

SerolyzeR - an R package for automated analysis of serological data

Nuno Sepulveda, nuno.sepulveda@pw.edu.pl

UseR! 2026 conference

Warsaw 07.07.2026



Faculty of Mathematics and Information Science

WARSAW UNIVERSITY OF TECHNOLOGY

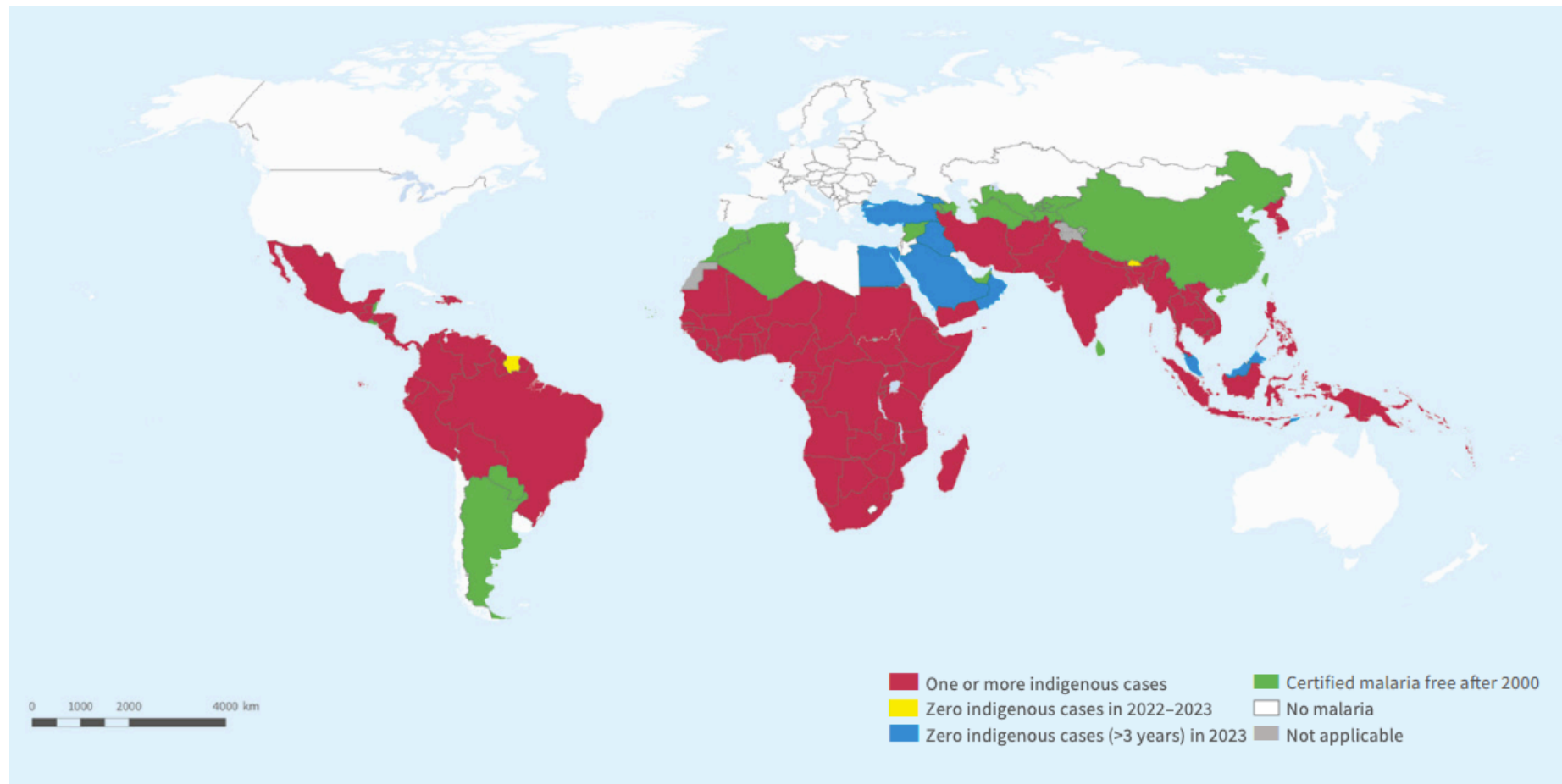
Disclaimer

Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or HADEA. Neither the European Union nor the granting authority can be held responsible for them



