-

Information and Communication Technologies

EPIWORK

Developing the Framework for an Epidemic Forecast Infrastructure

http://www.epiwork.eu

Project no. 231807

D1.5 Model parameterization, simulation, prediction and quantification of uncertainty

Period covered: months 36th - 54th
Start date of project: February 1st, 2009

Due date of deliverable: month 54th
Distribution: public

Date of preparation: Month 54th
Duration:
Actual submission date: August, 8th 2013
Status:

Project Coordinator: Alessandro Vespignani
Project Coordinator Organization Name: ISI Foundation
Lead contractor for this deliverable: Max Planck Gesellschaft zur Foerderung der Wissenschaften E.V.

CONTENTS

Deliverable 1.5 - Model parameterization, simulation, prediction and quantification of uncertainty

Introduction

Modelling and making future predictions from long term disease surveillance time-series is perhaps the most difficult yet ultimate goal of theoretical epidemiology. Moreover statistical modelling methodologies have changed enormously over the last ten years with modern Bayesian and computational approaches revolutionising the field. In this Deliverable, the various partners of Work Package 1 have explored state-of-the-art statistical methodologies for fitting historical disease time-series and in same cases predicting future scenarios. This work bridges the activities of other Work Packages and involves analysis of the internet surveillance data collected in the Epiwork project across Europe.

In this deliverable, the several contributions of partners involved in the WP1 are reported: Gulbenkian Institute (FGC-IGC), London School of Hygiene and Tropical Medicine (LSHTM), Tel Aviv University (TAU).

1. Influenza modelling and prediction for European countries (EpiWork surveillance data)

Mathematical models in biology are powerful tools for the study and exploration of complex dynamics. Nevertheless, bringing theoretical results to an agreement with experimental observations involves acknowledging a great deal of uncertainty intrinsic to our theoretical representation of a real system. Proper handling of such uncertainties is key to the successful usage of models to predict experimental or field observations. This problem has been addressed over the years by many tools for model calibration and parameter estimation. The IGC group developed a general framework for uncertainty analysis and parameter estimation that is designed to handle uncertainties associated with the modeling of dynamic biological systems while remaining agnostic as to the type of model used (Coelho et al 2011). The framework was applied to fit an SIR-like influenza transmission model to 7 years of incidence data in three European countries: Belgium, The Netherlands and Portugal (Coehlo et al. 2011; van Noort et al 2012). The basic reproduction number (R_0) was estimated to be consistent with the literature, providing a positive test for the method. From the estimates obtained for the remaining model parameters we were able to formulate a hypothesis relating the proportion of symptomatic cases among all influenza infections in a season to the timing of epidemic occurrence, this proportion being higher when the epidemic peaks in the colder and dryer Winter months. Figure 1 illustrates how hypothesis was further elaborated and parameters estimated to link the probability of developing symptoms upon infection to temperature and humidity variables measured by weather stations in the three countries (van Noort et al 2012).

