

Analytical approaches for antimalarial antibody responses to confirm historical and recent malaria transmission: an example from the Philippines

Maria Lourdes M. Macalinao,^{a,b,c} Kimberly M. Fornace,^{b,d} Ralph A. Reyes,^a Tom Hall,^b Alison Paolo N. Bareng,^a John H. Adams,^e Christèle Huon,^f Chetan E. Chitnis,^f Jennifer S. Luchavez,^a Kevin K. A. Tetteh,^b Katsuyuki Yui,^g Julius Clemence R. Hafalla,^b Fe Esperanza J. Espino,^a and Chris J. Drakeley^{b,*}



^aDepartment of Parasitology and National Reference Laboratory for Malaria and Other Parasites, Research Institute for Tropical Medicine, Department of Health, Muntinlupa City, Philippines

^bFaculty of Infectious and Tropical Diseases, Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, United Kingdom

^cSchool of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan

^dInstitute of Biodiversity, Animal Health & Comparative Medicine, University of Glasgow, Glasgow, United Kingdom

^eUniversity of South Florida, Tampa, FL 33612, USA

^fMalaria Parasite Biology and Vaccines Unit, Department of Parasites and Insect Vectors, Institut Pasteur, Paris, France

^gShionogi Global Infectious Diseases Division, Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan

Summary

Background Assessing the status of malaria transmission in endemic areas becomes increasingly challenging as countries approach elimination. Serology can provide robust estimates of malaria transmission intensities, and multiplex serological assays allow for simultaneous assessment of markers of recent and historical malaria exposure.

Methods Here, we evaluated different statistical and machine learning methods for analyzing multiplex malaria-specific antibody response data to classify recent and historical exposure to *Plasmodium falciparum* and *Plasmodium vivax*. To assess these methods, we utilized samples from a health-facility based survey (n = 9132) in the Philippines, where we quantified antibody responses against 8 *P. falciparum* and 6 *P. vivax*-specific antigens from 3 sites with varying transmission intensity.

Findings Measurements of antibody responses and seroprevalence were consistent with the 3 sites' known endemicity status. Among the models tested, a machine learning (ML) approach (Random Forest model) using 4 serological markers (PfGLURP R2, Etramp5.Ag1, GEXP18, and PfMSP1₁₉) gave better predictions for *P. falciparum* recent infection in Palawan (AUC: 0.9591, CI 0.9497–0.9684) than individual antigen seropositivity. Although the ML approach did not improve *P. vivax* infection predictions, ML classifications confirmed the absence of recent exposure to *P. falciparum* and *P. vivax* in both Occidental Mindoro and Bataan. For predicting historical *P. falciparum* and *P. vivax* transmission, seroprevalence and seroconversion rates based on cumulative exposure markers AMA1 and MSP1₁₉ showed reliable trends in the 3 sites.

Interpretation Our study emphasizes the utility of serological markers in predicting recent and historical exposure in a sub-national elimination setting, and also highlights the potential use of machine learning models using multiplex antibody responses to improve assessment of the malaria transmission status of countries aiming for elimination. This work also provides baseline antibody data for monitoring risk in malaria-endemic areas in the Philippines.

Funding Newton Fund, Philippine Council for Health Research and Development, UK Medical Research Council.

Copyright © 2023 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Malaria; Multiplex serology; Serosurveillance; Analytical approaches; Machine learning

*Corresponding author.

E-mail addresses: maloumacalinao@gmail.com (M.L.M. Macalinao), Kimberly.Fornace@glasgow.ac.uk (K.M. Fornace), rreyes.rmt@gmail.com (R. A. Reyes), thall@sgul.ac.uk (T. Hall), paolobareng@gmail.com (A.P.N. Bareng), ja2@usf.edu (J.H. Adams), christele.huon@pasteur.fr (C. Huon), chetan.chitnis@pasteur.fr (C.E. Chitnis), jluchavez@yahoo.com (J.S. Luchavez), Kevin.Tetteh@lshtm.ac.uk (K.K.A. Tetteh), katsu@nagasaki-u.ac.jp (K. Yui), Julius.Hafalla@lshtm.ac.uk (J.C.R. Hafalla), fe.espino2019@gmail.com (F.E.J. Espino), Chris.Drakeley@lshtm.ac.uk (C.J. Drakeley).

The Lancet Regional Health - Western Pacific 2023, #100792

Published Online XXX
<https://doi.org/10.1016/j.lanwpc.2023.100792>

Research in context

Evidence before this study

Serology has been shown to provide robust estimates of malaria transmission intensity in populations. Measurement of antibody levels against disease-specific and species-specific antigens through finger prick blood samples has the capacity for sensitive and high throughput monitoring that is further improved with the recent development of multiplex immunoassays. There is, however, a need to evaluate the methods of analysis and interpretation of multiplex serology data for use in classifying disease exposure profiles. We searched PubMed on 08 September 2022 using the terms ("malaria" OR "plasmodium") AND ("IgG" OR "antibody" OR "antigen" OR "markers") AND ("serolog*") AND (((("quantitative") OR ("multiplex") OR ("multivariate") OR ("model*")) AND ("analy*")), which returned 145 articles describing multiplex serology studies that not only focused on antimalarial antibodies, but also included seroprevalence studies for simultaneous assessment of multiple diseases, including SARS-CoV-2 and neglected tropical diseases. Although there have been numerous studies that utilized multiplex assays to assess serological exposure markers, most studies analyzed the markers individually, and only 3 studies explored the use of the quantitative data in multivariate statistical methods such as machine learning to see whether this method will improve classification or exposure predictions. Moreover, a PubMed search of malaria serology studies in the Philippines resulted in only 2 recent studies, both of which did not delve into evaluating serological markers of recent and historical malaria exposure against *Plasmodium falciparum* and *Plasmodium vivax*.

Added value of this study

Antibody data is usually analyzed individually through seropositivity cutoffs to produce binary outcomes, or estimation of seroconversion rates through reverse catalytic models fitted using maximum likelihood methods. Our study advances on previous work on serological markers with our assessment of the different analytical approaches for

interpreting multiplex antibody response data. By comparing the classical approaches to the more dynamic application of multiplex quantitative data in machine learning and ensemble learning, we are able to show that utilizing a combination of antibody measurements and available training data, we can potentially improve predictions for exposure to specific pathogens, may it be recent or historical. In our panel of *P. falciparum* markers, we identified a combination of 4 antigens that was able to accurately predict recent *P. falciparum* infection using a machine learning model. Our alternative approach in multiplex analysis was also able to confirm the absence of recent transmission of both *falciparum* and *vivax* malaria in our Philippine sites, which have not reported local cases in recent years. This is also the first study to describe the malaria exposure levels of areas of varying endemicity in the Philippines based on multiple *P. falciparum* and *P. vivax* serological markers, and can then provide the baseline antibody data for future malaria serosurveillance studies in the country.

Implications of all the available evidence

In this study we presented alternative approaches in analyzing multiplex serological data to make full use of the quantitative antimalarial antibody measurements for improved predictions of malaria exposure levels and transmission intensities. Although supervised classification algorithms would rely on the availability of training data, we have shown and validated that trained models can potentially be useful in areas with likely similar transmission patterns (e.g., Philippines and Malaysia), while its application in other settings needs to be further evaluated. Our study highlights the utility of quantitative antibody data from previously reported recent and historical malaria exposure markers in providing a sensitive assessment of the malaria situation of varied endemic settings. These recent advances in multiplex immunoassays allow for opportunities for an integrated disease surveillance.

Introduction

In malaria elimination areas, where reporting of local cases continually declines, assessing, and differentiating areas with residual transmission becomes increasingly challenging. In the case of the Philippines, which is aiming to eliminate malaria by 2030 following a sub-national elimination approach,¹ the country saw >70% decrease in malaria cases in the past decade, with only 2 of 81 provinces currently reporting local cases in 2021, and 19 in elimination phase. *P. falciparum* (Pf) contributes >80% to the total malaria cases, *P. vivax* (Pv) is at >20%, while the other species *Plasmodium malariae*, *Plasmodium ovale*, and *Plasmodium knowlesi* make up <1% of the cases.² Innovative tools that are capable of detecting both past and present infections can

potentially be used to confirm the presence or absence of malaria transmission in this kind of setting that employs subnational specific control approaches.³⁻⁶

Several studies have utilized serology and malaria-specific antibody responses to estimate malaria transmission intensities,^{7,8} showing that these represent a viable additional metric of both historical and recent exposure.⁹⁻¹⁴ Antibody prevalence alone and as age-adjusted seroconversion rates correlate with entomological and parasitological measures used in estimating malaria transmission.¹⁵⁻¹⁷ Many of the original studies examined single antigen platforms and antigens associated with cumulative exposure to infection such as Pf apical membrane antigen-1 (PfAMA1) and the 19 KDa fragment of Pf merozoite protein 1 (PfMSP1₁₉). More

recent advances in array and bead-based assay platforms allow for simultaneous analysis of antibody responses to multiple antigens.^{11,18–22} These approaches allow the inclusion of multiple targets that can represent diversity in the parasite and variation in individual immune response. Recent multi-antigen studies have identified markers associated with antibody responses describing recent and historical *P. falciparum* and *P. vivax* exposure.^{9,23–25} To fully realize the additional information provided by examining multiple antigenic targets, more advanced statistical approaches and algorithms such as machine learning can be employed to predict optimal combinations of antibody responses for the outcome of interest.

The overall aim of this study was to evaluate known malaria-specific *P. falciparum* and *P. vivax* serological markers for their predictive capacity to distinguish recent infections from historically exposed individuals. Specifically, we sought to evaluate different approaches to determining seropositivity for estimating malaria transmission intensities and exposure in areas of varying endemicity, and apply this using serological data from the Philippines. To achieve this, we: 1) evaluated analysis methods to determine seropositivity using malaria-specific antibody responses to single and multiple antigens, utilizing multiplex serological data from health facility surveys in 3 sites in the Philippines, and 2) analyzed antibody responses and estimated transmission intensities in relation to the supposed immune status of these populations.

Methods

Ethical approval

This study was reviewed and approved by the Research Institute for Tropical Medicine—Institutional Review Board in Manila, Philippines (RITM IRB 2016-04) and LSHTM Research Ethics Committee in London, United Kingdom (11597).

Data source: study sites and samples

The study was conducted in 3 municipalities in 3 provinces in the Philippines, representing areas of varying malaria endemicity: Rizal, Palawan, currently the most endemic municipality, and reported >60% of the total cases, with an annual parasite index (API) of 5.7 per 1000 risk population in 2018; Abra de Ilog, Occidental Mindoro, a municipality reporting sporadic local cases and with declining transmission (API of 0.38 in 2018); and Morong, Bataan with a last reported indigenous case in 2011 and declared malaria-free in 2019 (Fig. 1).