Funded by
the European Union

Malaria



World Malaria Report (2024). WHO

The big challenge of malaria elimination



Plasmodium falciparum

More virulent and lethal

Less diverse genome

No dormant infection

Anti-malarial drugs are efficacious

Easier to detect and control

Plasmodium vivax

Less virulent and lethal

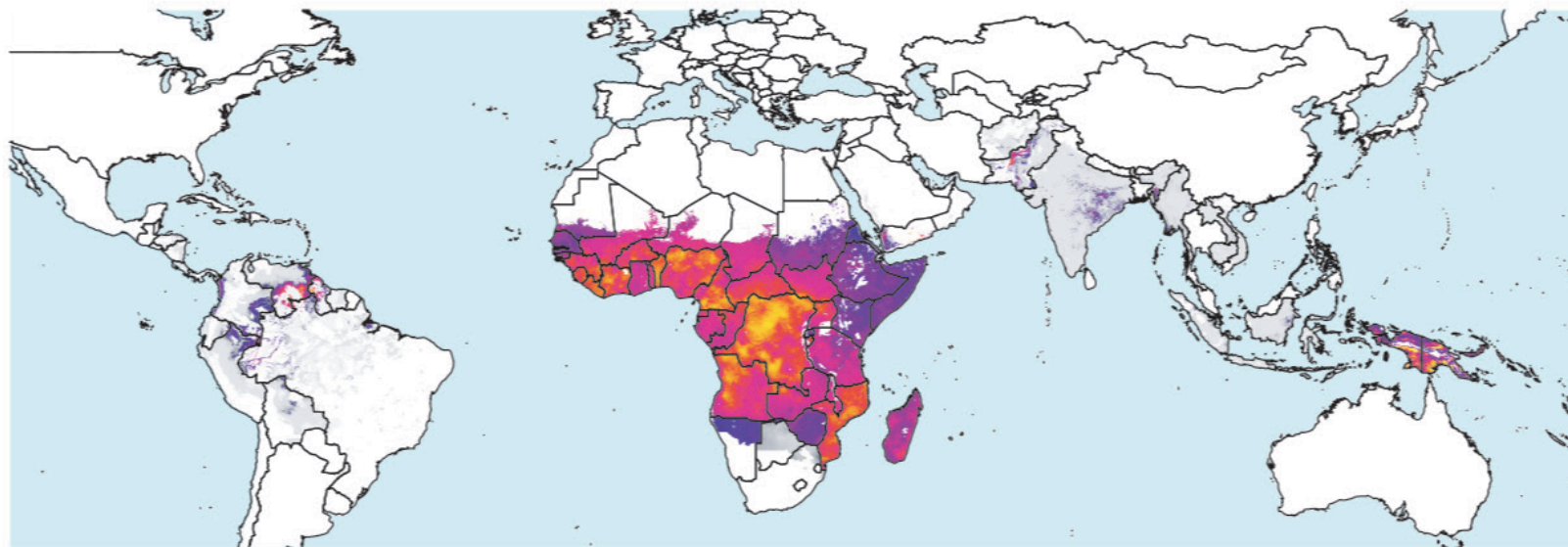
Highly diverse genome

Dormant infection/reactivation

Anti-malarial drugs are less efficacious

Less easy to detect and control

Global distribution of *P. falciparum*



P. falciparum incidence rate 2022
(new cases per 1000 people)



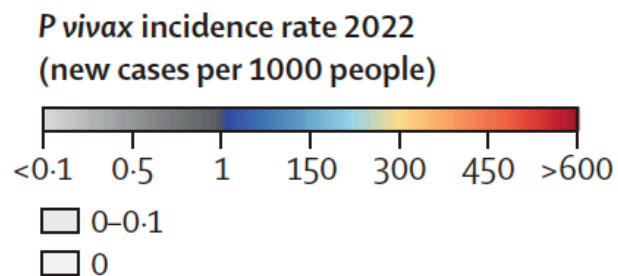
0

Sparsely populated

Weiss DJ et al. Lancet. 2025;405(10483):979-990.

Global distribution of *P. vivax*

Plasmodium vivax



Weiss DJ et al. Lancet. 2025;405(10483):979-990.