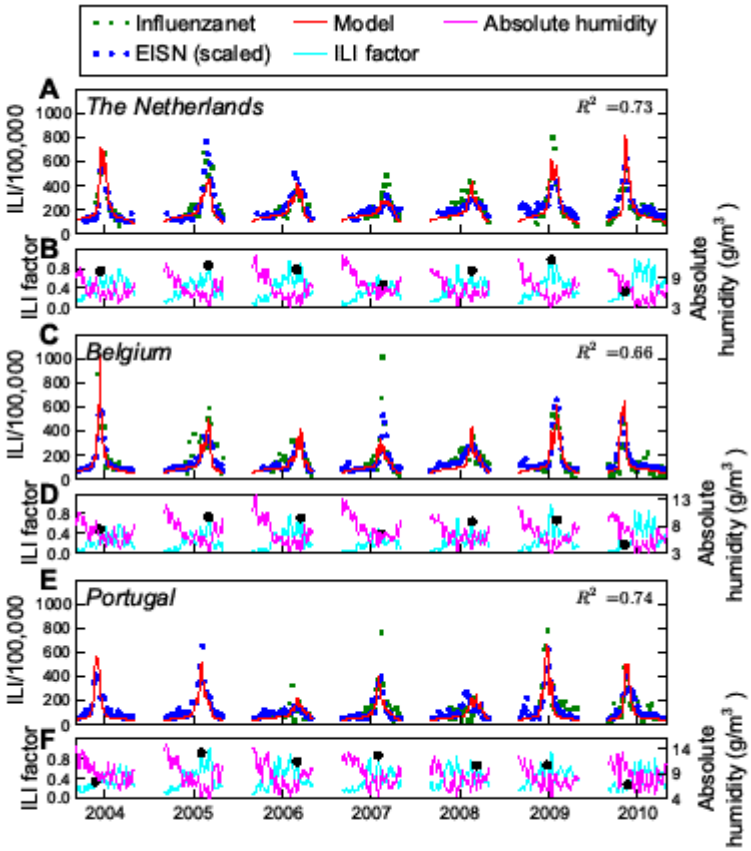

Figure 1: Fitting of influenza model based on the measured weekly absolute humidity. The ILI factor is determined by the typical weekly absolute humidity. (A) The Netherlands, (C) Belgium, and (E) Portugal (2003–2010). The measured absolute humidity and corresponding ILI factor are shown for (B) The Netherlands, (D) Belgium, and (F) Portugal. The black dots on the ILI factor curve indicate the time of the epidemic peaks.

2. Influenza modelling and prediction for Israeli data

In parallel, in a series of studies (Yaari et al 2013, Axelson et al. 2013, Huppert et al. 2013) the Tel Aviv University group were also heavily involved in modeling human influenza. Their interest was likewise concerned with the debate that surrounds the relative role of epidemic dynamics, viral evolution and climatic drivers in driving year-to-year variability of outbreaks. They maintain that the ultimate test of understanding is prediction, yet existing influenza models rarely forecast beyond a single year at best. The TAU group use a simple epidemiological model to reveal remarkable multiannual predictability based on high quality influenza surveillance data for Israel. Successful forecasts are driven by temperature, humidity, antigenic drift and immune loss. Essentially, influenza dynamics are a balance between large perturbations following significant antigenic jumps, interspersed with nonlinear epidemic dynamics tuned by climatic forcing.

The modeling is based on a high quality eleven-year dataset from June 2001 to January 2013 of daily ILI (influenza like illness) cases reported in Israel's largest city Tel Aviv (see Fig.2) where some 45% of the local population is covered by medical surveillance. Laboratory tests of ILI cases from sentinel clinics have shown the ILI dataset to be highly correlated with incidence of confirmed influenza cases. The high level of coverage and the data quality makes this one of the finest available influenza surveillance datasets by world standards.

They make use of the classical Susceptible-Infected-Recovered-Susceptible (SIRS) epidemic model where all individuals in the population are assumed to be in only one of three classes: Susceptible (S), Infected (I) and Recovered (R). As is usual, S individuals move to the I class after contact with an infective and transmission of the disease. Infective individuals eventually recover and move to the R class, while R individual eventually become susceptible once more closing the SIRS loop in a manner that mimics the effects of antigenic drift .

The model requires six basic parameters which we provide with the best fitting estimates found from the modelling procedure: i) The parameter lambda describes rate of loss-of-immunity due to antigenic drift and the rate of new susceptible individuals entering the population. ii) The reproduction number R_0 represents the number of infections transmitted by a typical individual over the lifetime of the disease in a wholly susceptible population. iii&iv) Seasonal forcing is included by modulating R_0 with the locally observed temperature and relative humidity time series (see Fig.2). The forcing delta(t) is specified by weights that control the influence of these climatic variables. v& vi) Two separate parameters for the epochal antigenic jumps are required to accommodate the early outbreak in 2003-4 from the effects of the new A/Fujian strain. Finally, the model requires only a single boundary condition, S(0) which represents the size of the susceptible population on the 1st of June 2001 at the beginning of the modeling period.

The simplicity of this elementary SIRS model would deem it unlikely to mimic Israel's complex influenza dynamics. Despite this, the model fit of the first eight years of the data (June 2001-June 2008) correlates with the observed data to an unusual degree (r=0.94) as shown in Fig.3. Since there are only six degrees of freedom, and one boundary condition, this is not even one parameter per year and we stress that a longer time series would not require more parameters. Also note that the large year-to-year variability in epidemic size and timing appears to be random in character, making it surprising that this mechanistic model can capture the intricacies of the data despite the major qualitative differences in the shape of the epidemic from year-to-year.

