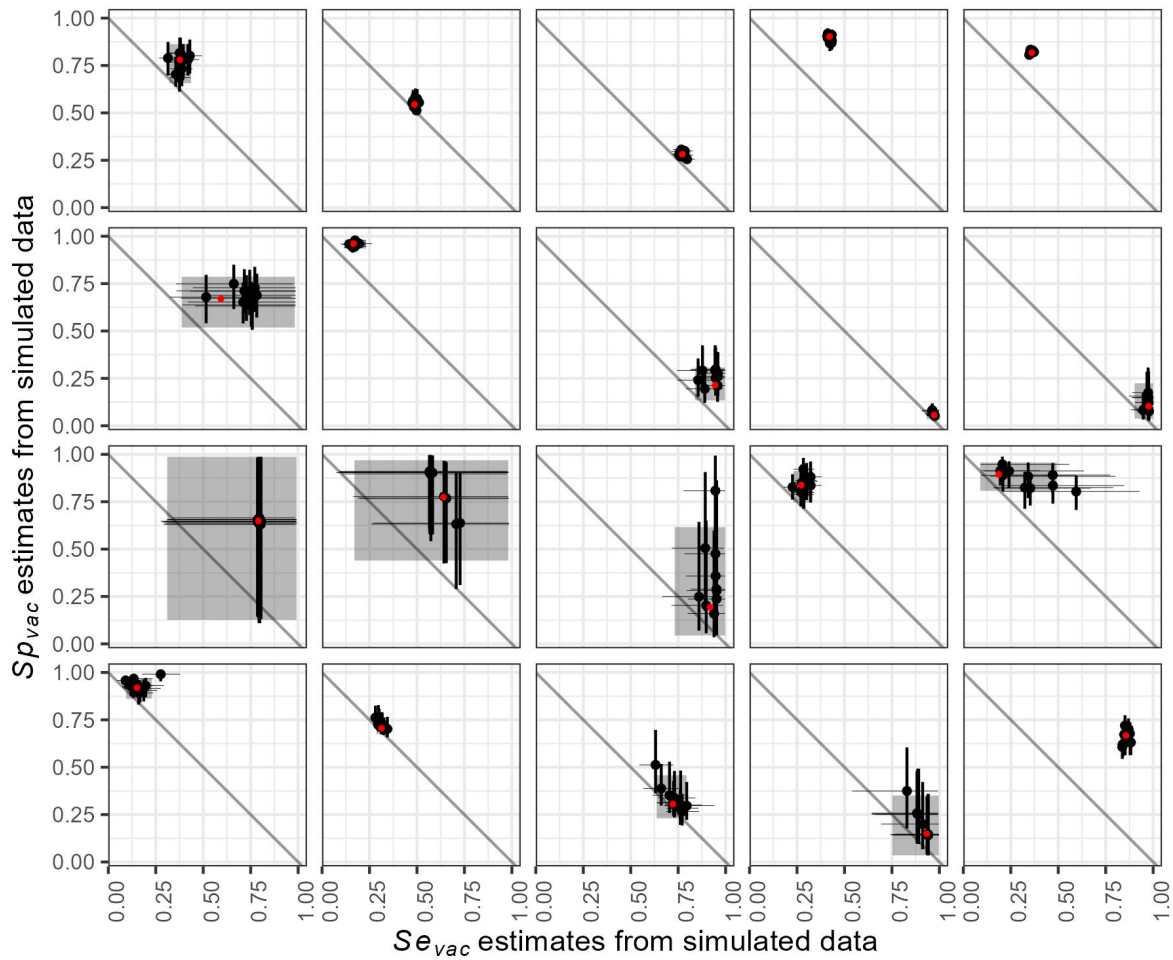


Figure S5: Estimation of  $Se_{vac,k}$  (x-axis) and  $Sp_{vac,k}$  (y-axis) for each of 20 countries (panels) from ten data sets simulated with posterior median parameter values. The ten black dots in each panel and their lines represent median estimates and 95% credible intervals obtained from our analysis of the ten simulated data sets. The red dots and gray areas represent median posterior estimates and their corresponding 95% credible intervals obtained from our analysis of empirical data. The red dots were also the parameter values used to generate the ten simulated data sets.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