Participants were recruited from June 2016 to June 2018 in a health facility-based rolling cross-sectional survey detailed in Reyes et al. (2021). Briefly, all health facility attendees were invited as study participants, who provided a finger-pricked blood sample for malaria

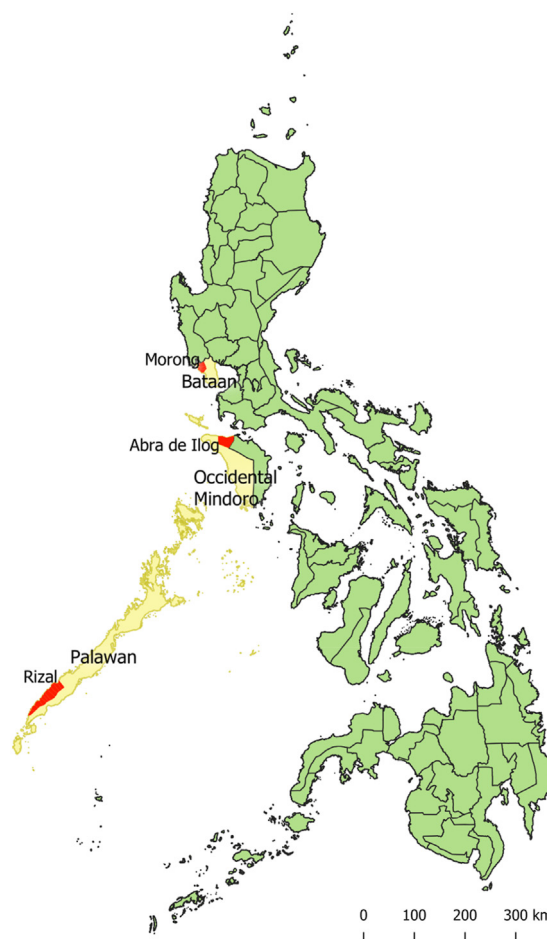


Fig. 1: Map showing the study sites, with red areas as the focused municipalities within the provinces marked yellow.

diagnosis through microscopy (blood smear), rapid diagnostic test (RDT), and a dried blood spot sample (DBS) on Whatman 3 MM CHR filter paper for malaria diagnosis by PCR and serological analysis. DBS samples were stored with desiccant at -20°C .

Multiplex bead-based assay of malaria-specific antibodies

Serological analysis was conducted using a multiplex bead-based assay as previously described,^{23,26,27} with an antigen panel that included 8 *P. falciparum*-specific and 6 *P. vivax*-specific recombinant antigens coupled to Magplex beads (Luminex Corp, Austin, TX, USA). The antigens were PfAMA1 (apical membrane antigen 1), PfMSP1₁₉ (merozoite surface protein), and their *P. vivax* homologues PvAMA1 and PvMSP1₁₉; PfGLURP R2 (glutamate rich protein), Etramp5.Ag1 (early transcribed membrane protein 5), PfSEA1 (schizont egress antigen), GEXP18 (Gametocyte exported protein 18), MSP2 CH150/9 (CH150/9 allele of MSP2), MSP 2 Dd2 (Dd2 allele of MSP2), PvEBP (erythrocyte binding protein),

PvRBP1a (reticulocyte binding protein 1a), PvRII and PvDBPRII (region II, Duffy binding protein). Antigen characteristics are shown in [Table 1](#).

The assay was conducted as described in Wu et al. (2019). Briefly, serum was eluted from DBS samples (~1uL from 3 mm punch; 1:400 final dilution). Approximately 1000 beads per antigen were added per well in 96-well flat bottom plates, which are incubated for 1.5 h on a shaker with 50uL sample and controls (2 positive controls for *P. falciparum* and *P. vivax*—pooled plasma from adults in hyperendemic malaria setting; 1 negative control—European malaria-naïve blood donors). After washing, samples are incubated for 1.5 h with 50uL 1:200 goat anti-human Fcγ-fragment-specific IgG (Jackson Immuno 109-116-098: conjugated to R-PE). Samples were washed, suspended in PBS, and read using a Luminex 200 or Magpix machine to measure IgG antibody levels reported as net median fluorescence intensity (MFI) values. During initial testing, strong correlation and comparable values were observed for samples assayed in both Luminex platforms. Plate-specific adjustments were performed based on the outcome of standard control curves generated from positive control pools included on each plate.

Data analysis

Statistical analyses were performed using R version 3.6.3 and Graphpad Prism 8. IgG antibody responses recorded as net MFI values were analyzed using different methods. Quantitative continuous antibody response data (reported as log₁₀ MFI values) were compared for different groups (*i.e.*, by age group, study site or recent malaria infection) using Student's t-test or one-way ANOVA Kruskal–Wallis test with Wilcoxon test for pairwise comparisons. Correlations between the

levels of antibody responses as well as age and malaria positivity were analyzed using Spearman's rank correlation.

Determining seropositivity rates using single antigen responses

Binary outcomes for seropositivity to each antigen were determined through the computation of cut-off values in 2 ways: 1) a 2-component finite mixture model (FMM); and 2) a reference negative population (NegPop) model. FMM was computed using the mixEM function in R mixtools package. For the NegPop model, European naïve samples were used as reference population. For both models, samples were considered seropositive for specific antigens if MFI values are higher than the antigen cutoff values (mean MFI plus 3 standard deviations). To evaluate the accuracy of these classifications, we used samples from known exposed and unexposed populations as validation data ([Table 2](#)), comprising *P. falciparum* (Pf+) and *P. vivax*-positive (Pv+) datasets, and a malaria-negative (neg) dataset collated from previous studies and this study. Sensitivity and specificity for identifying recent *P. falciparum* and *P. vivax* infection from this study and the validation dataset, and receiver operating characteristics (ROC) curves, were determined for each serological marker ([Supplementary Tables S1 and S2](#)).

Applying machine learning techniques for multiplex analysis of antigen responses

We next utilized machine learning (ML) techniques on multiplex serological data for determining recent and historical malaria exposure. Focusing on continuous covariates, the 8 Pf-specific antigens along with age were the covariates evaluated for *P. falciparum* exposure,

Antigen	Description	Gene ID	Reference
Pf-specific			
PfAMA1	apical membrane antigen 1	PF3D7_1133400	Collins et al., 2007 ²⁸
PfMSP1 ₁₉	19 kDa fragment of merozoite surface protein (MSP) 1	PF3D7_0930300	Burghaus and Holder, 1994 ²⁹
PfGlurp R2	glutamate rich protein R2	PF3D7_1035300	Theisen et al., 1995 ³⁰
Etramp5.Ag1	early transcribed membrane protein 5	PF3D7_0532100	van den Hoogen et al., 2019 ³¹ ; Spielmann et al., 2003 ³²
PfSEA1	schizont egress antigen	PF3D7_1021800	Raj et al., 2014 ³³
GEXP18	Gametocyte exported protein 18	PF3D7_0402400	Helb et al., 2015 ²³
MSP2 CH150	CH150/9 allele of MSP2	PF3D7_0206800	Polley et al., 2006 ³⁴
MSP2 Dd2	Dd2 allele of MSP2	PF3D7_0206800	Taylor et al., 1995 ³⁵
Pv-specific			
PvAMA1	apical membrane antigen 1	PVX_092275	Chuquiyauri et al., 2015 ³⁶
PvMSP1 ₁₉	merozoite surface protein	PVX_099980	França et al., 2016b ³⁷
PvRII	region II, Duffy binding protein	PVX_110810	França et al., 2016a ³⁸
PvDBP RII	region II, Duffy binding protein	PVX_110810	Ntumngia et al., 2012 ³⁹
PvEBP	erythrocyte binding protein	PVX_110835	Hester et al., 2013 ⁴⁰ ; Menard et al., 2013 ⁴¹
PvRBP 1a	reticulocyte binding protein 1a	PVX_098585	França et al., 2016a ³⁸

Table 1: List of antigens in the multiplex bead-based assay panel.

while for *P. vivax* exposure, only the 6 Pv-specific antigens were analyzed (age data not available). The training data used for these models are in Table 2. To account for the age-dependent cumulative increase in antibodies, the negative dataset for classifying recent infection included historical positives—survey participants aged >50 years from Malaysia⁴⁴ and Philippines⁴³ assumed to have had historical exposure, as well as a random selection of Pf or Pv-negatives from all age groups from this current study, all of which confirmed to be PCR-negative in addition to negative microscopy/RDT diagnosis. For classifying historical exposure, the training dataset included the historical positives as positives, and the same negatives from the validation dataset (Table 2).

The R package Super Learner (SL) optimized with AUC (Area under the ROC Curve) was used as an ensemble modelling algorithm to allow for evaluating multiple models simultaneously,^{23,42,45} namely Random Forest (RF), k-Nearest Neighbor (kNN), generalized boosted models (GBM), Support Vector Machine (SVM), and GLM with Lasso (glmnet). Feature selection with corP, that screens for univariate correlation, was also included for some component models (GBM, RF).

The SL model gives a prediction value that ranges from 0 to 1, and this can be used to obtain binary classification (*i.e.*, those with prediction values higher than 0.5 were considered positive). A 20-fold cross-validation was performed for internal validation to evaluate the performance of the fitted SL models using a withheld training dataset, and the whole training dataset. ROC curves were used to evaluate the outcome of predicted values for the samples and were then compared to single antigen performances. The performance of each ML model, or base learner, in the SL ensemble was also assessed. For additional assessment, external validation was performed for the classification models through an independently collated dataset,⁴⁶ which used the same panel of *P. falciparum* antigens (not available for *P. vivax* panel).