Identifying individuals with dormant infections



Plasmodium falciparum

More virulent and lethal

Less diverse genome

No dormant infection

Anti-malarial drugs are efficacious

Easier to detect and control

**Treating
people**



Plasmodium vivax

Less virulent and lethal

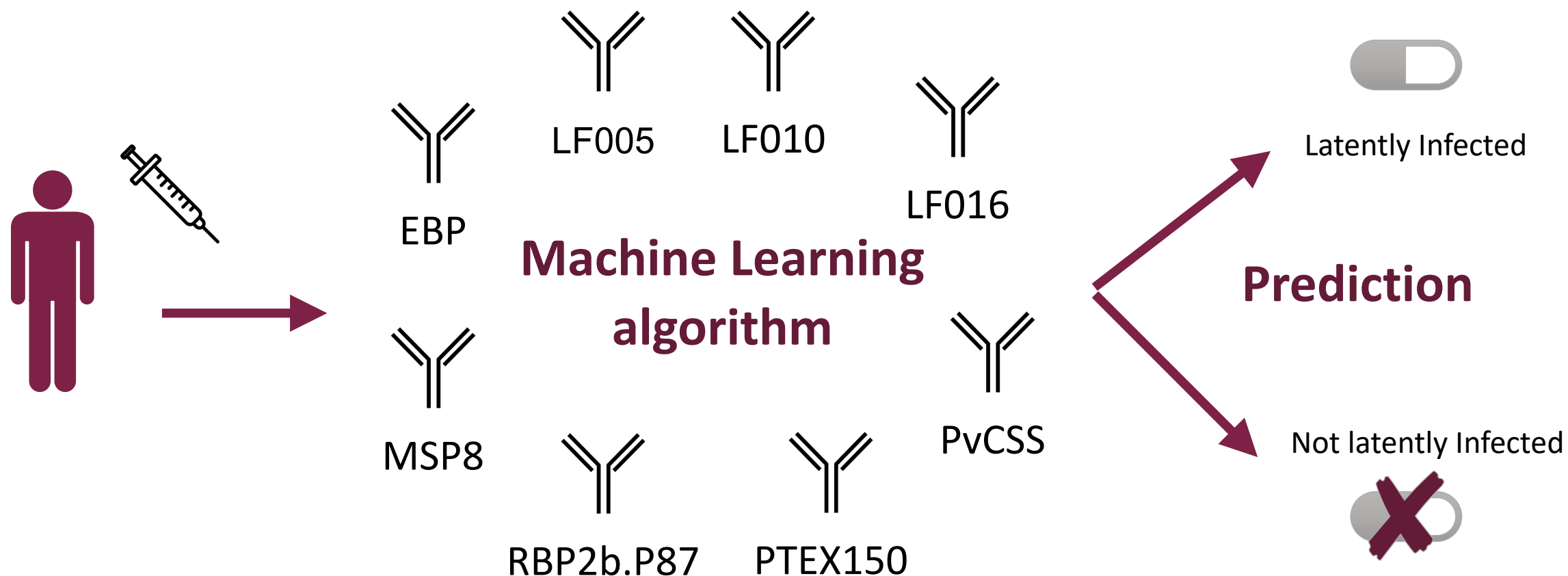
Highly diverse genome

Dormant infection/reactivation

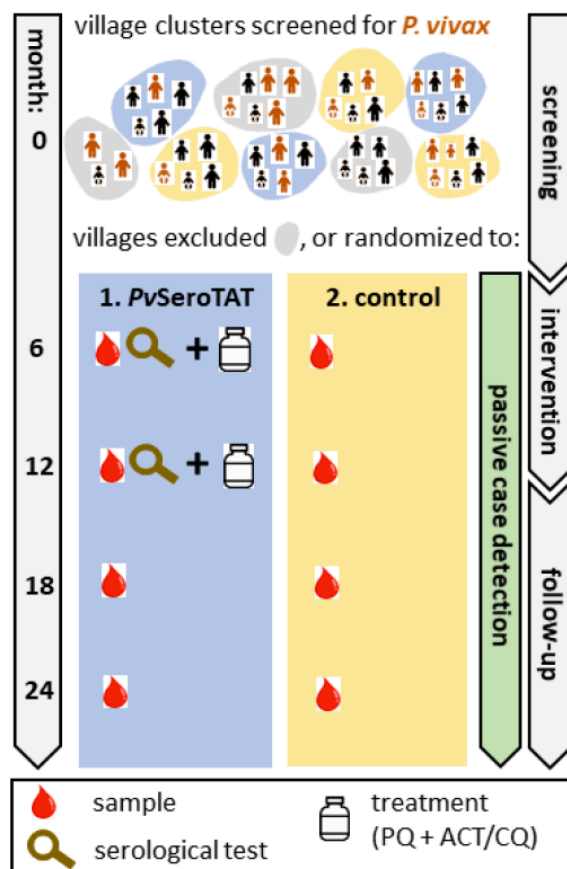
Anti-malarial drugs are less efficacious