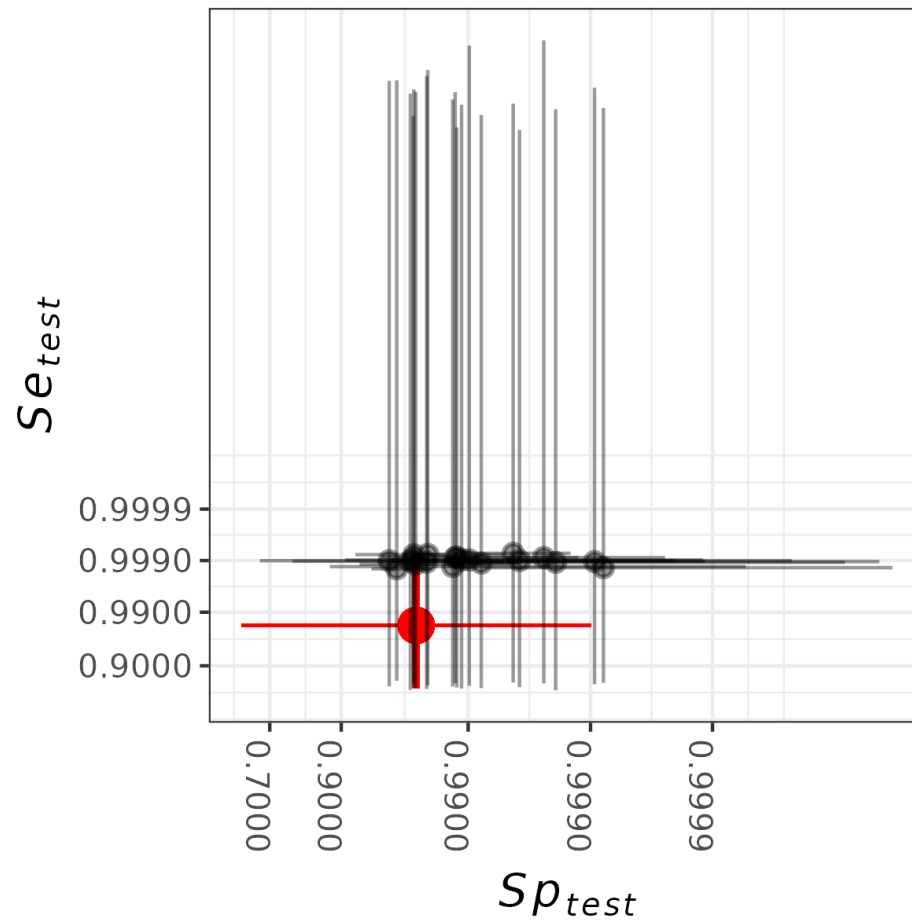


Figure S6: **Estimates of  $Se_{test}$  and  $Sp_{test}$  from the default model.** The points with their lines represent the median estimates and the 95% CrI of country-specific estimates  $Se_{test,k}$  and  $Sp_{test,k}$ . The red points and lines indicate the medians and 95% uncertainty intervals of the priors used.