Evaluating model performance in identifying recent Pf and Pv infections in this study

Using this study’s malaria-positive samples, we assessed the capability of the serological markers in identifying falciparum and vivax infections (Supplementary Table S1). The single antigen performances were

Training data used	N	Validation dataset ^a	Negative population model ^b	SuperLearner/ Random Forest:	SuperLearner/ Random Forest:	SuperLearner/ Random Forest:	Random Forest	Random Forest
				Recent Pf exposure	Recent Pv exposure	Historical Pf exposure	Historical Pf exposure	Historical Pv exposure
European naive controls	179	neg	-	-	-	-	-	-
Positive controls								
African <i>P. falciparum</i> -positive	4	Pf+	+					
African <i>P. vivax</i> -positive	10	Pv+		+				
This study								
<i>P. falciparum</i> -positive subset	568	Pf+	+					
<i>P. vivax</i> microscopy-positive subset	46	Pv+			+			
Below 10 y.o. from Bataan	202	neg	-	-	-	-	-	-
Historical positives aged ≥50 y.o. from Occidental Mindoro and Palawan	711		-			+		
PfAMA1 and PfMSP1 ₁₉ seropositive subset	512						+	
PvAMA1 and PvMSP1 ₁₉ seropositive subset	324							+
Randomly selected malaria-negatives from Palawan aged <50 y.o.	550		-	-				
Bataan study⁴²								
Below 10 y.o. from Bataan	73	neg	-	-	-	-	-	-
Malaysia study⁴³								
<i>P. falciparum</i> -positive subset	17	Pf+	+					
<i>P. vivax</i> -positive subset	37	Pv+			+			
Historical positives aged ≥50 y.o. from Malaysia 2	1581					+		
PfAMA1 and PfMSP1 ₁₉ seropositive subset	1479		-				+	
PvAMA1 and PvMSP1 ₁₉ seropositive subset	873		-					+

^aFinite mixture models (FMM) were derived from the entire ENSURE dataset (n = 9132), and indicated validation dataset was used to determine receiver operating characteristics (ROC) curves for both FMM and Negative population models. Seropositivity classifications from single antigens presented in Supplementary Tables S1 and S2 were assessed using combined Pf+ and neg datasets, and Pv+ and neg datasets, respectively. ^bFor the supervised classification models (negative population, Super Learner, and Random Forest models), + means it was classified as positive for the model, and - means it is considered for negative classification. Abbreviations: Pf- *P. falciparum*, Pv- *P. vivax*, neg - malaria-negative; y.o.- years old.

Table 2: Training and validation data used in the classification models.

compared with SL predictions, where models included 3 to 9 covariates combined based on the variable importance outcomes. The AUC values from the ROC analysis based on training data and ENSURE malaria-positives (test data) were then computed. With Super Learner giving weights to multiple ML models simultaneously, we also evaluated the individual ML models that were given the most weight or importance, namely RF and GBM.

Estimating seroconversion rates and historical exposure

Seropositivity of the cumulative exposure markers (AMA1, MSP1₁₉ for both *P. falciparum* and *P. vivax*) were used to estimate seroconversion rates (SCR) and seroreversion rates (SRR), by fitting the age-specific prevalence for each study site into reverse catalytic models using likelihood ratio methods, as previously described.^{3,47} The predicted time of change in transmission is analyzed. In addition to the AMA1 and MSP1₁₉ seroprevalence based on FMM models, RF classifications using the combined AMA1 and MSP1₁₉ data for both *P. falciparum* (Pf-RF:2-covar) and *P. vivax* (Pv-RF:2-covar) were also used to generate seroprevalence curves. The positive training data used for these RF models are subsets of the historical positives that were seropositive to both markers for each species based on FMM cutoff values (Table 2). The Super Learner approach was also used to predict Pf and Pv historical exposure using multiplex antibody data.

Role of the funding source

The funders of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. MLM, CD, and KF had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Heterogeneity of antibody responses from the 3 collection sites

A total of 9132 DBS samples (6572 for Palawan, 1683 for Occidental Mindoro, and 877 for Bataan) were available for serological evaluation. Females comprised >60% of survey participants, while all age groups were well-represented (Table 3). Malaria infections were detected only in Palawan, with 51.8% and 48.2% of infected males and females, respectively, aged under 10 years old. Of the 889 *Plasmodium*-positive samples confirmed through either microscopy, RDT and/or PCR, 58.0% had *P. falciparum*, 12.4% had *P. vivax*, 6.7% had mixed Pf + Pv infections, 6.1% had *P. malariae*, *P. ovale*, or *P. knowlesi*, while species identification could not be confirmed for 16.8% of the PCR-positive samples. From the Pf and Pv malaria-positive individuals (n = 707, including mixed infections), 388 (54.8%) were

symptomatic (had fever before or during the health facility visit), 62.1% of which were ≤10 years old. Conversely, 63.1% of the asymptomatic malaria infections were >10 years old.

Firstly, we compared the magnitude of antibody responses to *P. falciparum* and *P. vivax* antigens from the 3 collection sites. Palawan consistently had the highest antibody levels to all antigens and in all age groups, followed by individuals from Occidental Mindoro then Bataan (Fig. 2a and Supplementary Fig. S1). Expectedly, cumulative exposure markers PfAMA1, PfMSP1₁₉, PfGLURP R2, PvAMA1, and PvMSP1₁₉ had moderate associations with age (Spearman's coefficient ≥0.47, p < 0.0001), while recent exposure markers Etramp5.Ag1 and GEXP18 were less moderately associated (Spearman's coefficient of 0.39 and 0.42, respectively, p < 0.0001). PvMSP1₁₉ and PvDBPRII showed weak correlation (Spearman's coefficient of 0.13 and 0.12, respectively, p < 0.0001) with recent *P. vivax* infection (Fig. 2b). Similarly, mean antibody levels were higher in a species-specific manner (Fig. 3). Taken together, our results confirm the applicability of these serological markers in differentiating areas of varying malaria endemicity, and there were species-specific associations with recent infection.

Identifying markers of recent falciparum and vivax malaria exposure and recent infection

Next, we applied different analytical approaches to ascertain whether the serological markers can be used to predict recent or historical exposure. The resulting seropositivity cutoff values were comparable for FMM (unsupervised) and Negpop (supervised) classification models, with AUC values from the validation data ranging from 0.812 to 0.932 for the 8 *P. falciparum* markers, and 0.534 to 0.943 for the 6 *P. vivax* markers (Fig. 4a and b and Supplementary Tables S1 and S2). When applied to classifying study samples, some of the markers showed low predictive ability, with AUC values ranging from 0.497 to 0.756 for *P. falciparum* markers, and 0.513 to 0.731 for *P. vivax* markers. Etramp5.Ag1 and GEXP18 for *P. falciparum*, and PvMSP1₁₉ for *P. vivax*, had significantly higher AUC values (>0.735, p < 0.002) compared to the rest of the markers analyzed (Supplementary Tables S1 and S2). Density plots of antibody responses also show overlap of malaria-negative and positive distributions, contributing to the lower AUCs (Supplementary Fig. S2). Seroprevalence rates for all antigens were consistently low in Bataan in all age groups (Fig. 4a and b), with recent exposure markers (hollow circles) estimating lower rates.

We next applied ML methods for analyzing multiplex antibody response data. The Super Learner (SL) ensemble machine learning algorithm was used to simultaneously evaluate classifications from ML models, which included Random Forest (RF), k-Nearest Neighbor (kNN), generalized boosted models (GBM),

	Morong, Bataan (Malaria- eliminated)		Abra de Ilog, Occidental Mindoro (Near elimination)		Rizal, Palawan (Ongoing transmission)	
Number of analyzed samples	877		1683		6572	
Female, %	628	(71.6%)	1101	(65.4%)	3734	(56.8%)
Age, median (IQR)	26	(11–39)	28	(15–43)	13	(5–31)
Number (% in site)						
0–5	150	(17.1%)	150	(8.9%)	1900	(28.9%)
6–10	63	(7.2%)	154	(9.2%)	1055	(16.1%)
11–25	215	(24.5%)	445	(26.4%)	1488	(22.6%)
26 & up	449	(51.2%)	934	(55.5%)	2129	(32.4%)
<i>Plasmodium</i> -positive by microscopy/RDT/PCR (% in site) ^a	0	(0.0%)	0	(0.0%)	889	(13.5%)
<i>P. falciparum</i> -positive (by microscopy/RDT/PCR) ^a	0	(0.0%)	0	(0.0%)	595	(9.1%)
<i>P. vivax</i> -positive (by microscopy/RDT/PCR) ^a	0	(0.0%)	0	(0.0%)	172	(2.6%)
PCR-confirmed species ID, n (% in site) ^b						
<i>P. falciparum</i> , Pf mono-infection	0	(0.0%)	0	(0.0%)	516	(7.9%)
<i>P. vivax</i> , Pv mono-infection	0	(0.0%)	0	(0.0%)	110	(1.7%)
Pf + Pv mixed infection	0	(0.0%)	0	(0.0%)	60	(0.9%)
Other species (mono- or mixed infections)	0	(0.0%)	0	(0.0%)	55	(0.8%)
With fever or history of fever (% in site) ^c	69	(7.9%)	172	(10.2%)	2727	(41.5%)
Symptomatic Pf and Pv infections ^{a,c}	0	(0.0%)	0	(0.0%)	388	(5.9%)
Asymptomatic infections (no fever, <i>Plasmodium</i> -positive) ^{a,c}	0	(0.0%)	0	(0.0%)	428	(6.5%)
Asymptomatic microscopy-positive Pf infections	0	(0.0%)	0	(0.0%)	132	(2.0%)
Asymptomatic microscopy-positive Pv infections	0	(0.0%)	0	(0.0%)	35	(0.5%)

^aNumbers include all *Plasmodium* infections detected by microscopy, RDT and/or PCR. Some samples did not have enough material for further species identification by PCR, for which results were based on microscopy and/or RDT diagnosis, while some (n = 149) were confirmed *Plasmodium*-positive only. Pf and Pv numbers reported include mixed infections, which were specified in the PCR-confirmed species ID breakdown. ^bAll 5 *Plasmodium* species, including *P. malariae* (n = 50), *P. ovale* (n = 4) and *P. knowlesi* (n = 1), were detected in Palawan, with some as co-infections with Pf and/or Pv. ^cPresence of fever and/or history of fever (within the past 3 days of health facility visit) in malaria-positive participants was used to classify symptomatic infections.