For example, compare the large scale symmetrically shaped epidemic in 2004-5 with the small curtailed asymmetric outbreak in 2005-6. The model nevertheless captures both epidemics accurately. This is due to the complex interaction between the changing supply of new susceptible individuals arising due to loss of immunity in the population through antigenic drift, the strong transient dynamics following the appearance of a new strain and the timing of the climate signal each year.

Note that the years dominated by influenza B fall effortlessly within the model dynamics, which provides possible population-level observational support for the notion of broader cross-reactivity of the immune response to influenza, as previously reported.

The parameters obtained from fitting the first six years were used to run the model forwards in an attempt to predict the ILI cases over the following four years (June 2007 - June 2010). As shown in Fig.3, the model output closely traces most outbreaks beginning from June 1st 2008 with high fidelity, and without any refitting of parameters.

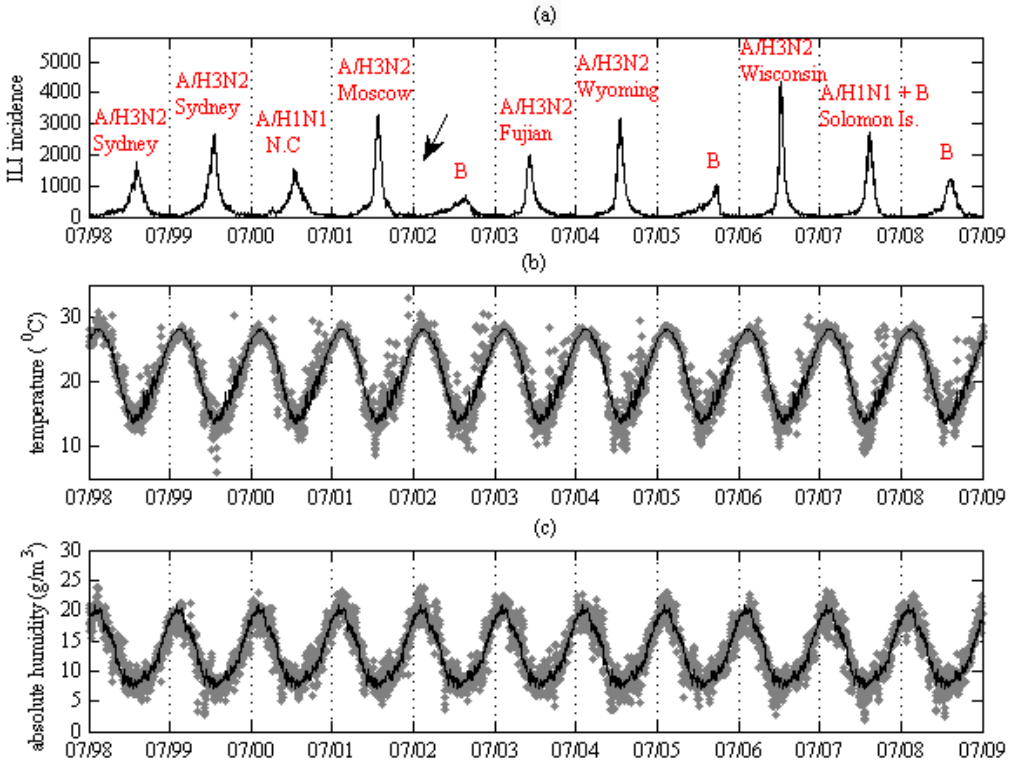

Fig.2 The smoothed daily ILI cases reported in Tel Aviv (bottom panel, light grey) with June 1st indicated on the x-axis for each year. Time series from the climate-driven deterministic SIRS model fitted to the ILI data (2001-2010) is shown in the bottom panel (red), with fit correlation r=0.94 to the period leading up to the pandemic in 2009. The climate data consists of centered, normalized daily temperature (top, dark blue) and relative humidity readings (top, light violet), both scaled.