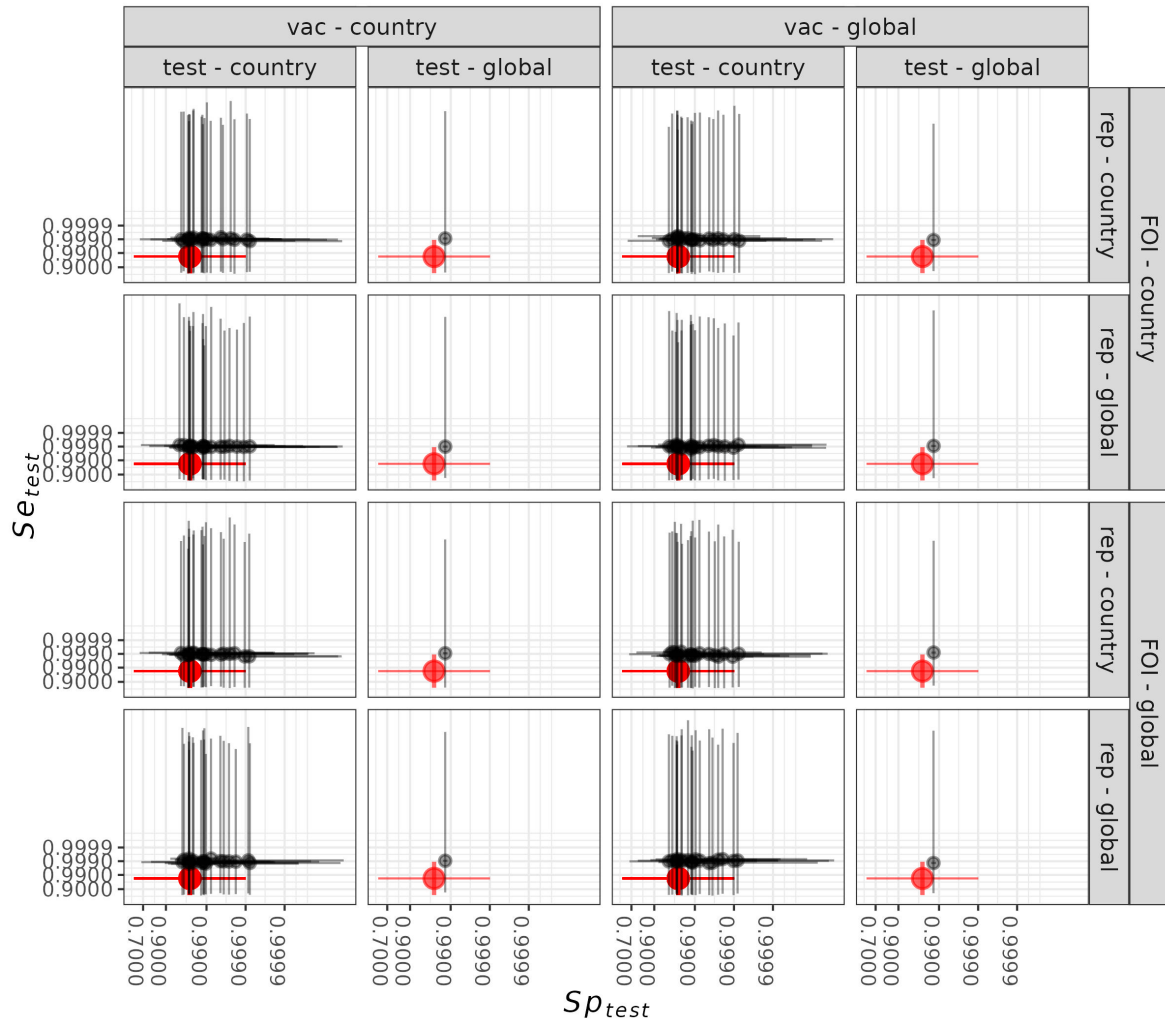


Figure S7: Estimates of  $Se_{test}$  (y-axis) and  $Sp_{test}$  (x-axis) from the 16 models (panels). In the columns and rows, geographic variability is indicated by whether parameters vary by country or do not vary and are thus global constants. The default model (top left) includes geographic variability in  $Se_{vac}$  and  $Sp_{vac}$  (vac),  $Se_{test}$  and  $Sp_{test}$  (test), force of infection (FOI), and reporting proportion (rep). The dots and lines represent posterior median estimates and the 95% credible intervals of country-specific estimates  $Se_{test,k}$  and  $Sp_{test,k}$ . The red points and lines indicate the medians and 95% uncertainty intervals of the priors.

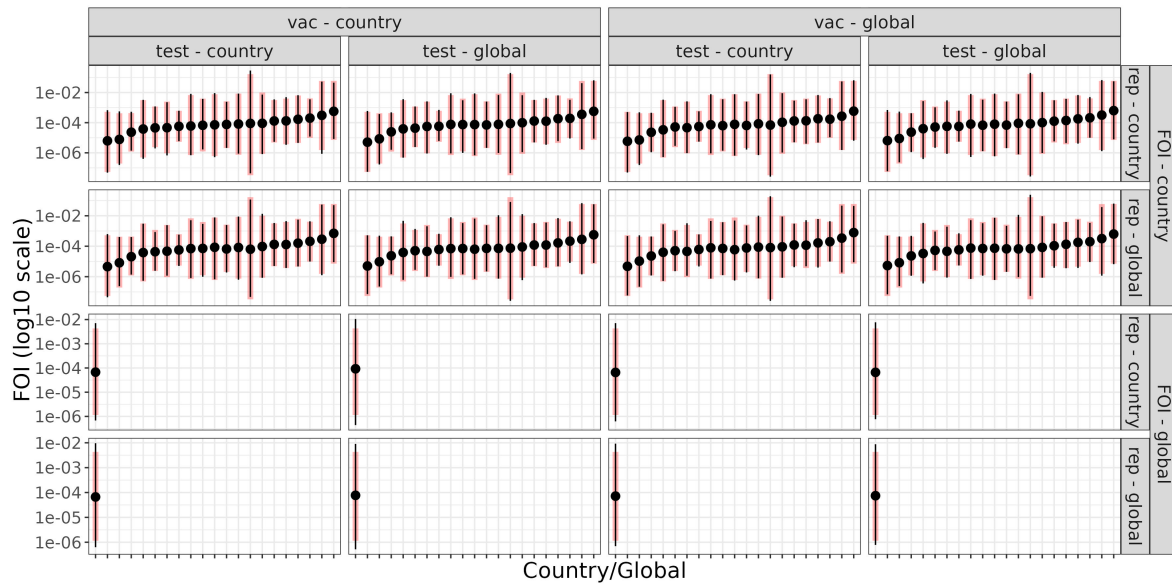


Figure S8: **Estimates of force of infection,  $\lambda_k$ , (y-axis) from the 16 models (panels).** Dots and lines represent posterior medians and 95% credible intervals of  $\lambda_k$  for each country. Vertical red lines represent the 95% intervals of the prior distributions, which were obtained from a previous study [26]. The default model (top left) includes geographic variability in  $Se_{vac}$  and  $Sp_{vac}$  (vac),  $Se_{test}$  and  $Sp_{test}$  (test), force of infection (FOI), and reporting proportion (rep). In the columns and rows, geographic variability is indicated by whether parameters vary by country or do not vary and are thus global constants.