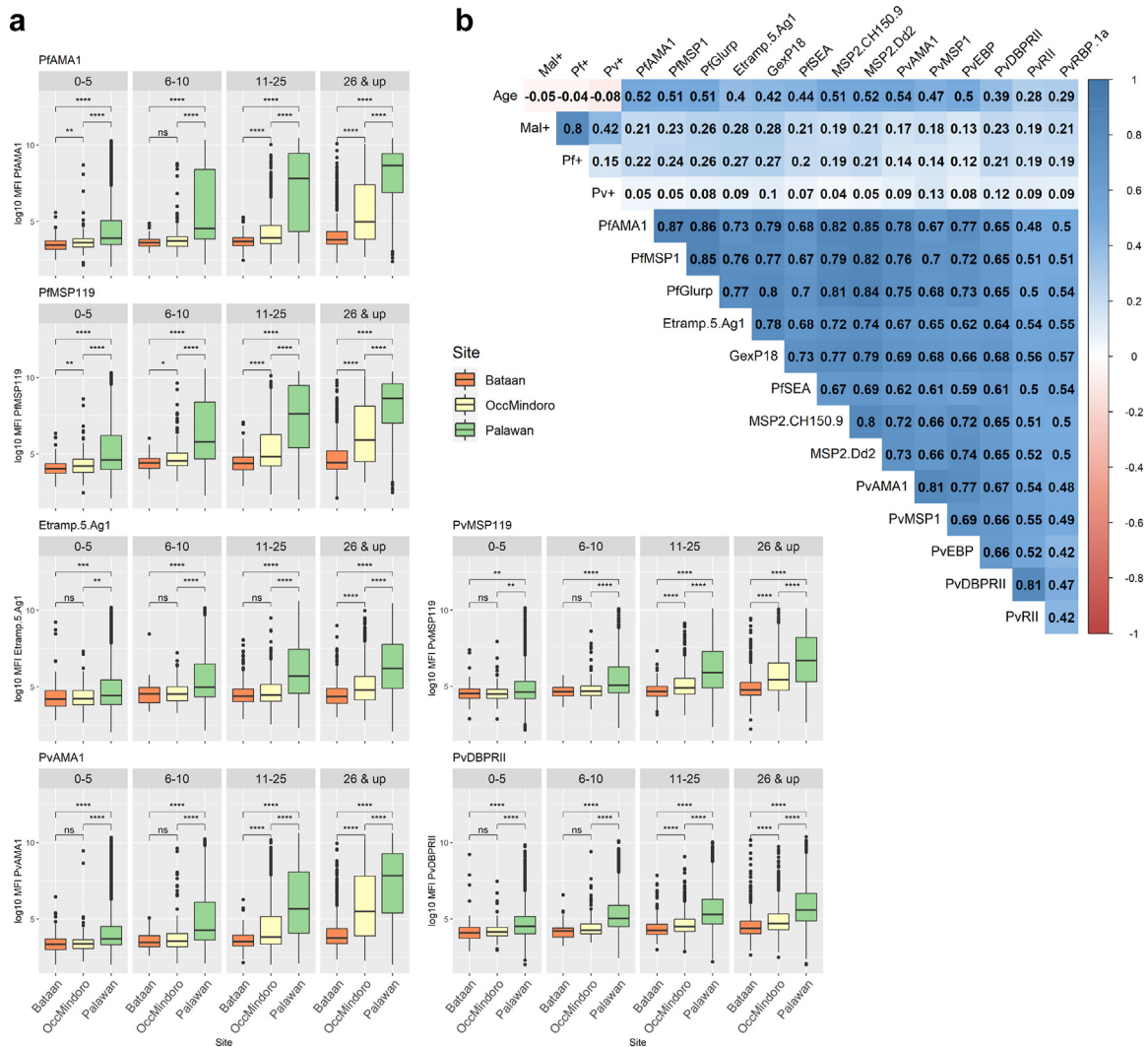
Table 3: Characteristics of study population by site.

Support Vector Machine (SVM), and GLM with Lasso (glmnet), and weights are applied to each learner after cross-validation. Covariates were assessed for their relative importance for the model, and different combinations in a model were assessed using the ROC curves of cross-validated predictions (Fig. 4c and d and Supplementary Tables S1 and S2). The 9-covariate model (SL: 9-covar; 8 antigens and age as covariates) had the highest AUC among all SL models, with a value of 0.9898 (95% CI 0.9874–0.9921) and 0.9197 (95% CI 0.9093–0.9302) for the binary classification predictions for the training data and test data, respectively, and removing the age as covariate (SL:8-covar model) decreased the prediction capacity (Supplementary Table S1). For predicting *P. vivax* recent exposure, the model incorporating the 6 Pv antigens as covariates (SL:6-covar) had a resulting AUC of 0.8857 (95% CI 0.8429–0.9284) and 0.6332 (0.5979–0.6686) for validation data and test data, respectively (Supplementary Table S2).

After a 20-fold nested cross-validation, generalized boosted models with or without feature selection (GBM and GBM_corP) were given the highest weights in the SL ensemble models for both species' recent exposure predictions (Supplementary Fig. S3a and 3b). However,

upon closer inspection of the classification accuracy of the individual learners for predicting *P. falciparum* recent infections, the Random Forest (RF) models were found to have the highest AUCs in both the validation data (not shown) and test data (Supplementary Fig. S3c). Computing for the AUCs of resulting binary classifications (Fig. 4c) also showed RF models with highly improved AUCs compared to FMM-based single antigen classifications and even the SL models, suggesting that RF is the best performing model in distinguishing recent falciparum malaria infections. On the other hand, vivax infection classification models had comparable AUCs ranging from 0.701 to 0.8 for test data (Supplementary Fig. S3d), and also for binary classifications for SL, RF, and GBM models (Fig. 4d and Supplementary Table S2).

Variable importance in the RF model for *P. falciparum* positivity (Fig. 4e) shows that responses to Etramp5.Ag1, PfGLURP R2, GEXP18 and PfMSP1₁₉ had the highest influence in the predictions, while PvMSP1₁₉ and PvAMA1 were the most predictive *P. vivax* antigens (Fig. 4f). RF.covar4 gave an AUC of 0.9983 for the training data and 0.9591 for the test data (Fig. 4c and Supplementary Table S1), and had comparable AUC with the RF 8-covariate model (RF.covar8),



showing that analysis using the RF model combining just the 4 predictive Pf antigens was sufficient for predicting recent *P. falciparum* infection or exposure.

The predictive performance of the resulting SL models were further compared with RF (both species) and GBM (for Pv) predictions (Fig. 4g and h). The rates of predicted recent *P. falciparum* infection based on SL and RF models all consistently gave 0%–0.2% prediction for Occidental Mindoro and Bataan, where no *P. falciparum* malaria cases were reported in recent years, thus accurately predicting the absence of recent falciparum malaria exposure and infection (Fig. 4g). SL

and RF models from *P. vivax* antigens also had <1% prediction rate for Bataan and <5% for Occidental Mindoro, also indicative of a predicted absence of recent *P. vivax* infections in these sites (Fig. 4h and Supplementary Table S2).

Model validation using external datasets

To further evaluate the performance of the RF:4-covar model for predicting *P. falciparum* recent exposure, external validation was performed using a collated test dataset (Supplementary Table S3) that included samples from a cross-sectional survey in Malaysia⁴⁸ (n = 8163; all

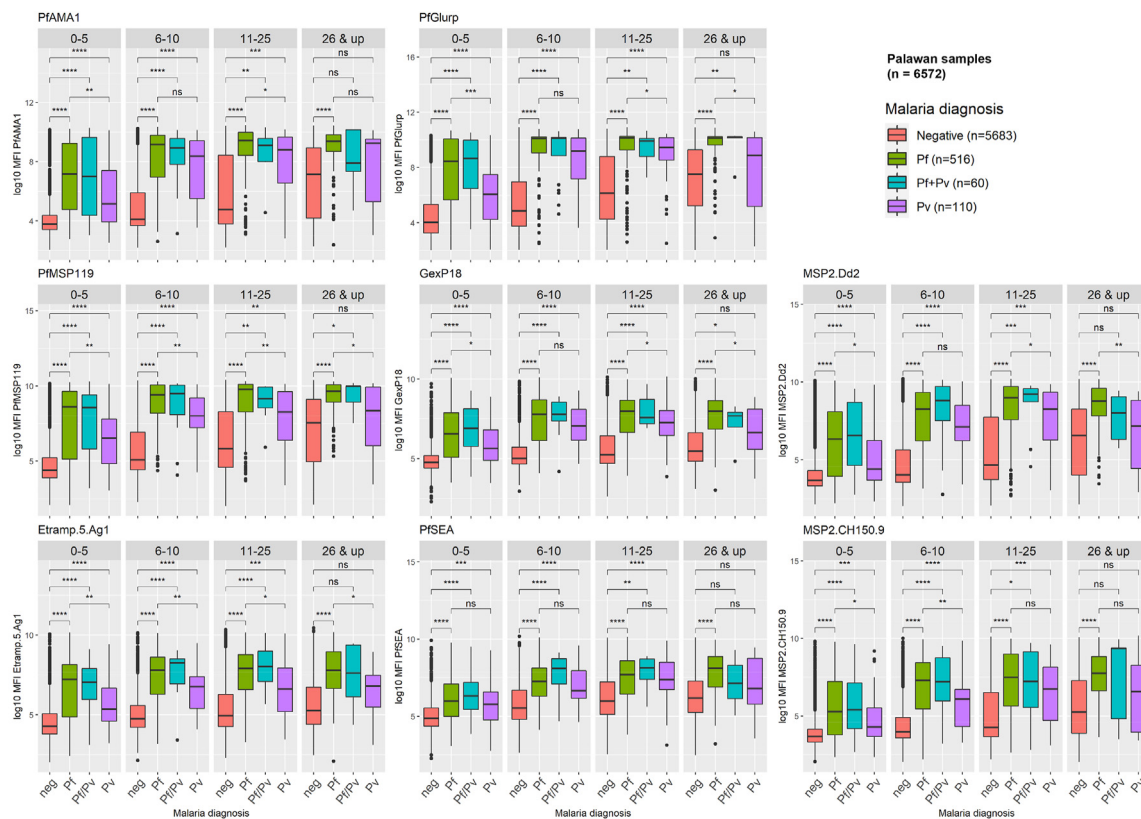
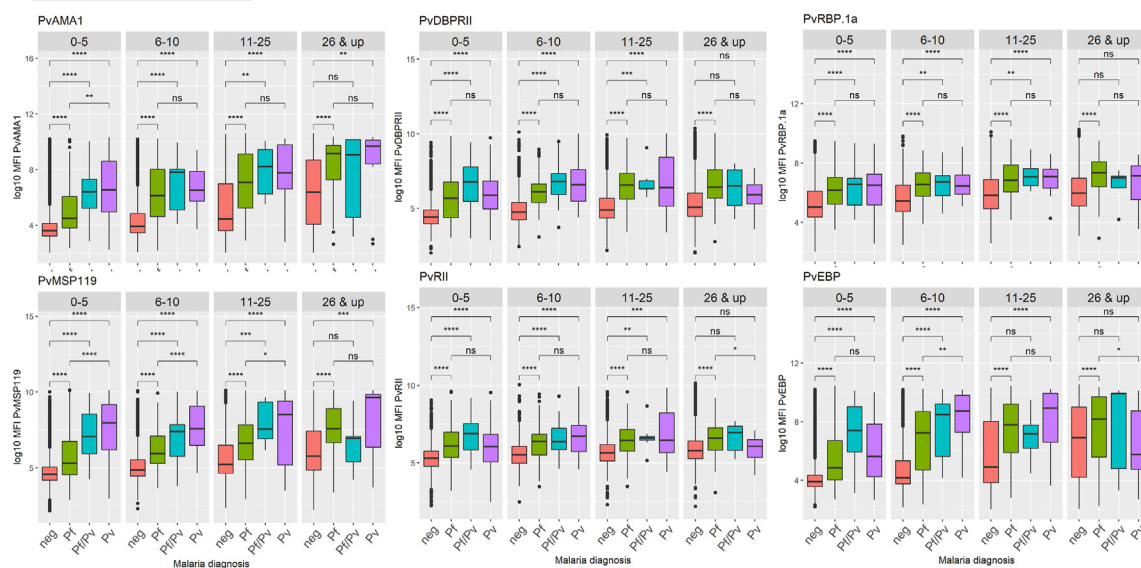
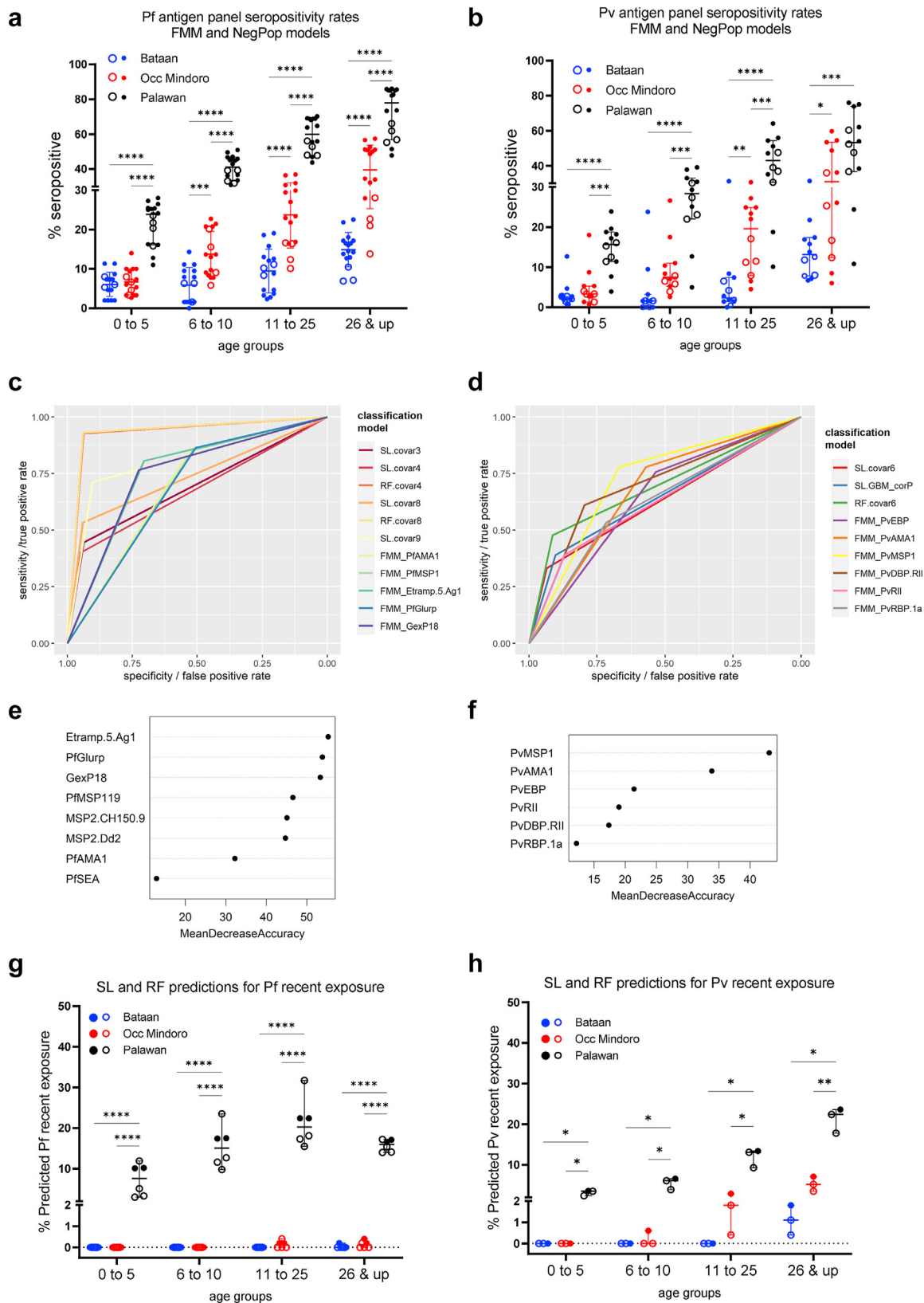
a *P. falciparum*-specific antigens**b** *P. vivax*-specific antigens