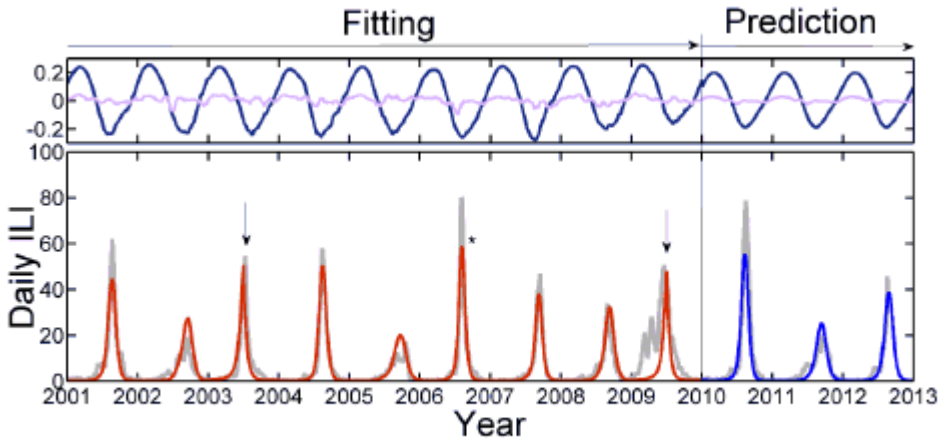

Fig.3 The extrapolation period from 1st June 2007 to 1st June 2010 is the model output (bottom, blue) driven solely by temperature and relative humidity (bottom panel, dark green). The pure prediction from 1st June 2010 is driven by the average climate of all years with fit r=0.93. The outbreaks dominated by influenza B are indicated, and the asterisk highlights the abnormally high amplitude of 2006-2007. The epochal jumps in antigenic drift are indicated with arrows.

3. Influenza modelling and prediction in the UK

The London School of Hygiene and Tropical Medicine has also been heavily involved in influenza modeling and prediction (Baguelin et al. 2013). They made use of a complex age and risk group specific mathematical model of flu which they fitted to multiple data sources on ILI and flu in the UK available over the period 1995/6 to 2008/9. A novel Monte Carlo Markov Chain procedure was developed and used for the purpose

The LSHTM group took the modeling a step further than predictions. They attempted to examine different vaccination policies. As the epidemiology of flu in the UK has been disturbed by vaccination for many years, the first step was to use the fitted model to reconstruct flu dynamics in the absence of any vaccination. From this point it was possible to use the model to predict forward what would have happened if different alternative vaccination policies were applied. The outcomes of this model were influential in reshaping UK policy in that as of September 2013, children will now be vaccinated rather than chiefly the elderly over-65 age group in the population.