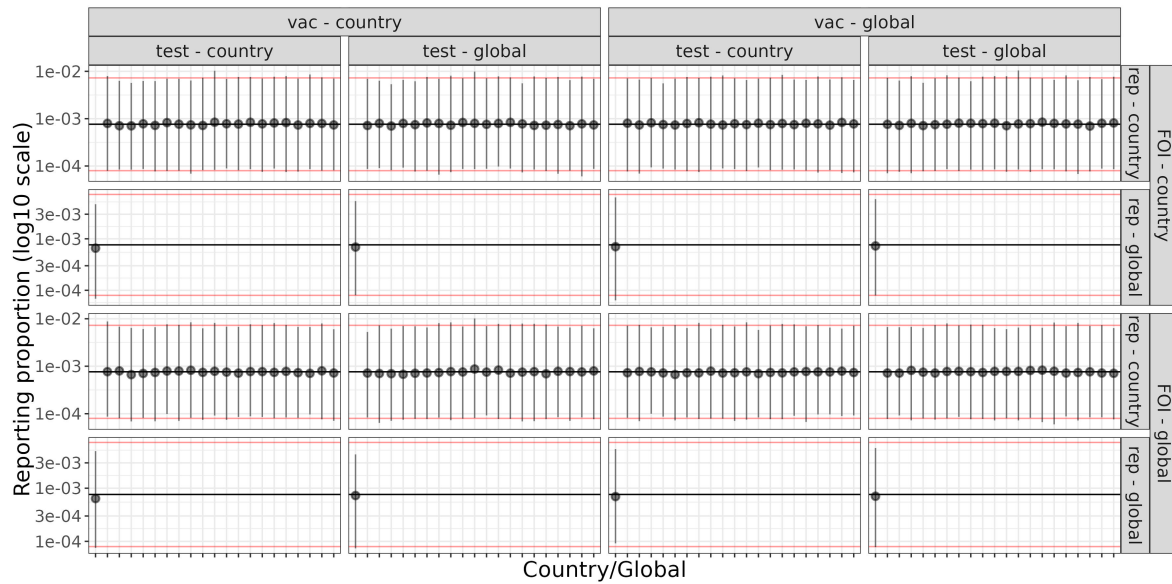


Figure S9: **Estimates of reporting proportion,  $\rho_k$ , from the 16 models (panels)**. Black dots and lines represent posterior medians and 95% credible intervals of  $\rho_k$  for each country (or for the global scale, depending on the model). Red lines represent the 95% interval and the horizontal black lines represent the median of the prior distribution, which was obtained from a previous study [26]. The default model (top left) includes geographic variability in  $Se_{vac}$  and  $Sp_{vac}$  (vac),  $Se_{test}$  and  $Sp_{test}$  (test), force of infection (FOI), and reporting proportion (rep). In the columns and rows, geographic variability is indicated by whether parameters vary by country or do not vary and are thus global constants.

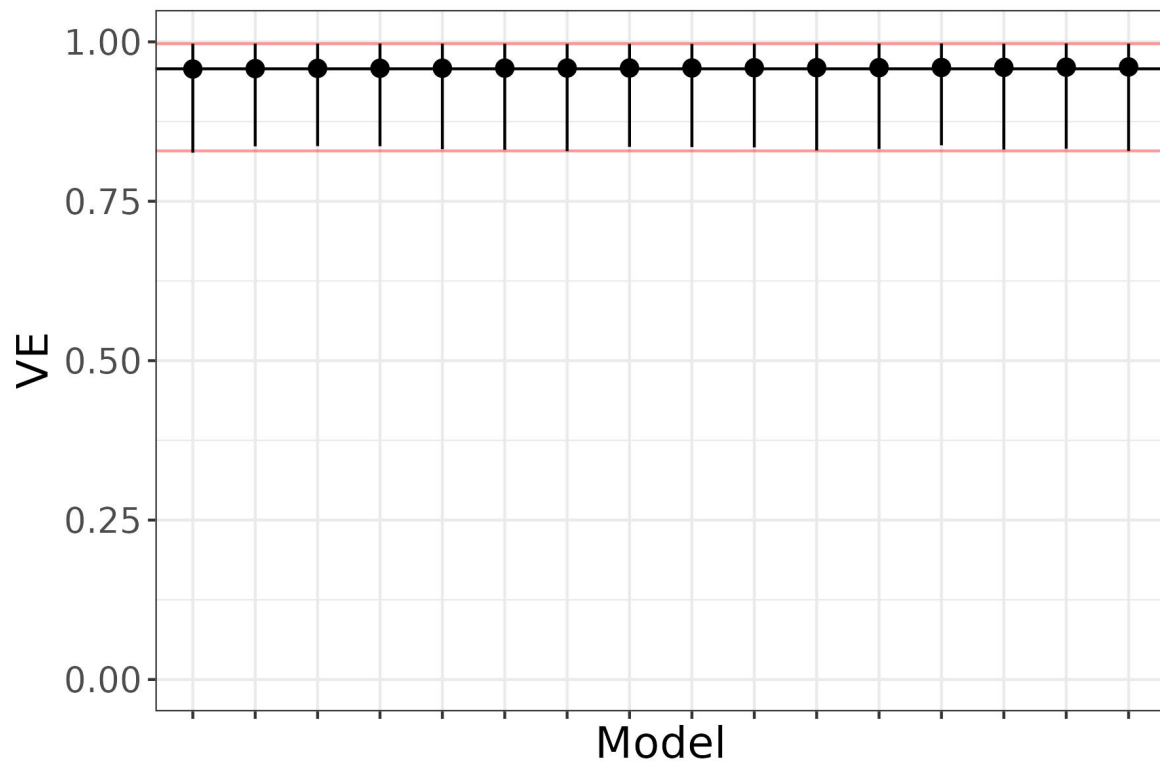


Figure S10: **Estimates of vaccine efficacy,  $VE$ , (y-axis) from all 16 models (x-axis).** Black dots and lines depict posterior medians and 95% credible intervals for  $VE$ . The horizontal red lines represent the 95% interval of the prior distribution, and the black line represents the median of the prior distribution.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