Fig. 3: Serological markers exhibit species-specific association with recent infection in varying levels. **a** and **b**) Comparison of antibody titers of Palawan samples (n = 6572) for *Plasmodium*-positive (Pf: *P. falciparum*-positive, Pf/Pv: Pf and Pv mixed infection, Pv: *P. vivax*-positive) and negative (neg) samples by age group in each *P. falciparum*-specific and *P. vivax*-specific antigens.



Pf-negative), a 2017 cross-sectional survey in Bataan⁴³ (n = 1926; all malaria-negative), PCR-validated malaria cases (Pf = 47, non-Pf = 428), and malaria-naïve or malaria-negative samples (n = 506). The model prediction resulted in an AUC of 0.93 (CI 0.8817–0.9782), and was able to predict 0.47% prevalence of infection for the Malaysia cross-sectional survey, and 0.10% for the Bataan cross-sectional survey. It also correctly identified 41 of 47 (87.2%) of the *P. falciparum* PCR-confirmed cases from the malaria-positive subset.

Estimating *P. falciparum* and *P. vivax* historical exposure

AMA1 and MSP1₁₉ responses for *P. falciparum* and *P. vivax* have been widely used to assess historical exposure by estimating seroconversion and seroreversion rates. Using reverse catalytic models and maximum likelihood tests on the age-specific seroprevalence based on the FMM model seropositivity rates for PfAMA1, PfMSP1₁₉, PvAMA1, and PvMSP1₁₉, we estimated the time of interruption of transmission in our study area (Fig. 5 and Table 4). The RF algorithm was also employed to generate models that combine both antigens (2-covariate models Pf-RF:2-covar and Pv-RF:2-covar), and the predicted binary outcomes were also used in the reverse catalytic models. Most of the seroconversion curves were best fit with a model assuming a change in transmission based on log likelihood tests ($p < 0.001$), except for PvMSP1₁₉ ($p = 0.507$) and the Pv-RF:2-covar ($p = 0.285$) model (Table 4). Based on these models, the time of change in transmission in Bataan was estimated at 29–33 years for *P. falciparum* and 28–29 years for *P. vivax*. Occidental Mindoro had varied estimates of the change point (7 and 12 years). For these 2 provinces, the seroconversion rates suggest a decrease in transmission, while an increase of the seroconversion rates for Palawan after the 2-year change point, suggested an increase in transmission in this study site.

We further explored whether the use of SL model can provide improved estimates of historical exposure.

The exposure prediction values from the SL model (8 *P. falciparum* markers) were graphed for each sample by site and age (Fig. 6a and d), and compared with seropositivity rates of PfAMA1 (Fig. 6c), which is shown to have the highest influence in the RF model prediction (Fig. 6b). However, comparing the SL-predicted rates with AMA1 and MSP1₁₉ seropositivity rates (Fig. 6d vs 6e) show that SL overestimated historical exposure, with higher-than-expected positivity rates in younger age groups from Bataan and Occidental Mindoro, where transmission is reportedly absent. The seropositivity rates based on the 2-covariate RF models with AMA1 and MSP1₁₉ seem to provide better estimates of historical exposure to both *P. falciparum* and *P. vivax* based on known malaria prevalence of the sites than the SL models (Fig. 6e and f).

Discussion

In this study we assessed the utility of multiple antigen-specific antibody responses and different statistical methods to estimate both recent malaria infection and historical transmission in areas of different endemicity in the Philippines. Using a multiplex bead-based assay with a panel of 8 *P. falciparum*-specific and 6 *P. vivax*-specific recombinant antigens on DBS samples, we showed how the use of combined antibody data can improve predictions of recent or historical malaria exposure. In particular, machine learning model predictions of malaria exposure were consistent with reported levels of transmission in our sites, which highlight the potential for advanced analysis of multiplex serological data to provide accurate complementary data on incidence levels that could be used by control programs at small spatial scales.

In many countries including the Philippines, malaria-endemic provinces apply for malaria-free status when indigenous cases have not been reported for a set number of years, typically 3 or more. More rapid, sub-national demonstration of the absence of exposure could

Fig. 4: Analysis of serological markers through machine learning methods improves classifications for recent *P. falciparum* and *P. vivax* exposure or recent infection. a and b Seropositivity rates of sample population by site and age group based on cutoff values from finite mixture models (FMM) and negative population model (NegPop) for Pf and Pv antigens (detailed in Supplementary Tables S1 and S2). Hollow dots represent the FMM and Negpop seropositivity rates for the two reported recent exposure markers Etramp5.Ag1 and GEXP18 for Pf (a), and PvMSP1₁₉ and PvDBP.RII for Pv (b) panel. Lines with error bars represent median with 95% CI. **c and d** Receiver operating characteristic (ROC) curves for the antibody responses to single antigens for individual antigens, and for the SL models (shown in both as binary outcomes of seropositivity/prediction values). **e and f** Variable importance of the 8 Pf-specific (e) and 6 Pv-specific (f) antigens based on the Random Forest model. **g** Predicted rates of recent Pf exposure based on analysis of the continuous antibody responses of the 8-antigen panel using Super Learner. Each hollow dot represents differences in the number of covariates used for the model (3, 4, 8, 9), as well as 2 showing rates from prediction values of the Random Forests (RF) component model (RF.covar4 and RF.covar8 in solid dots). **h** Predicted rates of recent Pv exposure based on analysis of the continuous antibody responses of the 6-antigen panel using Super Learner. Dots represent the positivity rates from the prediction values of the SL model, and the individual predictions from the 2 most weighted base learners in the resulting model—RF and GBM. RF predictions (RF.covar6) are shown as solid dots (SL: super learner, RF: random Forest, covar#: number of covariates included in the model, FMM: finite mixture models; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ with significance assessed by one-way ANOVA followed by Tukey's multiple comparison).