Less easy to detect and control

Basic idea



Two clinical trials (Madagascar and Ethiopia)



Box 2: Cluster Randomised Trial design summary.

- **Sites.** (i) Ethiopia; (ii) Madagascar.
- **Arms.** (i) Control; (ii) Intervention: *PvSeroTAT*.
- **Clusters screened per site.** 36
- **Clusters per arm per site.** 12
- **Total clusters.** (2 sites)*(2 arms)*(12 clusters) = 48
- **Participants per cluster.** 400 (200 – 600)
- **Samples per cluster.** 200
- **Target prevalence.** 3% (1% – 10%) *P. vivax* PCR prevalence.
- **Coefficient of variation.** $k = 0.4$ (0.3 – 0.5)
- **Intervention diagnostic test.** Serological test for recent *P. vivax* infection.
- **Intervention treatment.** Primaquine for 7 days at 1 mg/kg with G6PD testing. Directly observed.
- **Intervention rounds.** 2 rounds 6 months apart.
- **Expected effect size.** 60% (50 – 70%) reduction in *P. vivax* PCR prevalence.
- **Duration of follow-up.** 12 months.



PvSTATEM - New tools for the control and elimination of *P. vivax* ~~malaria~~



Study sites: Ethiopia, Madagascar

Malaria eradication requires addressing all malaria species, including *P. vivax* that resists elimination efforts because of its dormant liver forms ("hypnozoites"). These forms cannot be detected with traditional methods, but they can reactivate ("relapse infection") causing illness and new transmission from people to mosquitoes.

The *P. vivax* Serological Testing and Treatment in Ethiopia and Madagascar (PvSTATEM) project aims to address this challenge.

We have developed a diagnostic test to identify people likely to carry hypnozoites. These people can be targeted for treatment with primaquine to clear hypnozoites and prevent relapse infection. This combination of a novel diagnostic test and primaquine treatment provides a new intervention for malaria control and elimination.