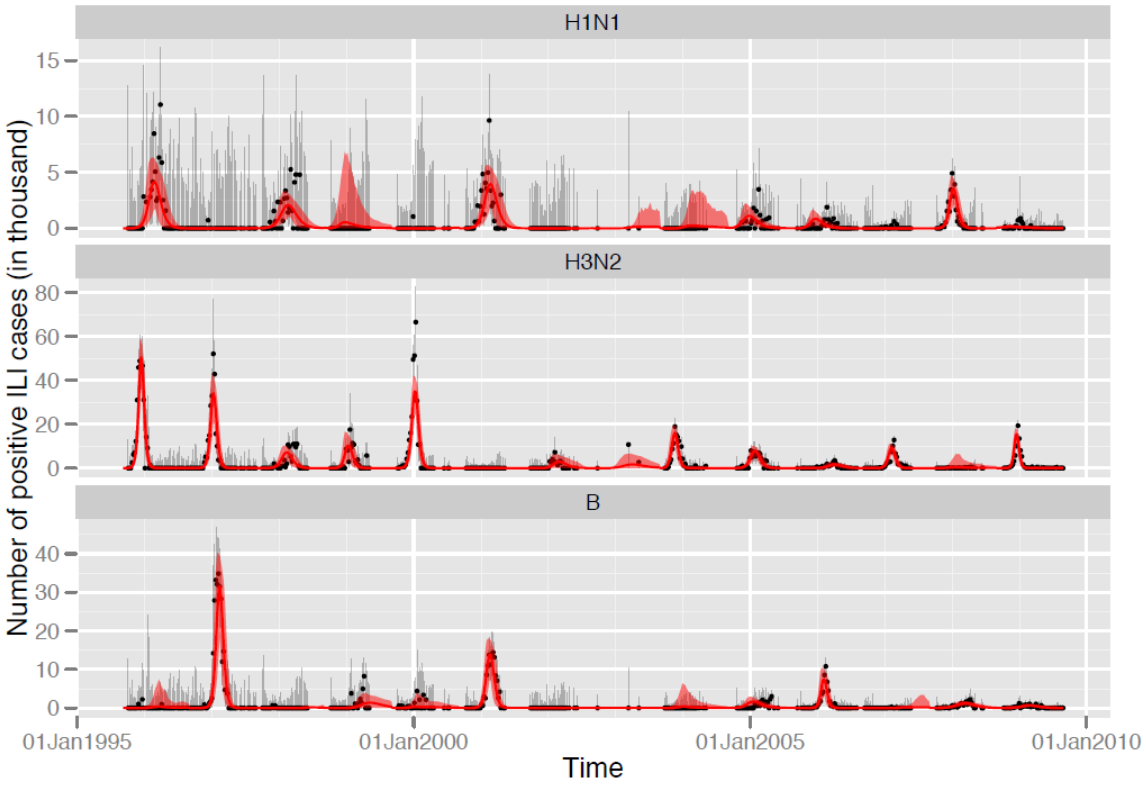

Fig.4 The model fit is compared to the data on overall incidence (not by age group) for the different strains over the 14 seasons

3. Refining model fitting and estimation of parameters for other diseases: pneumococcus, tuberculosis and malaria

i) Streptococcus pneumoniae (pneumococcus)

Day care centre attendees play a central role in maintaining the circulation of Streptococcus pneumoniae (pneumococcus) in the population. The prevalence of pneumococcal carriage is highest in DCC attendees but varies across countries and is found to be consistently lower in Finland than in Portugal. Longitudinal data about serotype-specific carriage in day care centre attendees in Portugal and Finland were analysed by the IGC group with a continuous-time event history model in a Bayesian framework. The monthly rates of within-room transmission, community acquisition and clearing carriage were estimated and interpreted (Figure 5), and the observed levels of carriage prevalence and longitudinal patterns in carriage acquisition and clearance were adequately predicted (Pessoa et al. 2013). The model was further developed to describe colonization dynamics by pneumococcus before and after vaccination through an SIS framework. Fitting these models to data from Portugal (2001-2007), using single and multiple carriage prevalence, the IGC group estimates transmission and competition coefficients and vaccine efficacy (Gjini et al 2013). This study raises fundamental questions about the sensitivity of established estimation procedures to the adopted model framework.