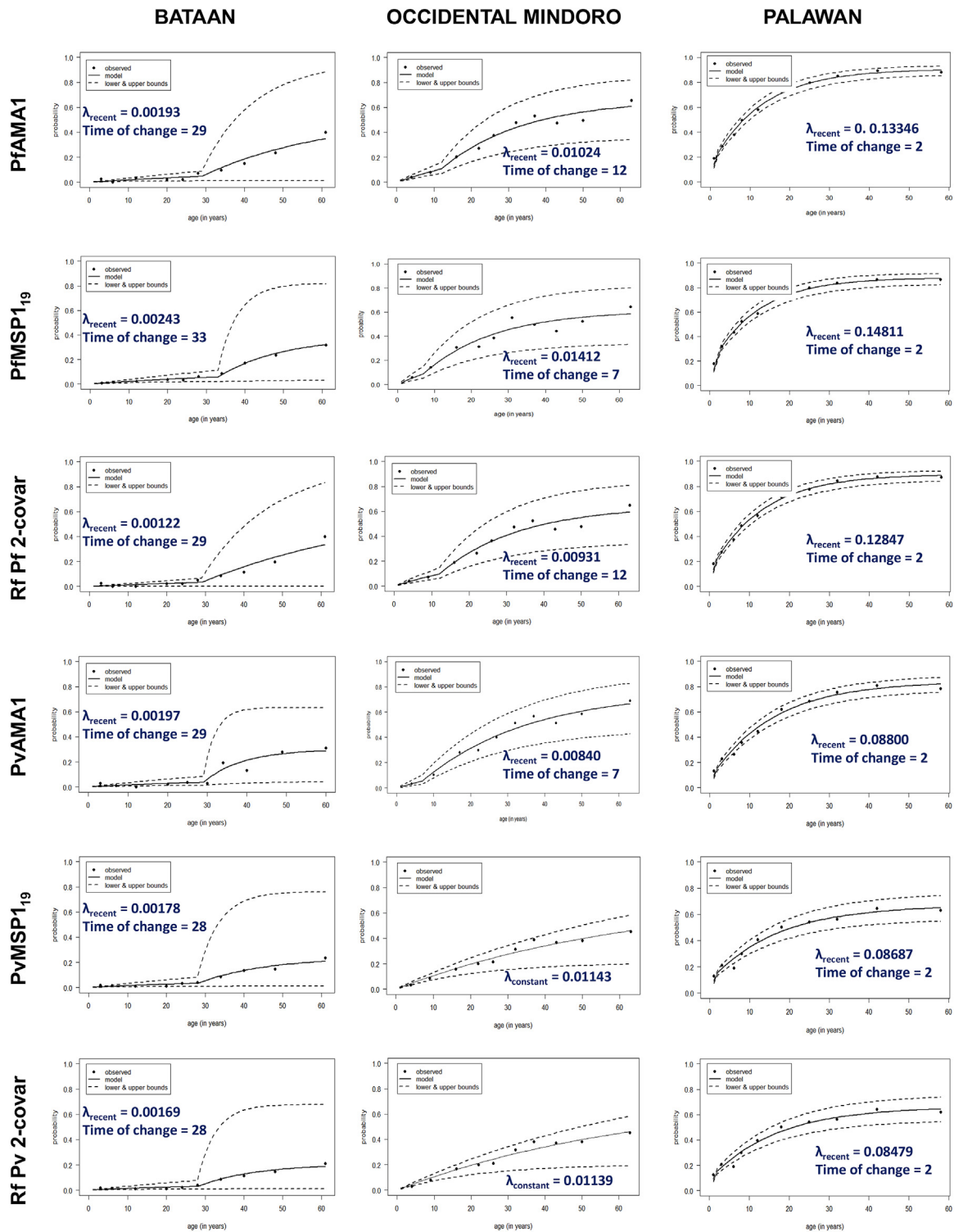


Fig. 5: Seroconversion curves based on reverse catalytic models using AMA1 and MSP1₁₉ antibody responses provide accurate estimates of historical exposure. Age-specific seroprevalence was based on finite mixture models and Random Forest models (using both antigens: RF 2-cover models) for each species. Solid lines represent the fit of the reversible catalytic models, dashed lines represent 95% confidence intervals,

		Bataan	Occidental Mindoro	Palawan
PfAMA1	SCR _{historical}	0.02102 (0.00599–0.07375)	0.03306 (0.02150–0.05082)	0.06953 (0.06251–0.07736)
	SCR _{recent}	0.00193 (0.00112–0.00332)	0.01024 (0.00707–0.01483)	0.13346 (0.11802–0.15092)
	SRR	0.010002 (0.0010–0.10353)	0.01239 (0.00575–0.02673)	0.00841 (0.00626–0.01129)
	Time of change in transmission (yrs)	29	12	2
PfmSP1 ₁₉	SCR _{historical}	0.04042 (0.00836–0.19516)	0.03399 (0.02221–0.05204)	0.07521 (0.06658–0.08495)
	SCR _{recent}	0.00243 (0.00148–0.00396)	0.01412 (0.00842–0.02367)	0.14811 (0.13109–0.16734)
	SRR	0.01959 (0.00564–0.06796)	0.01779 (0.00926–0.03416)	0.01221 (0.00946–0.01575)
	Time of change in transmission (yrs)	33	7	2
RF with PfAMA1 and PfmSP1 ₁₉	SCR _{historical}	0.01513 (0.00426–0.05370)	0.03338 (0.02172–0.05130)	0.06836 (0.06133–0.07619)
	SCR _{recent}	0.00122 (0.00064–0.00234)	0.00931 (0.00630–0.01377)	0.12847 (0.11333–0.14564)
	SRR	0.00463 (0.00002–0.91264)	0.01327 (0.00639–0.02755)	0.00901 (0.00674–0.01204)
	Time of change in transmission (yrs)	29	12	2
PvAMA1	SCR _{historical}	0.07078 (0.01619–0.30943)	0.02760 (0.02094–0.03638)	0.05210 (0.04652–0.05834)
	SCR _{recent}	0.00197 (0.00102–0.00377)	0.00840 (0.00451–0.01564)	0.08800 (0.07591–0.10202)
	SRR	0.03138 (0.01524–0.06460)	0.00715 (0.00269–0.01898)	0.01074 (0.00801–0.01441)
	Time of change in transmission (yrs)	29	7	2
PvMSP1 ₁₉	SCR _{historical}	0.02858 (0.00465–0.17570)	0.01143 (0.00921–0.01419)	0.03760 (0.03172–0.04458)
	SCR _{recent}	0.00178 (0.00092–0.00347)	–	0.08687 (0.07507–0.10052)
	SRR	0.02941 (0.00862–0.10026)	0.00404 (0.00047–0.03486)	0.02146 (0.01599–0.02880)
	Time of change in transmission (yrs)	28	–	2
RF with PvAMA1 and PvMSP1 ₁₉	SCR _{historical}	0.03387 (0.00559–0.20521)	0.01139 (0.00919–0.01414)	0.04023 (0.03472–0.04661)
	SCR _{recent}	0.00169 (0.00084–0.00340)	–	0.08479 (0.07313–0.09833)
	SRR	0.03498 (0.01265–0.09673)	0.00388 (0.00042–0.03598)	0.02227 (0.01674–0.02963)
	Time of change in transmission (yrs)	28	–	2

Seroconversion rates (SCR) and seroreversion rates (SRR) are presented for *P. falciparum* and *P. vivax* based on reverse catalytic models using age-specific seroprevalence from finite mixture models with AMA1 and MSP1₁₉ and Random Forest (RF) 2-covariate models. If the best reverse catalytic model to fit the age-adjusted seropositivity data was one with no change point in transmission, only 1 SCR is indicated, while if the best fit is one with a change point, the SCR before the change point is indicated as SCR_{historical} and the SCR after the change point is indicated as SCR_{recent}. The estimated change point in transmission (in years) is also indicated.

Table 4: Seroconversion and seroreversion rates from reverse catalytic models in each study site.

aid in this process. The use of multiple antigens to assess exposure to malaria infection circumvents some issues related to genetic diversity in the parasite and variation in the human immune response to different antigenic targets. The approach also allows application of more advanced statistical analysis to examine optimal combinations of antibody responses in predicting specific outcomes. Here, in addition to the well-studied antigens AMA1 and MSP1₁₉, we screened other antigenic markers that have been reported to accurately predict recent exposure in studies in Africa and the Caribbean.^{14,23,49} PfGLURP R2, Etramp5.Ag1 and GEXP18 have been identified as recent exposure markers for *P. falciparum*^{23,49} and PvMSP1₁₉ for *P. vivax*.²⁵ Our RF variable importance results also identified these markers to be the most predictive in our panel.

In assessing the individual performances of the serological markers, we show that the antibody responses to all antigens were species-specific, consistent with previous observations.^{50,51} Although we

observed differences in antibody levels between genders (Supplementary Figs. S4 and S5) and between patients and companions (Supplementary Fig. S6), particularly for Palawan, these factors had little influence overall on recent and historical exposure prediction rates based on the ML model. For some recent exposure markers, we observed <10% seropositivity in younger age groups from malaria-free Bataan, suggesting background seropositivity. When we employed machine learning,^{23,25,52} our models were able to confirm the absence of recent infection in Occidental Mindoro and Bataan (although there was lower precision for *P. vivax* exposure prediction), and, using the Random Forest predictions from a 4-covariate model including PfGLURP R2, Etramp5.Ag1, GEXP18, and PfmSP1₁₉, accurately identified >92.0% of the *Plasmodium*-positive study samples from this study, whether detected through microscopy, RDT and/or PCR. It also correctly identified 87.2% of *P. falciparum* positives from a Malaysian dataset. The observation that this relatively simple model, which did not

and dots represent the observed proportions of seropositives per age divided into 10% centiles. For models assuming a change point in transmission, only the recent seroconversion rates and change point estimates (in years) are shown, while the historical seroconversion rates and seroreversion rates are detailed in Table 4.

