Deliverable

D1.1 Automated tools for rapid assessment of key transmission parameters and rates of spread estimates

Version: 1

Due: Month 24

Completed: Month 24

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Contents

Deliverable Description

D1.1 Automated tools for rapid assessment of key transmission parameters and rates of spread estimates

The aim of Deliverable 1.1 was to search and catalogue tools for the rapid assessment of key transmission parameters and rate of spread estimates. Model-fitting in real time has proven potential to inform risk assessment and policy responses to outbreaks, provided there is rapid access to epidemiological and genome sequence data (e.g. Volkova et al. 2011). Data inputs include spatio-temporal distribution of cases (and non-cases), contact tracing and individual risk factors. Literature describing models is available for various infections, for instance foot and mouth disease, avian and human influenza, MRSA, but formal validation in an epidemic situation is an iterative process that takes time. Implementing these models for real-time outbreak use in the appropriate software (e.g. specific programming language or software program) for further use by others is often a job that requires specialist skills in programming and/or mathematics and can take up valuable time. Especially in an outbreak situation it is of the essence that key transmission parameters and (initial) rates of spread are estimated as quickly as possible to assess the potential of that outbreak. Estimation of the basic reproduction rate (R_0), for example, might help with a rapid assessment of the transmission potential of a disease. In general, for an epidemic to occur in a susceptible population R_0 must be > 1. Thus, if an initial estimate clearly shows a R_0 much larger than 1, immediate action can (and should) be taken to combat the disease and prevent further spread.

Thus, it is clearly highly beneficial when tools that can be used for rapid assessment of key parameters and estimating rates of spread are readily available when needed. In order to find out which tools (models / frameworks) are available for both COMPARE project partners and others, an online search was performed to identify relevant open-source models or modelling frameworks that can be readily used (or adapted) to investigate an outbreak situation. A catalogue was created of these models/frameworks and specific information on the models was extracted from the model/framework documentation to help identify the correct model for a specific situation.

The resulting catalogue consists of a total of 21 open-source models with a description of the models, link to the source code or website and key information about the model that is available. The results are bundled in a datasheet and made available on the DTU COMPARE Sharepoint site to all project partners. Furthermore, the catalogue will be made available online for worldwide access in order to promote “open-science”. Additional features like the possibility to amend existing information and add new models/frameworks are planned for future releases of the online version of the catalogue.

In conclusion, the obtained catalogue of existing open-source models / frameworks cannot only contribute to other WPs within the COMPARE project but also outside the COMPARE project by making the catalogue available online.

Methods

The online search was conducted in three separate search engines:

- Google (and Google Scholar)
- MEDLINE
- Github.com

To narrow down the results to open-source, readily available frameworks and a combination of different terms were used to obtain as many results as possible. The results were then manually reviewed to extract findings that matched the criteria (i.e. an open-source readily available model or model framework). In total, 21 models or model frameworks were identified and added to the catalogue. In addition, more detailed information on the models was subsequently identified from the documentation or from the source-code and added to the catalogue.

Results

Table A1 in the Annex I Section shows the results from the online search for open-source modeling frameworks.

Conclusions

In conclusion, the catalogue of existing open-source models / frameworks created in this Work Package will contribute to other WPs within the COMPARE project but also outside the COMPARE project by making the catalogue available online. Furthermore, it will be made publicly available with the intention to add tools to update existing data and add new models or model frameworks when they become available to build a useful repository of modelling software that can be used rapidly in the event of a disease outbreak.

References:

Volkova VV, Bessell PR, Woolhouse MEJ et al. Evaluation of risks of foot-and-mouth disease in Scotland to assist with decision making during the 2007 outbreak in the UK. *Vet Rec* 2011; **169**: 124-U41.

Annex I. Catalogue of open-source models/model frameworks

Model / Framework Title	URL of Source Code / Project page	Description	Language Used	Model Type	Intended Use	(Meta-)Data required	Outputs	Notes
FRED (Framework for Reconstructing Epidemiological Dynamics)	http://fred.publichealth.pitt.edu/	FRED supports research on the dynamics of infectious disease epidemics and the interacting effects of: mitigation strategies, viral evolution, and personal health behaviour. The system uses agent-based modelling based on census-based synthetic populations that capture the demographic and geographic distributions of the population, as well as detailed household, school, and workplace social networks. Multiple circulating and evolving strains can be simulated. Mitigation strategies in the framework include vaccination, anti-viral drugs, and school closure policies. FRED supports models of health behaviour change to facilitate the study of critical personal health behaviours such as vaccine acceptance, personal hygiene and spontaneous social distancing. FRED is available through open source in the hopes of making large-scale agent-based epidemic models more useful to the policy-making community, the research community, and as a teaching tool for students in public health.	C++	Agent-based model	Policy-making community, the research community, and as a teaching tool for students in public health	Synthetic population data if the model is used for country (countries) outside the USA. Disease specific parameters.	Incidence, Susceptible, Exposed, Infectious Symptomatic, Immune and Recovered Agents, Number of schools closed, Disease ID, Attack rate and Reproductive rate	census based synthetic populations for USA included

