**Team FORSEA: Members and Methods**

**Our team**

We are an international team of 4 people:

Esther Van Kleef (Dutch)1: Postdoctoral research fellow working on optimal antibiotic use and understanding and predicting its effects on selection and transmission of bacterial resistance.

Rene Niehus (German)1: Postdoctoral research fellow investigating the driving factors of antibacterial resistance and forecasting Dengue

James Watson (UK)1: Postdoctoral research fellow working on statistical modelling of Malaria epidemiology

Chris Mercado (Philippines)1: MSc student working on spatial risk mapping of malaria.

1Mahidol Oxford Tropical Medicine Research Unit (MORU), Bangkok, Thailand, Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom

**Our methods**

*Optimising an ensemble*

For the point forecasts we developed an ensemble model that combines **LASSO** and **Random Forests**. Each model was optimised individually by evaluation on a test data set. We used the mean squared errors (MSE) to determine model quality.

*From point forecasts to bin probabilities*

We use the mean point forecasts of each model (LASSO and Random Forest) and the real observations for one season (52 weeks) to determine the distributions of the residuals. On log-scale, the residuals are roughly normally distributed. We then determine for each model the standard error of the residuals, and then we use the mean forecast and the residual standard error to compute the bin probabilities.

*Weighing the models*

To combine the two models, we take the bin probabilities from each add them with a weighing factor. We use 1/MSE for each forecast week to weigh the 2 different models.

As covariates in our models we use (with various different lags):

• Season week

• Sin of season week

• Previous cases

• Change in cases

• School holidays

• Average national temperature

• Google search requests for “fever”

• Google search requests for “headache”

• Google search requests for “doctor”

• Google search requests for “shivering”

• Google search requests for “cough”

*Peak prediction, season onset and peak incidence*

For the season peak week, intensity and start week predictions, we also used penalised LASSO regression but taking as predictors the fluctuations in weighted ILI throughout the months of Sep-Nov.