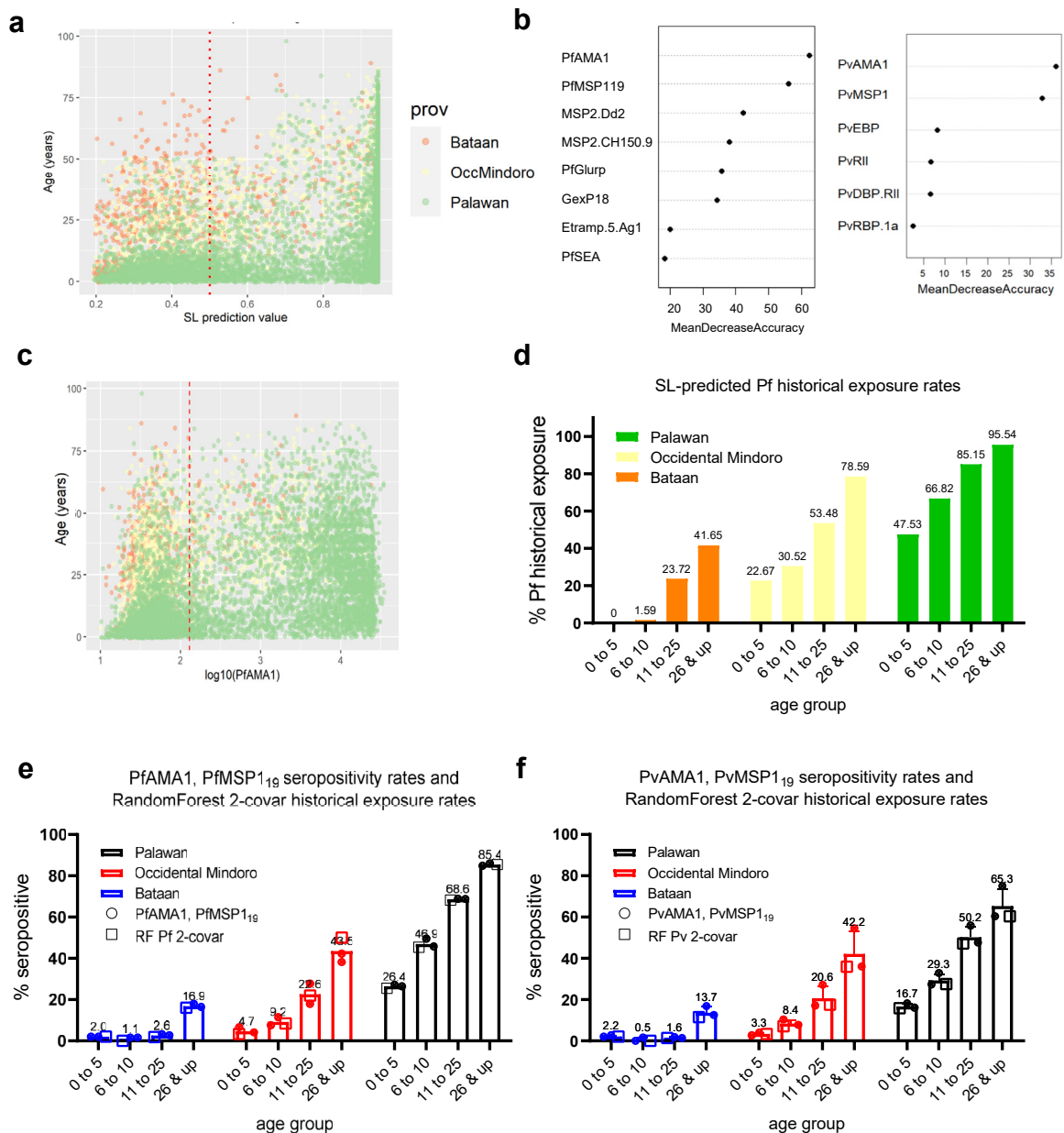


Fig. 6: Cumulative exposure markers confirm historical *P. falciparum* and *P. vivax* exposure, and heterogeneity of transmission in the 3 sites. **a)** Plot of Super Learner prediction values for Pf historical exposure by site and age, using the model with 8 Pf-specific serological markers as covariates. Red dotted line represents positivity cutoff at 0.5. **b)** Variable importance from the Random Forest model of the 8 Pf-specific antigens and 6 Pv-specific antigens in predicting historical exposure for each species. **c)** Distribution of antibody responses to PfAMA1 by site and age of individuals ($n = 9132$). Red dashed line represents the seropositivity cutoff value from the FMM model. **d and e)** Summary graphs per age category per site of SL-predicted Pf historical exposure (**d**) and PfAMA1, PfMSP1 seropositivity rates graphed with estimated historical exposure rates using the Random Forest model with PfAMA1, PfMSP1₁₉ as covariates (**e**). **f)** Summary graph of PvAMA1, PvMSP1 seropositivity rates with estimated historical exposure rates using the Random Forest model with PvAMA1, PvMSP1₁₉ as covariates per age category per site.

require epidemiological variables such as age, was able to provide accurate predictions for different datasets, shows its potential as a robust indicator for recent infections, in areas with varying levels of transmission.

We expected that the Super Learner model would improve the predictions with its ensemble approach; however, for predicting *P. falciparum* recent infections in particular, RF-based models had the better AUC for both training and test data. This was not the case for the

P. vivax predictions, which showed similar AUCs (<0.7) for the final SL model and its component models. It is likely that the low predictive ability we observed from the *P. vivax* panel is a limitation of the panel itself rather than the analytical approaches used. Notwithstanding, the SL approach was able to provide the means for evaluating multiple supervised ML models simultaneously, and, for the purpose of accurately distinguishing recent exposure from historical exposure with high sensitivity, the RF model seems to provide accurate estimates.

This study had other limitations. For classifying recent and historical exposure, we had no data on the malaria history of the survey participants, such that we instead focused on predicting for recent infections based on our cross-sectional data. Still, external validation of our prediction algorithm for *P. falciparum* infection confirms its promising performance in accurately predicting malaria infections. Also, since the study analyzed samples from a health facility-based survey, whether or not it accurately represents the populations needs to be assessed. The bias towards sampling health-seeking individuals was partially addressed by also recruiting companions of patients who visited the health facilities.⁵³ When we used our RF.covar4 model to predict recent exposure in the 2017 cross-sectional survey also conducted in Morong, Bataan, it also predicted 0.10% recent exposure—similar with our health facility-based data, despite observed differences in antibody levels between the studies (Supplementary Tables S1 and S3 and Fig. S7). The 29 to 30-year change point in transmission for Bataan can also be deemed consistent with the reported change point at 22 years in this previous survey.⁴³

Although we cannot validate the predicted historical exposure, we are able to show that consistent with previous studies,^{12,20,24,54,55} AMA1 and MSP1₁₉ seropositivity rates (individually or combined) for *P. falciparum* and *P. vivax* seemed to agree with the known malaria situation of our study sites. Changes in SCR over time also reflect this, with Bataan having the lowest, followed by Occidental Mindoro, and highest for Palawan. Our results are able to confirm the absence of *P. falciparum* and *P. vivax* transmission in malaria-free Bataan for the age groups 10 and below, providing evidence for its malaria elimination status. For Occidental Mindoro, the interruption in transmission from 7 to 12 years ago may be explained by the extensive malaria control activities implemented through the Global Fund for Malaria since the 2000s. Our seroconversion curve results also suggest an increase in malaria transmission in Rizal, Palawan during the time of our survey (2016–2018), and upon checking the records of reported malaria cases, there was indeed a steady rise of the API in this municipality from 25.43 in 2014, 33.97 in 2015, to 46.09 in 2016 and 56.78 in 2018.

Despite the mentioned limitations of the study, our results clearly show the potential use of multiplex antibody responses and applications of machine learning approaches in assessing malaria transmission for countries aiming for malaria elimination. Of interest for countries aiming for elimination is evaluating the impact of a decrease in transmission on the immunity or vulnerability of a population, especially in differing endemic settings.^{56,57} Immunological cohort studies are needed to investigate this further. In a sub-national elimination setting such as the Philippines, both recent and historical *P. falciparum* and *P. vivax* exposure metrics were indicative of the absence of recent transmission in Bataan and Occidental Mindoro, and also identified recent infections in Rizal, Palawan, thus showing its ability in assessing the malaria situation in varying areas of endemicity. Our study provides baseline immunological data for monitoring risk populations in the Philippines. This serological surveillance approach can aid in devising control measures by malaria elimination programs, as well as provide evidence of the effectiveness of programs being implemented.

Contributors

CJD, FEJE, JCRH, and MLMM conceived and designed this study. MLMM, KMF, and CJD analyzed and interpreted the data. MLMM wrote the first draft of the manuscript with support from KMF, CJD, JCRH, KY, and FEJE. JSL, RAR, MLMM, and APNB supervised the data and sample collection, and the conduct of assays. KKAT, JHA, CH, and CEC provided the recombinant antigens used for the multiplex bead-based assay. MLMM and RAR performed the serological assays with support from TH and KKAT. All authors read and approved the final manuscript.

Data sharing statement

All data analyzed in this study are available from authors upon reasonable request and with permission of relevant institutional review boards.

Declaration of interests

The authors declare no competing interests.

Acknowledgements

This study was supported by the Newton Fund, Philippine Council for Health Research and Development, and UK Medical Research Council through funding received for the ENSURE project (MR/N019199/1). We also express our gratitude to the ENSURE team, especially Carol Joy Sarsadiaz, Hennessey Sabanal, Beulah Boncayao, and Elaine de la Fuente, who assisted in the conduct of serological assays, as well as Dr. Catalino Demetria of RITM Special Pathogens Laboratory, who provided technical assistance and RITM Immunology laboratory, Natural Products Laboratory-DPSM in University of the Philippines-Manila, for access to their Luminex 200 instruments. Also, we thank the local governments and health staff of Palawan, Occidental Mindoro, and Bataan for their invaluable support to our research. This work is also supported by the Nagasaki University's "Doctoral Program for World-leading Innovative and Smart Education" for Global Health, KENKYU SHIDO KEIHI. KMF was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and Royal Society (grant no 221963/Z/20/Z).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.lanwpc.2023.100792>.