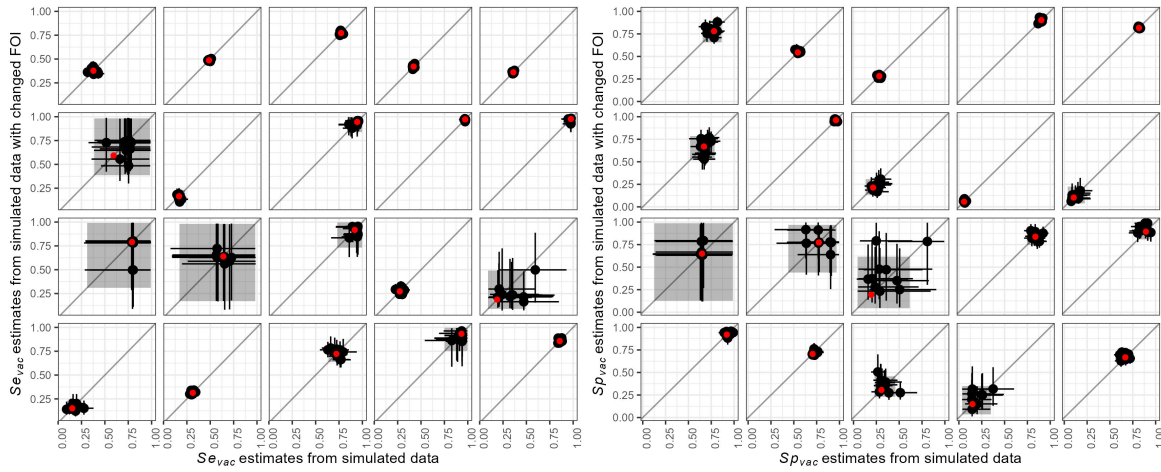


Figure S11: Estimates of  $Se_{vac,k}$  (left) and  $Sp_{vac,k}$  (right) from simulated data with (y-axis) or without (x-axis) changed values of  $\lambda_k$  for each of 20 countries (panels). In the unchanged scenario, we simulated ten data sets with posterior median parameter values. In the changed scenario, we simulated ten data sets with  $\lambda_k$  set to values of either  $10^{-5}$ ,  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$ , or  $10^{-1}$ . The ten black dots in each panel and their lines represent median estimates and 95% credible intervals obtained from our analysis of the ten simulated data sets. The red dots and gray areas represent median posterior estimates and their corresponding 95% credible intervals obtained from our analysis of empirical data.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

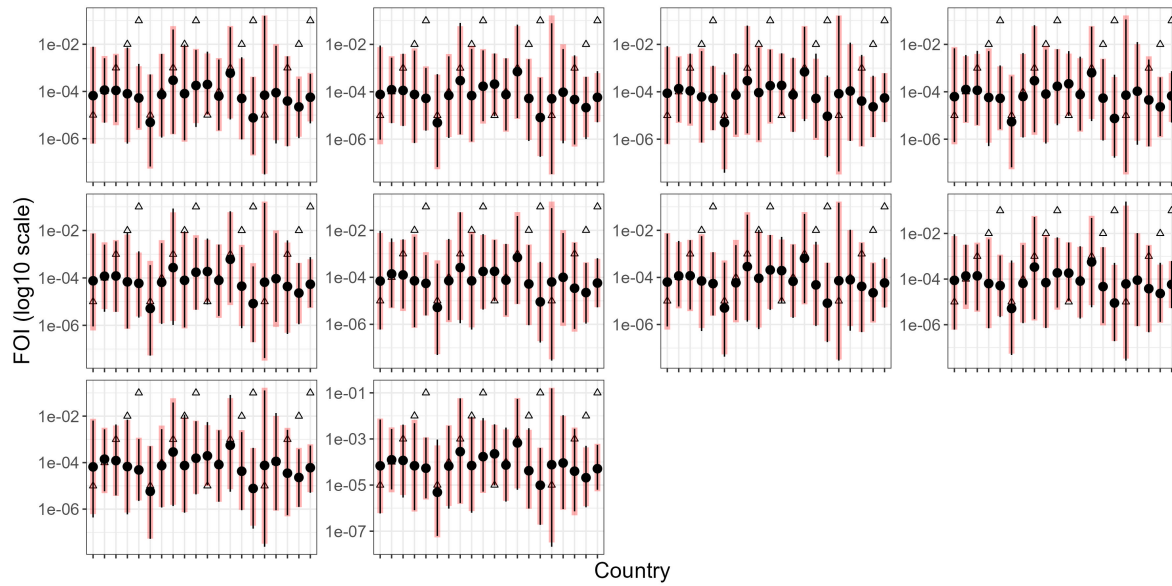


Figure S12: Estimates of force of infection,  $\lambda_k$ , (y-axis) for each of 20 countries (x-axis) in ten simulated data sets (panels). Black dots and vertical lines denote posterior median and 95% credible intervals. Vertical red lines represent the 95% intervals of the prior distributions, sourced from a previous study [26]. Triangles correspond to the values of  $\lambda_k$  used to generate the simulated data.