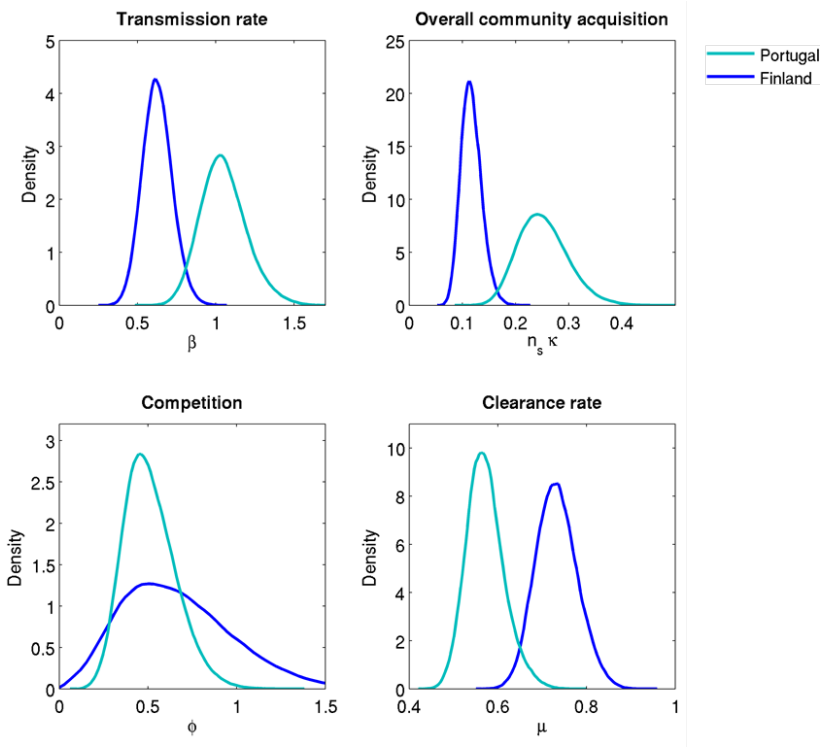

Figure 5. Posterior distributions of the pneumococcus transmission parameters. The rate parameters are presented per month.

ii) Tubersculosis

Two competing mathematical model of tuberculosis transmission were fitted by the IGC group to data from 14 molecular epidemiology studies, enabling the selection of the best performing model and the estimation of relevant epidemiological parameters. The IGC group used Gauss-Newton algorithm to fit the data and F-test to compare the models. Model selection indicated that heterogeneity is an important component of tuberculosis transmission and should be represented in models (Gomes et al. 2012). The model was subsequently applied to more detailed data to estimate the real TB burden in Portugal, indicating that tuberculosis transmission is in a close vicinity of the reinfection threshold and should, therefore, present relatively high sensitivity to control measures (Lopes et al. 2013).

iii) Malaria

A mathematical model was constructed by the IGC group to investigate the effects of vector control campaigns against malaria. The model was fitted to data, consisting of parasite prevalence and specific antibodies, collected in Principe Island, West Africa, before and during a vector control campaign (Bandeiras et al. 2013). The model illustrates well the consequences of loss of immunity due to reduced transmission in this setting and can be applied to other regions where similar data may be available.

Dose-response models are widely used to relate parasite loads with risk of infection for the host. They can also be used to estimate heterogeneity in host-susceptibility to infection. The IGC group applied these models to survival curves of insect hosts challenged with different viral titers. By replicating the procedure in populations of insects that are carriers or non-carriers of the symbiont bacterium Wolbachia, the IGC group estimates the distribution of protection conferred by Wolbachia against viral infection among their insect hosts as illustrated in Figure 6 (Pessoa et al 2013b). The methods developed by the IGC group are applicable to the estimation of parameters relevant to the efficacy of Wolbachia releases as a control measure against vector borne diseases.

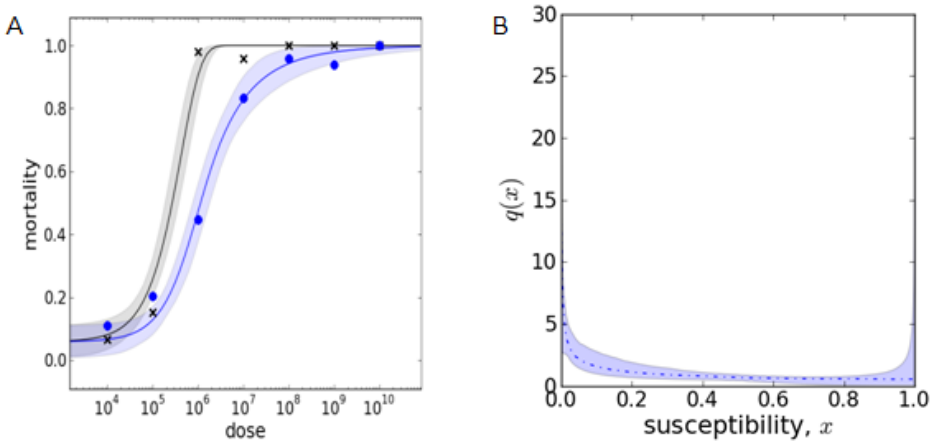

Figure 6. Fitting of dose-response model to mortality data. Wolb$^-$ and Wolb$^+$ are represented by black and blue, respectively. The shaded areas represent the 95% CI from bootstrap. A, Points are observed mortalities, and curves represent the probability of death conditional to challenge dose according to the best fitting model. B, Susceptibility distribution associated to the best fitting dose-response model.