References

- 1 World Health Organization. *Progress towards subnational elimination in the Philippines*. Geneva PP - Geneva: World Health Organization; 2014.
- 2 World Health Organization. *World malaria report 2019*. Geneva, Switzerland. 2019.
- 3 Drakeley CJ, Corran PH, Coleman PG, et al. Estimating medium- and long-term trends in malaria transmission by using serological markers of malaria exposure. *Proc Natl Acad Sci U S A*. 2005;102(14):5108–5113.
- 4 Cotter C, Sturrock HJW, Hsiang MS, et al. The changing epidemiology of malaria elimination: new strategies for new challenges. *Lancet*. 2013;382(9895):900–911.
- 5 malERA Refresh Consultative Panel on Characterising the Reservoir and Measuring Transmission. malERA: An updated research agenda for characterising the reservoir and measuring transmission in malaria elimination and eradication. *PLoS Med*. 2017;14(11):e1002452.
- 6 World Health Organization. *A framework for malaria elimination*. 2017.
- 7 Fowkes FJ, Richards JS, Simpson JA, Beeson JG. The relationship between anti-merozoite antibodies and incidence of *Plasmodium falciparum* malaria: a systematic review and meta-analysis. *PLoS Med*. 2010;7(1):e1000218.
- 8 Folegatti PM, Siqueira AM, Monteiro WM, Lacerda MV, Drakeley CJ, Braga EM. A systematic review on malaria sero-epidemiology studies in the Brazilian Amazon: insights into immunological markers for exposure and protection. *Malar J*. 2017;16(1):107.
- 9 van den Hoogen LL, Griffin JT, Cook J, et al. Serology describes a profile of declining malaria transmission in Farafenni, the Gambia. *Malar J*. 2015;14(1):416.
- 10 Pothin E, Ferguson NM, Drakeley CJ, Ghani AC. Estimating malaria transmission intensity from *Plasmodium falciparum* serological data using antibody density models. *Malar J*. 2016;15(1):79.
- 11 Kerkhof K, Sluydts V, Willen L, et al. Serological markers to measure recent changes in malaria at population level in Cambodia. *Malar J*. 2016;15(1):529.
- 12 Idris ZM, Chan CW, Mohammed M, et al. Serological measures to assess the efficacy of malaria control programme on Ambae Island, Vanuatu. *Parasit Vectors*. 2017;10(1):204.
- 13 Ssewanyana I, Arinaitwe E, Nankabirwa JI, et al. Avidity of anti-malarial antibodies inversely related to transmission intensity at three sites in Uganda. *Malar J*. 2017;16(1):67.
- 14 Wu L, Mwesigwa J, Affara M, et al. Sero-epidemiological evaluation of malaria transmission in the Gambia before and after mass drug administration. *BMC Med*. 2020;18(1):331.
- 15 Corran P, Coleman P, Riley E, Drakeley C. Serology: a robust indicator of malaria transmission intensity? *Trends Parasitol*. 2007;23(12):575–582.
- 16 Staniscic DI, Fowkes FJ, Koinari M, et al. Acquisition of antibodies against *Plasmodium falciparum* merozoites and malaria immunity in young children and the influence of age, force of infection, and magnitude of response. *Infect Immun*. 2015;83(2):646–660.
- 17 Niass O, Saint-Pierre P, Niang M, et al. Modelling dynamic change of malaria transmission in holoendemic setting (Dielmo, Senegal) using longitudinal measures of antibody prevalence to *Plasmodium falciparum* crude schizonts extract. *Malar J*. 2017;16(1):409.
- 18 Fouda GG, Leke RF, Long C, et al. Multiplex assay for simultaneous measurement of antibodies to multiple *Plasmodium falciparum* antigens. *Clin Vaccine Immunol*. 2006;13(12):1307–1313.
- 19 Ondigo BN, Park GS, Gose SO, et al. Standardization and validation of a cytometric bead assay to assess antibodies to multiple *Plasmodium falciparum* recombinant antigens. *Malar J*. 2012;11:427.
- 20 Perraut R, Richard V, Varela ML, et al. Comparative analysis of IgG responses to *Plasmodium falciparum* MSP1p19 and PF13-DBL1alpha1 using ELISA and a magnetic bead-based duplex assay (MAGPIX(R)-Luminex) in a Senegalese meso-endemic community. *Malar J*. 2014;13(1):410.
- 21 Koffi D, Toure AO, Varela ML, et al. Analysis of antibody profiles in symptomatic malaria in three sentinel sites of Ivory Coast by using multiplex, fluorescent, magnetic, bead-based serological assay (MAGPIX). *Malar J*. 2015;14(1):509.
- 22 Wu L, Hall T, Ssewanyana I, et al. Optimisation and standardisation of a multiplex immunoassay of diverse *Plasmodium falciparum* antigens to assess changes in malaria transmission using sero-epidemiology. *Wellcome Open Res*. 2019;4:26.
- 23 Helb DA, Tetteh KK, Felgner PL, et al. Novel serologic biomarkers provide accurate estimates of recent *Plasmodium falciparum* exposure for individuals and communities. *Proc Natl Acad Sci U S A*. 2015;112(32):E4438–E4447.
- 24 Wu L, Mwesigwa J, Affara M, et al. Antibody responses to a suite of novel serological markers for malaria surveillance demonstrate strong correlation with clinical and parasitological infection across seasons and transmission settings in the Gambia. *BMC Med*. 2020;18(1):304.
- 25 Longley RJ, White MT, Takashima E, et al. Development and validation of serological markers for detecting recent *Plasmodium vivax* infection. *Nat Med*. 2020;26(5):741–749.
- 26 Coutts SP, King JD, Pa'au M, et al. Prevalence and risk factors associated with lymphatic filariasis in American Samoa after mass drug administration. *Trop Med Health*. 2017;45(1):22.
- 27 van den Hoogen LL, Presume J, Romilus I, et al. Quality control of multiplex antibody detection in samples from large-scale surveys: the example of malaria in Haiti. *Sci Rep*. 2020;10(1):1135.
- 28 Collins CR, Withers-Martinez C, Bentley GA, Batchelor AH, Thomas AW, Blackman MJ. Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *J Biol Chem*. 2007;282(10):7431–7441.
- 29 Burghaus PA, Holder AA. Expression of the 19-kilodalton carboxy-terminal fragment of the *Plasmodium falciparum* merozoite surface protein-1 in *Escherichia coli* as a correctly folded protein. *Mol Biochem Parasitol*. 1994;64(1):165–169.
- 30 Theisen M, Vuust J, Gottschau A, Jepsen S, Høgh B. Antigenicity and immunogenicity of recombinant glutamate-rich protein of *Plasmodium falciparum* expressed in *Escherichia coli*. *Clin Diagn Lab Immunol*. 1995;2(1):30–34.
- 31 van den Hoogen LL, Walk J, Oulton T, et al. Antibody responses to antigenic targets of recent exposure are associated with low-density parasitemia in controlled human *Plasmodium falciparum* infections. *Front Microbiol*. 2019;9(January):1–11.
- 32 Spielmann T, Ferguson DJ, Beck HP, etramps, a new *Plasmodium falciparum* gene family coding for developmentally regulated and highly charged membrane proteins located at the parasite-host cell interface. *Mol Biol Cell*. 2003;14(4):1529–1544.
- 33 Raj DK, Nixon CP, Nixon CE, et al. Antibodies to PfSEA-1 block parasite egress from RBCs and protect against malaria infection. *Science*. 2014;344(6186):871–877.
- 34 Polley SD, Conway DJ, Cavanagh DR, et al. High levels of serum antibodies to merozoite surface protein 2 of *Plasmodium falciparum* are associated with reduced risk of clinical malaria in coastal Kenya. *Vaccine*. 2006;24(19):4233–4246.
- 35 Taylor RR, Smith DB, Robinson VJ, McBride JS, Riley EM. Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of the immunoglobulin G3 subclass. *Infect Immun*. 1995;63(11):4382–4388.
- 36 Chuquiyauri R, Molina DM, Moss EL, et al. Genome-Scale protein microarray comparison of human antibody responses in *Plasmodium vivax* relapse and reinfection. *Am J Trop Med Hyg*. 2015;93(4):801–809.
- 37 Franca CT, Hostetler JB, Sharma S, et al. An antibody screen of a *Plasmodium vivax* antigen library identifies novel merozoite proteins associated with clinical protection. *PLoS Negl Trop Dis*. 2016;10(5):e0004639.
- 38 Franca CT, He WQ, Gruszczyn J, et al. *Plasmodium vivax* reticulocyte binding proteins are key targets of naturally acquired immunity in young Papua New Guinean children. *PLoS Negl Trop Dis*. 2016;10(9):e0005014.
- 39 Ntumngia FB, Schloegel J, Barnes SJ, et al. Conserved and variant epitopes of *Plasmodium vivax* Duffy binding protein as targets of inhibitory monoclonal antibodies. *Infect Immun*. 2012;80(3):1203–1208.
- 40 Hester J, Chan ER, Menard D, et al. De novo assembly of a field isolate genome reveals novel *Plasmodium vivax* erythrocyte invasion genes. *PLoS Negl Trop Dis*. 2013;7(12):e2569.
- 41 Menard D, Chan ER, Benedet C, et al. Whole genome sequencing of field isolates reveals a common duplication of the Duffy binding protein gene in Malagasy *Plasmodium vivax* strains. *PLoS Negl Trop Dis*. 2013;7(11):e2489.
- 42 Hubbard A, Munoz ID, Decker A, et al. Time-dependent prediction and evaluation of variable importance using superlearning in high-dimensional clinical data. *J Trauma Acute Care Surg*. 2013;75(1 Suppl 1):S53–S60.

- 43 van den Hoogen LL, Bareng P, Alves J, et al. Comparison of commercial ELISA kits to confirm the absence of transmission in malaria elimination settings. *Front Public Health*. 2020;8:480.
- 44 Fornace KM, Brock PM, Abidin TR, et al. Environmental risk factors and exposure to the zoonotic malaria parasite *Plasmodium knowlesi* across northern Sabah, Malaysia: a population-based cross-sectional survey. *Lancet Planet Health*. 2019;3(4):e179–e186.
- 45 Arnold BF, van der Laan MJ, Hubbard AE, et al. Measuring changes in transmission of neglected tropical diseases, malaria, and enteric pathogens from quantitative antibody levels. *PLoS Neglected Trop Dis*. 2017;11(5):1–20.
- 46 Fornace KM, Herman LS, Abidin TR, et al. Exposure and infection to *Plasmodium knowlesi* in case study communities in Northern Sabah, Malaysia and Palawan, the Philippines. *PLoS Negl Trop Dis*. 2018;12(6):e0006432.
- 47 Sepulveda N, Stresman G, White MT, Drakeley CJ. Current mathematical models for analyzing anti-malarial antibody data with an eye to malaria elimination and eradication. *J Immunol Res*. 2015;2015:738030.
- 48 Fornace KM, Surendra H, Abidin TR, et al. Use of mobile technology-based participatory mapping approaches to geolocate health facility attendees for disease surveillance in low resource settings. *Int J Health Geogr*. 2018;17(1):21.
- 49 van den Hoogen LL, Stresman G, Presume J, et al. Selection of antibody responses associated with *Plasmodium falciparum* infections in the context of malaria elimination. *Front Immunol*. 2020;11(May):928.
- 50 Rogier E, Nace D, Dimbu PR, et al. Framework for characterizing longitudinal antibody response in children after *Plasmodium falciparum* infection. *Front Immunol*. 2021;12(March):617951.
- 51 Chotirat S, Nekkab N, Kumpitak C, et al. Application of 23 novel serological markers for identifying recent exposure to *Plasmodium vivax* parasites in an endemic population of western Thailand. *Front Microbiol*. 2021;12:643501.
- 52 Rosado J, Pelleau S, Cockram C, et al. Multiplex assays for the identification of serological signatures of SARS-CoV-2 infection: an antibody-based diagnostic and machine learning study. *Lancet Microbe*. 2021;2(2):e60–e69.
- 53 Reyes RA, Fornace KM, Macalinao MLM, et al. Enhanced health facility surveys to support malaria control and elimination across different transmission settings in the Philippines. *Am J Trop Med Hyg*. 2021;104(3):968–978.
- 54 Rosas-Aguirre A, Speybroeck N, Llanos-Cuentas A, et al. Hotspots of malaria transmission in the Peruvian Amazon: rapid assessment through a parasitological and serological survey. *PLoS One*. 2015;10(9):e0137458.
- 55 Biggs J, Raman J, Cook J, et al. Serology reveals heterogeneity of *Plasmodium falciparum* transmission in northeastern South Africa: implications for malaria elimination. *Malar J*. 2017;16(1):48.
- 56 Fowkes FJI, Boeuf P, Beeson JG. Immunity to malaria in an era of declining malaria transmission. *Parasitology*. 2016;143:139–153.
- 57 World Health Organization. *WHO malaria policy advisory committee meeting: meeting report, October 2017*. Geneva, Switzerland. 2017.