It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

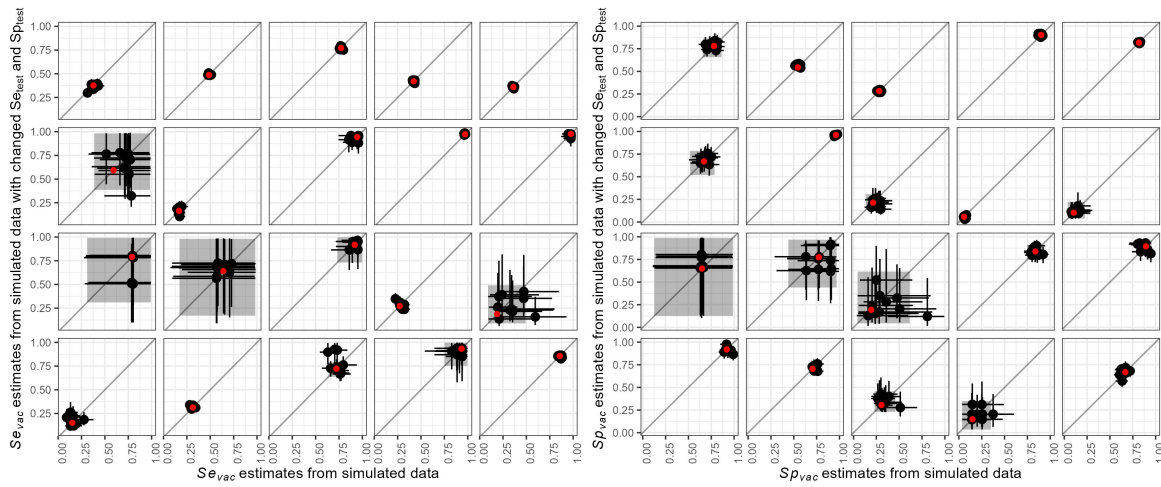


Figure S13: Estimates of  $Se_{vac,k}$  (left) and  $Sp_{vac,k}$  (right) from simulated data with (y-axis) or without (x-axis) changed values of  $Se_{test,k}$  and  $Sp_{test,k}$  for each of 20 countries (panels). In the unchanged scenario, we simulated ten data sets with posterior median parameter values. In the changed scenario, we simulated ten data sets with  $Se_{test,k}$  and  $Sp_{test,k}$  set to 0.7. The ten black dots in each panel and their lines represent median estimates and 95% credible intervals obtained from our analysis of the ten simulated data sets. The red dots and gray areas represent median posterior estimates and their corresponding 95% credible intervals obtained from our analysis of empirical data.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

Table S1: Correlation between  $Se_{vac,k}$  and  $Sp_{vac,k}$  across countries in eight models that account for geographic variability in  $Se_{vac}$  and  $Sp_{vac}$ .

$Se_{test}$ and $Sp_{test}$	FOI	Reporting proportion	Correlation
country	country	country	-0.852 (-0.924 - -0.711)
country	country	global	-0.853 (-0.926 - -0.720)
country	global	country	-0.854 (-0.925 - -0.726)
country	global	global	-0.855 (-0.922 - -0.728)
global	country	country	-0.855 (-0.925 - -0.709)
global	country	global	-0.854 (-0.924 - -0.718)
global	global	country	-0.856 (-0.925 - -0.727)
global	global	global	-0.857 (-0.923 - -0.726)

