**Refining the characterization of the Sanofi Pasteur dengue vaccine’s efficacy profile using machine learning**

Dorigatti I1, Donnelly CA1, Laydon DJ1, Small R2, Jackson N3, Coudeville L4, Ferguson NM1

1MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London W2 1PG

2Sanofi Pasteur, 2501 Discovery Dr, Orlando, FL 32826, USA

3Sanofi Pasteur, 1541 Avenue Marcel Mérieux, 69280 Marcy l'Étoile, France

4Sanofi Pasteur, 2 Avenue Pont Pasteur, 69367 Lyon Cedex 07, France

Corresponding author: i.dorigatti@imperial.ac.uk

**Abstract**

**Background**

CYD-TDV is the first licensed dengue vaccine for individuals 9 – 45 (or 60) years of age. Based on a subset of subjects for which baseline serostatus was measured with PRNT50 (12% of trial subjects) in the phase-2b and phase-3 trials, the vaccine was found to confer good protection against virologically-confirmed dengue in subjects pre-exposed to dengue but lower protection in individuals seronegative to dengue at baseline. A complete characterization of how efficacy varies by pre-vaccination dengue serostatus is lacking and can be informed by machine learning.

**Methods**

We applied Boosted Regression Tree to impute the baseline serostatus for a subset of the subjects with missing information (i.e. those with post-dose 3 PRNT50 titres) in the phase-2b and phase-3 trials of CYD-TDV. By merging individual-level imputations with group-level predictions of baseline serostatus, we provide more precise estimates of how efficacy varies with serostatus, age and serotype during active surveillance.

**Findings**

Over all ages and in the 9-16 age-group, we found that CYD-TDV is protective against dengue serotypes 1, 3 and 4 regardless of baseline serostatus, while efficacy against dengue serotype 2 is positive for dengue pre-exposed subjects. Most notably, we find support for age variation in efficacy independent of serostatus.