<http://pvstatem.eu>

Project partners



France

UK

Ethiopia

Madagascar



Ireland

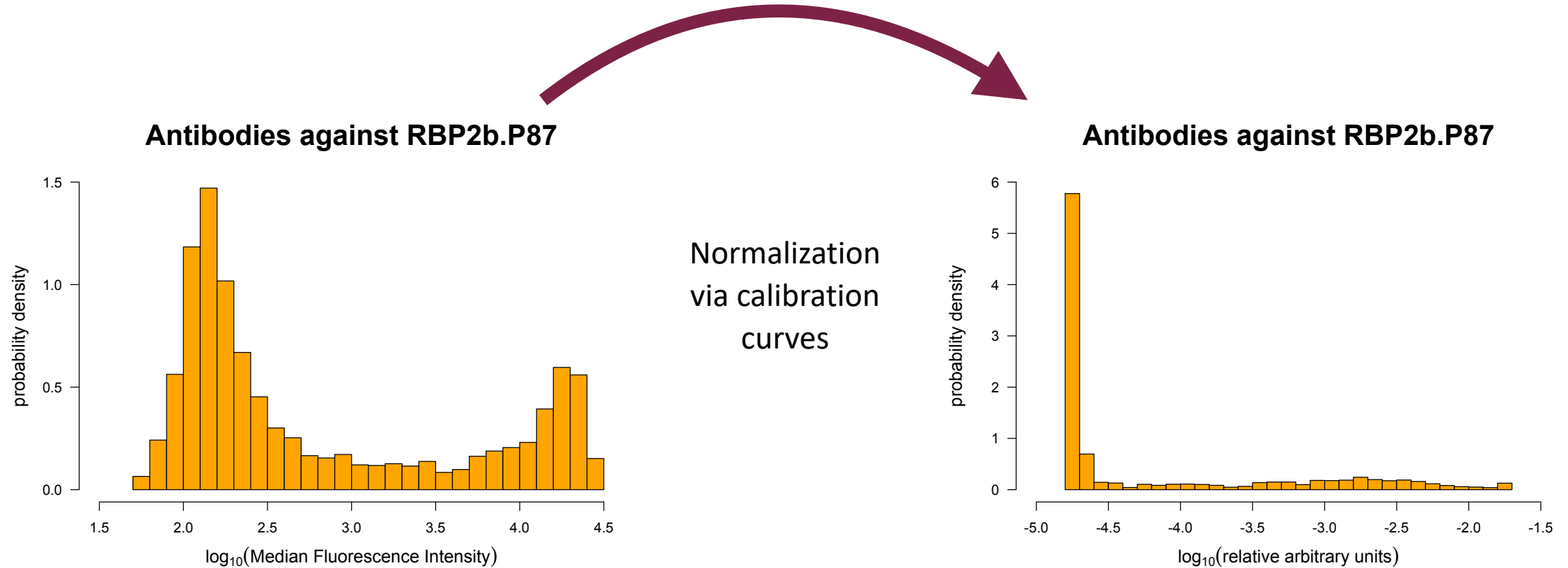
Switzerland

Italy

Australia

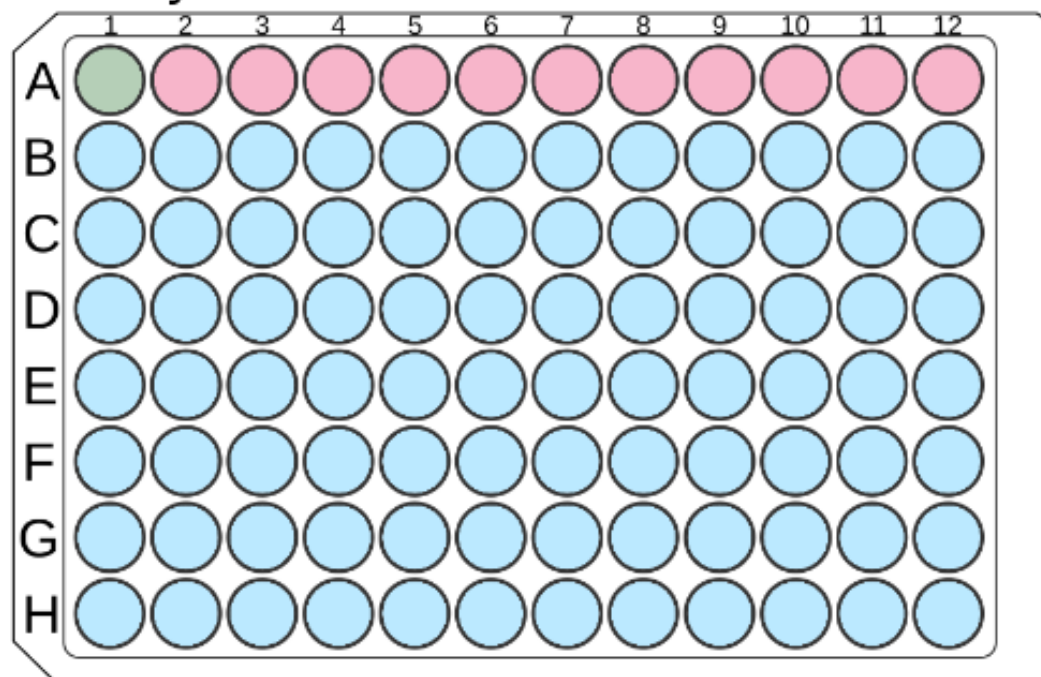
Poland

Serological data is central to the project

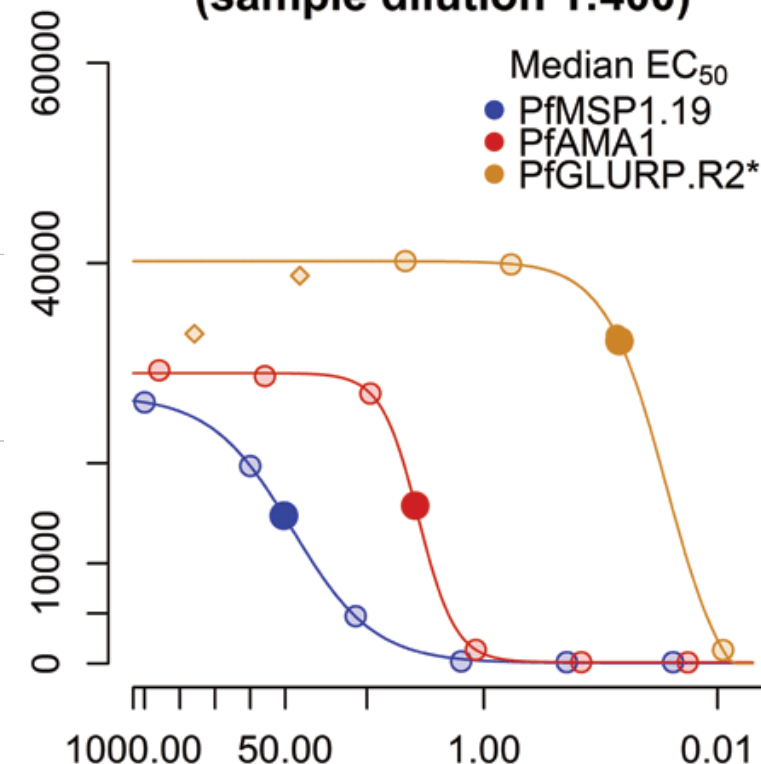


Data is generated in batches

Layout of CovidOISExPONENT



CP3 control
(sample dilution 1:400)



Challenges

Ensure high quality within a batch (calibration curve/bead counts)

Ensure comparability among batches (calibration curve/data normalisation)

Ensure feasibility to large number of samples (e.g., $n \sim 3500$ samples at baseline in Madagascar) and of analytes (e.g., ~ 8 antibodies in PvSTATEM)

No software available for processing serological data (**Opportunity**)

serostat

SerolyzeR package for R



SerolyzeR: Reading, Quality Control and Preprocessing of MBA (Multiplex Bead Assay) Data

Speeds up the process of loading raw data from MBA (Multiplex Bead Assay) examinations, performs quality control checks, and automatically normalises the data, preparing it for a user, who does not necessarily have experience with R language. The package is developed within the project 'PvSTATEM', which is an international project aiming for malaria