## References

- [1] WHO. Global framework for immunization monitoring and surveillance, November 2007.
- [2] Xiang Li, Christinah Mukandavire, Zulma M. Cucunubá, Susy Echeverria Londono, Kaja Abbas, Hannah E. Clapham, Mark Jit, Hope L. Johnson, Timos Papadopoulos, Emilia Vynnycky, Marc Brisson, Emily D. Carter, Andrew Clark, Margaret J. de Villiers, Kirsten Eilertson, Matthew J. Ferrari, Ivane Gamkrelidze, Katy A. M. Gaythorpe, Nicholas C. Grassly, Timothy B. Hallett, Wes Hinsley, Michael L. Jackson, Kévin Jean, Andromachi Karachaliou, Petra Klepac, Justin Lessler, Xi Li, Sean M. Moore, Shevanthi Nayagam, Duy Manh Nguyen, Homie Razavi, Devin Razavi-Shearer, Stephen Resch, Colin Sanderson, Steven Sweet, Stephen Sy, Yvonne Tam, Hira Tanvir, Quan Minh Tran, Caroline L. Trotter, Shaun Truelove, Kevin van Zandvoort, Stéphane Verguet, Neff Walker, Amy Winter, Kim Woodruff, Neil M. Ferguson, and Tini Garske. Estimating the health impact of vaccination against ten pathogens in 98 low-income and middle-income countries from 2000 to 2030: a modelling study. *The Lancet*, 397(10272):398–408, January 2021. Publisher: Elsevier.
- [3] Felicity T. Cutts, Hector S. Izurieta, and Dale A. Rhoda. Measuring Coverage in MNCH: Design, Implementation, and Interpretation Challenges Associated with Tracking Vaccination Coverage Using Household Surveys. *PLOS Medicine*, 10(5):e1001404, May 2013. Publisher: Public Library of Science.
- [4] Felicity T. Cutts, Pierre Claquin, M. Carolina Danovaro-Holliday, and Dale A. Rhoda. Monitoring vaccination coverage: Defining the role of surveys. *Vaccine*, 34(35):4103–4109, July 2016.
- [5] Vaccine Funding Guidelines. page 65, 2021.
- [6] M. Carolina Danovaro-Holliday, Claudia Ortiz, Shea Cochi, and Cuauhtémoc Ruiz-Matus. Electronic immunization registries in Latin America: progress and lessons learned. *Revista Panamericana De Salud Publica = Pan American Journal of Public Health*, 35(5-6):453–457, 2014.
- [7] Xavier Bosch-Capblanch, Olivier Ronveaux, Vicki Doyle, Valerie Remedios, and Abdallah Bchir. Accuracy and quality of immunization information systems in forty-one low income countries. *Tropical medicine & international health: TM & IH*, 14(1):2–10, January 2009.
- [8] P. Bolton, E. Holt, A. Ross, N. Hughart, and B. Guyer. Estimating vaccination coverage using parental recall, vaccination cards, and medical records. *Public Health Reports (Washington, D.C.: 1974)*, 113(6):521–526, 1998.
- [9] Melody Miles, Tove K. Ryman, Vance Dietz, Elizabeth Zell, and Elizabeth T. Luman. Validity of vaccination cards and parental recall to estimate vaccination coverage: A systematic review of the literature. *Vaccine*, 31(12):1560–1568, March 2013.
- [10] Katy A. M. Gaythorpe, Kévin Jean, Laurence Cibrelus, and Tini Garske. Quantifying model evidence for yellow fever transmission routes in Africa. *PLOS Computational Biology*, 15(9):e1007355, September 2019.
- [11] Moritz U. G. Kraemer, Nuno R. Faria, Robert C. Reiner, Nick Golding, Birgit Nikolay, Stephanie Stasse, Michael A. Johansson, Henrik Salje, Ousmane Faye, G. R. William Wint, Matthias Niedrig, Freya M. Shearer, Sarah C. Hill, Robin N. Thompson, Donal Bisanzio, Nuno Taveira, Heinrich H. Nax, Bary S. R. Pradelski, Elaine O. Nsoesie, Nicholas R. Murphy, Isaac I. Bogoch, Kamran Khan, John S. Brownstein, Andrew J. Tatem, Tulio de Oliveira, David L. Smith, Amadou A. Sall, Oliver G. Pybus, Simon I. Hay, and Simon Cauchemez. Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study. *The Lancet Infectious Diseases*, 17(3):330–338, March 2017. Publisher: Elsevier.
- [12] Leocadia Kwagonza, Ben Masiira, Henry Kyobe-Bosa, Daniel Kadobera, Emily B. Atuheire, Bernard Lubwama, Atek Kagirita, Edson Katushabe, John T. Kayiwa, Julius J. Lutwama, Joseph C. Ojwang, Issa Makumbi, Alex Riolerus Ario, Jeff Borchert, and Bao-Ping Zhu. Outbreak of yellow fever in central and southwestern Uganda, February–may 2016. *BMC Infectious Diseases*, 18(1):548, November 2018.