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

STEM Project (Spatiotemporal Epidemiological Modeller)	http://www.eclipse.org/stem/index.php	STEM is designed to make it easy for developers and researchers to plug in their choice of models. It comes with a large number of existing compartment models and a new model building framework that allows users to rapidly extend existing models or to create entirely new models. The model building framework provides a simple graphical users interface and automatically generates all of the model code and hot injects the code into STEM at runtime. In many cases, no knowledge of Eclipse or Java is required. The STEM code generator even allows users to build models affected by changes in climate data.	Java	Compartmental model building framework	Model building framework.	Dependent on the model that is implemented in STEM. Can range from simple parameter estimates and initial values to transport networks and spatial maps.	Dependent on the model that is implemented in STEM. Can range from simple output of (fraction of) infected / recovered to animated time series of spatial maps with the selected information.	Comes with large number of existing models.
EpiFire	https://github.com/tjhladish/EpiFire/wiki	EpiFire is a fast C++ applications programming interface for simulating the spread of epidemics on contact networks	C++	Contact network model		Description of the network (although a network can be generated). Transmission rate, number of index cases, infectious period.	Node status in time, epidemic curves, epidemic size, expected R_0. Basic network analysis.	
Generic EbolaResponse (ER) : modelling the spread of disease impact & interven	http://stacks.cdc.gov/view/cdc/24900	Spreadsheet-based model that (1) allows user to estimate the number of Ebola Virus Disease (EVD) cases in a community, (2) tracks patients` susceptibility to disease through infectivity, incubation, recovery, and death, and (3) calculates the spread of EVD and its impact for 300 days. The timeline of 300 days is longer than most of the previously recorded outbreaks of EVD or Marburg Virus Disease.	Microsoft Excel	Susceptible, Incubation, Infectious, Recovery (SIIR) Model	Ebola outbreak modelling	Population size, number of index cases, Infectious period, Incubation distribution	Daily clinical cases, Daily hospital admission, Daily beds-in-use, Cumulative Ebola cases	Ebola specific

tion.								
Global epidemic and mobility model (GLEAM)	http://www.gleamviz.org/	The global epidemic and mobility model, GLEAM, combines real-world data on populations and human mobility with elaborate stochastic models of disease transmission to deliver analytic and forecasting power to address the challenges faced in developing intervention strategies that minimize the impact of potentially devastating epidemics.	Java	stochastic transmission model, grid based model	Pandemic modelling	Disease specific parameter values	Maps showing spread of infection in time, influence of airport networks	Many datasets build-in (i.e. population, airports, flight patterns, commuting networks, etc.)
Zika modelling	http://www.zika-model.org/index.html	Data-driven global stochastic and spatial vector-host epidemic model to provide a quantitative analysis of the spreading pattern of ZIKV in Latin America.	GLEAM (Java)	stochastic , vector-host epidemic model	ZIKV outbreak modelling	See GLEAM	See GLEAM	Zika Virus specific.
Epi Info	https://www.cdc.gov/epiinfo/index.html	Epi Info™ is a public domain suite of interoperable software tools designed for the global community of public health practitioners and researchers. It provides for easy data entry form and database construction, a customized data entry experience, and data analyses with epidemiologic statistics, maps, and graphs for public health professionals who may lack an information technology background. Epi Info™ is used for outbreak	Unknown (Source not available?)	statistical modelling, outbreak investigation	Data entry, epidemiological statistics	Surveillance data	Statistical reports, maps & graphs for public health professionals	

		investigations; for developing small to mid-sized disease surveillance systems; as analysis, visualization, and reporting (AVR) components of larger systems; and in the continuing education in the science of epidemiology and public health analytic methods at schools of public health around the world.						
FlexDx Xpert Scale Up	https://github.com/JJPennington/FlexDx-Xpert-Scale-Up	FlexDx is a tool that allows non-expert users to define their local situation according to four key parameters: TB incidence, proportion of new TB cases that are multidrug-resistant (MDR), adult human immunodeficiency virus (HIV) prevalence, and per-patient cost of first-line therapy. This tool then links a decision analysis describing nine different TB diagnostic strategies to a transmission model that provides five-year estimates of TB incidence, mortality, and control costs in any user-specified scenario.	Python	Compartmental model specific for tuberculosis	TB outbreak modelling	See accompanying publications for full description and data requirements: https://elifesciences.org/content/3/e02565	See accompanying publications for full description and possible outputs: https://elifesciences.org/content/3/e02565	Tuberculosis specific
InfluSim	https://sourceforge.net/projects/influsim/	A deterministic pandemic influenza simulator. Computes the effect of interventions like antiviral treatment of cases and social distancing. To be used for pandemic preparedness planning by health care offices.	Java	Deterministic pandemic influenza model	Influenza modelling	Age distributions, contact matrix, children contacts, Hospital parameters, Disease parameters (Infectious periods, Severity, Case fatality), Treatments resources, Social distancing parameters	Infection dynamics, Costs, Resource use.	Influenza specific