Conclusion

The partners in WP1 have developed new sophisticated approaches for modelling influenza based on time series surveillance data. The partners gained from exchanging ideas between themselves at Epiwork review meetings (and elsewhere) and this resulted in the different groups addressing similar challenges but in their own unique ways with their own different datasets. Deliverable 1.5 has generated a set of exciting new tools for predicting epidemic dynamics that go well beyond the current state-of-the-art.

These tools integrate into the larger Epiwork project. The IGC group led the way by modelling surveillance data from Epiwork's flu internet surveillance. Epiwork internet surveillance was also an essential factor that made possible the work of the LSHT group. Finally the prediction tools of all groups have applications in other components of the EPIWORK project and can be tested on simulation data of the GLEAM VIZ computational platform.

Epiwork publications produced as part of Deliverable 1.5.

Gulbenkian Institute:

Coelho FC, Codeço CT, Gomes MGM, A Bayesian framework for parameter estimation in dynamical models. PLoS ONE 6(5): e19616. (2011) [Epiwork 30%]

van Noort SP, Aguas R, Ballesteros S, Gomes MGM, The role of weather on the relation between influenza and influenza-like illness. J Theor Biol 298: 131-137. (2012) [Epiwork 30%]

Pessoa D, Hoti F, Syrjänen R, Sá-Leão R, Tarja Kaijalainen, Gomes MGM, Auranen K, Comparative analysis of Streptococcus pneumoniae transmission in Portuguese and Finnish day-care centers. BMC Infectious Diseases 13: 180. (2013) [Epiwork 30%]

Gjini E, Valente C, Sá-Leão R, Gomes MGM (2013) Pneumococcal colonization, vaccination and symmetry between serotypes (submitted). [Epiwork 30%]

Gomes MGM, Aguas R, Lopes JS, Nunes MC, Rebelo C, Rodrigues P, Struchiner CJ, How host selection governs tuberculosis reinfection. Proc R Soc Lond B 279: 2473-2478. (2012) [Epiwork 30%]

Lopes JS, Rodrigues P, Fonseca-Antunes A, Gomes MGM, Interpreting simple measures of tuberculosis transmission: A case study on the Portuguese population (in preparation). (2013) [Epiwork 30%]

Bandeiras C, Trovoada MJ, Gonçalves LA, Marinho CRF, Turner L, Hviid L, Penha-Gonçalves C, Gomes MGM, Modeling malaria infection and immunity against variant surface antigens in Principe Island, West Africa (in preparation). (2013) [Epiwork 30%]

Pessoa D, Souto-Maior C, Lopes JS, Gjini E, Codeço CT, Ceña B, Marialva M, Zwerschke D, Teixeira L, Gomes MGM, Inferring distributions of antiviral protection conferred by the Wolbachia symbiont in Drosophila melanogaster (submitted). (2013b) [Epiwork 30%]

Tel Aviv University:

Huppert A., Barnea, O., Katriel, G., Yaari, R, Roll, U. & Stone, L. Modeling and Statistical Analysis of the Spatio-Temporal Patterns of Seasonal Influenza in Israel. PL0S ONE 7(10): e45107 (2012) [Epiwork 60%]

Barnea, Huppert, Katriel, Stone. Spatio-temporal synchrony of influenza in cities across Israel: The "Israel is One City" hypothesis. (submitted). [Epiwork 60%]

Axelson J, Yaari R., Grenfell B., Stone L. Multiannual forecasting of seasonal influenza dynamics reveals climatic and evolutionary drivers. PNAS (under revision). [Epiwork 60%]

Yaari R, Katriel G, Huppert A, Axelsen JB, Stone L , Modelling seasonal influenza : the role of weather and punctuated antigenic drift. Journal of The Royal Society Interface 10. doi:10.1098/ rsif.2013.0298 (2013) [Epiwork 60%]

Katriel, G. Stochastic discrete-time age-of-infection epidemic models Int. J. Biomath., 06 (2013), 1250066 [Epiwork 60%]

LSHTM group
Baguelin, Camacho, Edmunds et al (submitted)