- [13] World Health Organization. *A global strategy to eliminate yellow fever epidemics (EYE) 2017–2026*. World Health Organization, Geneva, 2018.
- [14] A. D. Barrett. Yellow fever vaccines. *Biologicals: Journal of the International Association of Biological Standardization*, 25(1):17–25, March 1997.
- [15] Christopher JL Murray, Bakhuti Shengelia, Neeru Gupta, Saba Moussavi, Ajay Tandon, and Michel Thieren. Validity of reported vaccination coverage in 45 countries. *The Lancet*, 362(9389):1022–1027, September 2003.
- [16] Jackson, Michael L and Rothman, Kenneth J. Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness, February 2015.
- [17] Neal D. Goldstein, Igor Burstyn, E. Claire Newbern, Loni P. Tabb, Jennifer Gutowski, and Seth L. Welles. Bayesian Correction of Misclassification of Pertussis in Vaccine Effectiveness Studies: How Much Does Underreporting Matter? *American Journal of Epidemiology*, 183(11):1063–1070, June 2016.
- [18] Erik K. Johnson, Rebecca Kahn, Yonatan H. Grad, Marc Lipsitch, and Daniel B. Larremore. Test negative designs with uncertainty, sensitivity, and specificity. preprint, *Epidemiology*, June 2021.
- [19] Paolo Eusebi, Niko Speybroeck, Sonja Hartnack, Jacob Stærk-Østergaard, Matthew J. Denwood, and Polychronis Kostoulas. Addressing misclassification bias in vaccine effectiveness studies with an application to Covid-19. *BMC Medical Research Methodology*, 23:55, February 2023.
- [20] Ulrike Baum, Sangita Kulathinal, and Kari Auranen. Exposure misclassification bias in the estimation of vaccine effectiveness. *PLoS ONE*, 16(5):e0251622, May 2021.
- [21] Tom De Smedt, Elizabeth Merrall, Denis Macina, Silvia Perez-Vilar, Nick Andrews, and Kaatje Bollaerts. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. *PLOS ONE*, 13(6):e0199180, June 2018. Publisher: Public Library of Science.
- [22] Tini Garske, Maria D. Van Kerkhove, Sergio Yactayo, Olivier Ronveaux, Rosamund F. Lewis, J. Erin Staples, William Perea, Neil M. Ferguson, and for the Yellow Fever Expert Committee. Yellow Fever in Africa: Estimating the Burden of Disease and Impact of Mass Vaccination from Outbreak and Serological Data. *PLoS Medicine*, 11(5):e1001638, May 2014. Publisher: Public Library of Science.
- [23] Arran Hamlet, Kévin Jean, Sergio Yactayo, Justus Benzler, Laurence Cibrelus, Neil Ferguson, and Tini Garske. POLICI: A web application for visualising and extracting yellow fever vaccination coverage in Africa. *Vaccine*, 37(11):1384–1388, March 2019.
- [24] Hugo Muench. *Catalytic Models in Epidemiology*. Harvard University Press, 1959. Google-Books-ID: rXIHA AAAIAAJ.
- [25] Katy AM Gaythorpe, Kaja Abbas, John Huber, Andromachi Karachaliou, Niket Thakkar, Kim Woodruff, Xiang Li, Susy Echeverria-Londono, VIMC Working Group on COVID-19 Impact on Vaccine Preventable Disease, Matthew Ferrari, Michael L Jackson, Kevin McCarthy, T Alex Perkins, Caroline Trotter, and Mark Jit. Impact of COVID-19-related disruptions to measles, meningococcal A, and yellow fever vaccination in 10 countries. *eLife*, 10:e67023, June 2021. Publisher: eLife Sciences Publications, Ltd.
- [26] T. Alex Perkins, John H. Huber, Quan M. Tran, Rachel J. Oidtmann, Magdalene K. Walters, Amir S. Siraj, and Sean M. Moore. Burden is in the eye of the beholder: Sensitivity of yellow fever disease burden estimates to modeling assumptions. *Science Advances*, 7(42):eabg5033, October 2021. Publisher: American Association for the Advancement of Science.
- [27] Kévin Jean, Christl A. Donnelly, Neil M. Ferguson, and Tini Garske. A Meta-Analysis of Serological Response Associated with Yellow Fever Vaccination. *The American Journal of Tropical Medicine and Hygiene*, 95(6):1435–1439, December 2016.

- [28] Lawrence Henry Ofosu-Appiah, Dodzi Kofi Amelor, Bright Ayensu, Ernest Akyereko, Nafisah Is-sah Rabiwu, David Opare, Godfred Owusu-Okyere, Dennis Odai Laryea, Franklin Asiedu-Bekoe, and Julius Abraham Addo Mingle. An evaluation of the diagnostic performance characteristics of the Yellow Fever IgM immunochromatographic rapid diagnostic test kit from SD Biosensor in Ghana. *PLOS ONE*, 17(1):e0262312, January 2022. Publisher: Public Library of Science.
- [29] Cristina Domingo, Rémi N. Charrel, Jonas Schmidt-Chanasit, Hervé Zeller, and Chantal Reusken. Yellow fever in the diagnostics laboratory. *Emerging Microbes & Infections*, 7(1):1–15, December 2018. Publisher: Taylor & Francis .eprint: <https://doi.org/10.1038/s41426-018-0128-8>.
- [30] Matthias Niedrig, Oliver Kürsteiner, Christian Herzog, and Karen Sonnenberg. Evaluation of an indirect immunofluorescence assay for detection of immunoglobulin M (IgM) and IgG antibodies against yellow fever virus. *Clinical and vaccine immunology: CVI*, 15(2):177–181, February 2008.
- [31] Stan Development Team. RStan: the R interface to Stan, 2022.
- [32] R Core Team. R: A Language and Environment for Statistical Computing, 2022.
- [33] World Health Organization. The immunization data quality self-assessment (DQS) tool, 2005.
- [34] World Health Organization. *Data quality review*. World Health Organization, Geneva, 2017.
- [35] Heather M. Scobie, Michael Edelstein, Edward Nicol, Ana Morice, Nargis Rahimi, Noni E. MacDonald, M. Carolina Danovaro-Holliday, and Jaleela Jawad. Improving the quality and use of immunization and surveillance data: Summary report of the Working Group of the Strategic Advisory Group of Experts on Immunization. *Vaccine*, 38(46):7183–7197, October 2020.
- [36] Yellow Fever (YF) reported cases and incidence.
- [37] Chigozie Edson Utazi, Andrew Tatem. Precise mapping reveals gaps in global measles vaccination coverage, December 2020.
- [38] K. M. O’Reilly, A. Cori, E. Durray, M. Z. Wadood, A. Bosan, R. B. Aylward, and N. C. Grassly. A New Method for Estimating the Coverage of Mass Vaccination Campaigns Against Poliomyelitis From Surveillance Data. *American Journal of Epidemiology*, 182(11):961–970, December 2015.