Version: 1.4.1
Imports: [dplyr](#), [ggplot2](#), [nplr](#), [R6](#), [readxl](#), [stringi](#), [stringr](#), [grid](#), [png](#), [tools](#), [ggrepel](#), [lubridate](#), [R.utils](#), [svglite](#), [fs](#), [scales](#), [rlang](#), [patchwork](#), [cellranger](#)
Suggests: [knitr](#), [qpdf](#), [furr](#), [future](#), [qs2](#), [rmarkdown](#), [testthat](#) (≥ 3.0.0)
Published: 2026-02-20
DOI: [10.32614/CRAN.package.SerolyzeR](#)
Author: Jakub Grzywaczewski [aut, cre], Tymoteusz Kwiecinski [aut], Mateusz Nizwantowski [aut], Przemyslaw Biecek [ths], Nuno Sepulveda [ths]
Maintainer: Jakub Grzywaczewski <jakubzgrzywaczewski at gmail.com>
BugReports: <https://github.com/mini-pw/SerolyzeR/issues>
License: [BSD 3 clause](#) + file [LICENSE](#)
URL: <https://github.com/mini-pw/SerolyzeR>, <https://mini-pw.github.io/SerolyzeR/>
NeedsCompilation: no
Materials: [README](#)
CRAN checks: [SerolyzeR results](#)

Documentation:

Reference manual: [SerolyzeR.html](#) , [SerolyzeR.pdf](#)
Vignettes: [Basic SerolyzeR functionalities](#) (source, [R code](#))

Downloads:

Package source: [SerolyzeR_1.4.1.tar.gz](#)
Windows binaries: r-devel: [SerolyzeR_1.4.1.zip](#), r-release: [SerolyzeR_1.4.1.zip](#), r-oldrel: [SerolyzeR_1.4.1.zip](#)
macOS binaries: r-release (arm64): [SerolyzeR_1.4.1.tgz](#), r-oldrel (arm64): [SerolyzeR_1.4.1.tgz](#), r-release (x86_64): [SerolyzeR_1.4.1.tgz](#), r-oldrel (x86_64): [SerolyzeR_1.4.1.tgz](#)
Old sources: [SerolyzeR archive](#)

Current
version
1.4.1

ggplot2
(Plotting)

nlpr
(N-parameter logistic
regression)

Software development team



Tymoteusz Kwiecinski

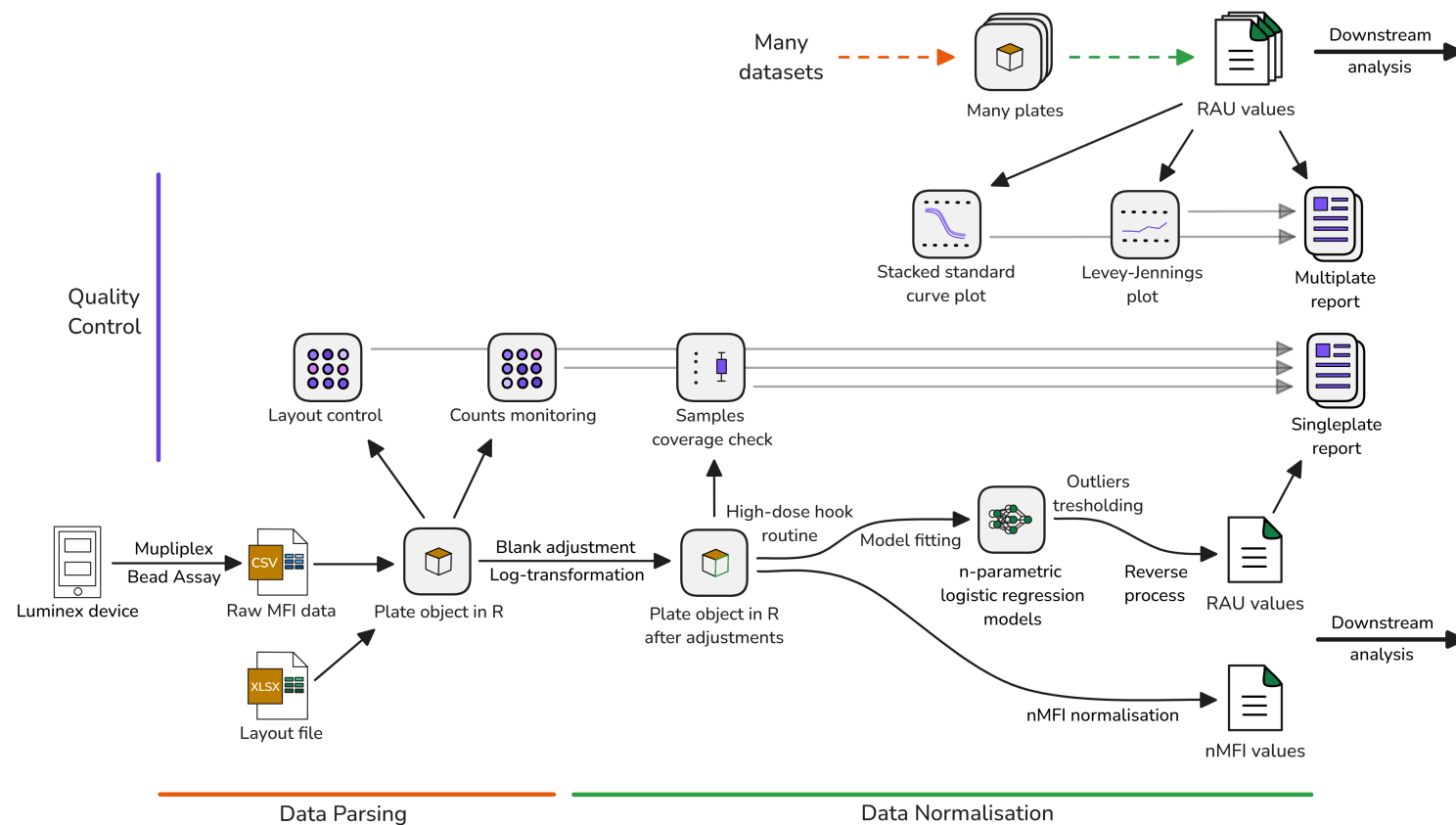


Jakub Grzywaczewski



Mateusz Nizwantowski

Overview



Very important!
Feedback and testing from lab researchers with different R skills/knowledge.

Very important!
Wiki page with comprehensive tutorials.

Very important!
Shiny apps.

Take-home messages

Research-driven opportunities

Collect information on research protocols related to serological data and their processing.

Standardise data labelling and quality control protocols

Creation of a new package with automated functionality on serological data processing.

To work with strong research groups who are in the forefront of generation data

Possible Impact

Increase data quality and the quality of the subsequent research output

Utilisation of the package beyond PvSTATEM (e.g., sero-surveillance of multiple infectious diseases).

Opportunity to impact public health



ORIGINAL RESEARCH article

Front. Public Health, 25 October 2022
Sec. Infectious Diseases – Surveillance, Prevention and Treatment
Volume 10 - 2022 | <https://doi.org/10.3389/fpubh.2022.924316>

Assessing seroprevalence and associated risk factors for multiple infectious diseases in Sabah, Malaysia using serological multiplex bead assays

YuYen L. Chan ^{1*} Catriona L. Patterson ¹
Jeffrey W. Priest ² Gillian Stresman ¹ Timothy William ³
Tock H. Chua ⁴ Kevin Tetteh ¹ Patrick Lammie ⁵
Chris Drakeley ¹ Kimberly M. Fornace ^{6,7}

communications medicine

Article

A Nature Portfolio journal

<https://doi.org/10.1038/s43856-025-01162-5>

Investigation of the sero-epidemiology of vaccine preventable diseases and common viral infections in French populations

Emma Bloch ¹, Gaëlle Baudemont ¹, Françoise Donnadieu ¹, Laura Garcia ¹, Stéphane Pelleau ¹, SeroPed Study Consortium ¹, Milieu Intérieur Consortium ¹, Lluís Quintana-Murci ^{2,3,4,5}, Darragh Duffy ^{4,5,6}, Arnaud Fontanet ^{4,6} & Michael White ¹

nature communications



Article

<https://doi.org/10.1038/s41467-025-62305-9>

Multiplex bead assays enable integrated serological surveillance and reveal cross-pathogen vulnerabilities in Zambezia Province, Mozambique

Received: 12 December 2024

Accepted: 18 July 2025

Published online: 26 August 2025

Check for updates

Andrea C. Carcelen ¹, Celso Monjane ², Sophie Bérubé ³, Saki Takahashi ⁴, Thebora Sultane ², Imelda Chelene ², Gretchen Cooley ⁵, E. Brook Goodhew ⁵, Catriona Patterson ⁶, Kevin Tetteh ⁶, Manuel Mutambe ², Melissa M. Higdon ¹, George Mwinnyaa ¹, Gilberto Nhapi ⁷, Pedro Duce ⁷, Diana L. Martin ⁵, Christopher Drakeley ⁶, William J. Moss ^{1,4} & Ivalda Macicame ²

The Journal of Infectious Diseases

MAJOR ARTICLE



Whose Line Is it Anyway? Defining Seropositivity Cutoffs for Infectious Disease Surveillance

Michael T. White ¹, Gaëlle Baudemont ^{1,2}, Françoise Donnadieu ¹, Arona Sabene Diatta ², Ibrahima Sarr ², Aissatou Diagne ², Aissatou Toure-Balde ², Fatoumata Diene Sarr ², Joseph Faye ², Cheikh Sokhna ⁴, Hubert Bassene ⁴, Inès Vigan-Womas ², Stéphane Pelleau ¹ and Makhtar Niang ²

Acknowledgements



Prof. Przemysław Biecek



Jakub Grzywaczewski



Agata Balak-Kuśmierska



PvSTATEM family!