Sispread (Simulation of Infectious Spreading)	http://sispread.sourceforge.net/	Sispread allows to simulate the dynamics of a hypothetic infectious disease within a contact network of connected people. It is intended to help people concerned by public health to easily perform epidemic simulations and to analyse their results.	c/c++	Contact network model	public health	Type of model (SI, SIS, SIR), transmission rate, recovery rate, index cases	Network files, outbreak data files	only generated networks possible
EpiModel	http://www.epimodel.org/	EpiModel is an R package that provides tools for simulating and analysing mathematical models of infectious disease. Supported epidemic model classes include deterministic compartmental models, stochastic individual contact models, and stochastic network models. Disease types include SI, SIR, and SIS epidemics with and without demography, with utilities available for expansion to construct and simulate epidemic models of arbitrary complexity. The network model class is based on the statistical framework of temporal exponential random graph models (ERGMs) implemented in the Statnet suite of software for R.	R	Deterministic compartmental models, stochastic individual contact models, stochastic network models		Parameters specific for used model class, initial conditions	Simulation data. I.e. compartment values and flow sizes for each simulation	
Vimes	http://github.com/thibautjombart/vimes	Visualisation and Monitoring of Epidemics, including some outbreak detection algorithms.	R	Visualisation and monitoring tool	Match cases based on different data types (i.e. location, time of onset, genetic sequences)	Different sources of data that should be matched.	Clusters of similar cases (graphs)	

Outbreaker2	http://github.com/thibautjombart/outbreaker2	Inferring transmission chains by integrating epidemiological and genetic data.	R	Tool to integrate epidemiological and genetic data to infer transmission chains	Minimum: DNA-sequences (in DNAbin format), collection dates and generation time distribution. Optional: case distance matrix	transmission tree ("who infected whom"); dates of infection; mutation rate per generation of infection; proportion of the outbreak sampled; effective reproduction numbers over time or at an individual level; incidence curves; mutation rate per unit of time.
EpiEstim	http://github.com/annecori/EpiEstim	Quantifying transmissibility throughout an epidemic from incidence time series.	R	Estimate key parameters from incidence time series	Incidence data	reproduction number R
Incidence	http://github.com/reconhub/incidence	Computation, handling, visualisation and simple modelling of incidence.	R	Incidence modelling	Incidence data	Various showing incidence data, results of model fitting
Surveillance project	http://surveillance.r-forge.r-project.org/	To provide open source software for the temporal and spatio-temporal visualization, modelling and monitoring of epidemic phenomena. This includes count, binary and categorical data time series as well as continuous-time processes having discrete or continuous spatial resolution.	R	Models for epidemic phenomena	Surveillance data	Various options for visualisation, likelihood inference and simulation

Generic Disease Modelling System	http://www.globalburden.org https://github.com/aflaxman/gbd	This is the soon-to-be-finalized version of DisMod III, the Generic Disease Modelling System. It was developed for the Institute of Health Metrics and Evaluation at the University of Washington, as be part of the Global Burden of Disease 2010 Study.	Python	descriptive epidemiological meta-regression tool	Prevalence data, including: Region, Country, GBD Cause , Year Start, Year End, Parameter : one of 'Prevalence', 'Incidence', 'Remission', 'Mortality', 'RR' (relative risk), 'SMR' (standardized mortality ratio), 'With Condition Mortality', 'Duration', 'cause-specific mortality', or 'Excess mortality', Age, Sex	Various measures of burden of disease	Package still under development. Documentation is sparse
GLOBAL EPIDEMIC MODEL (GEM)	http://www.epimodels.org/drupal/?q=node/12	The Global Epidemic Model (GEM) uses airline travel data and population data from 2000 to create a realistic air travel network among the major metropolitan areas of the world. The purpose of GEM is to model the time course of a possible epidemic as it spreads across the globe from an initial small outbreak in one city. GEM also allows for testing of various disease intervention strategies.	Java	pandemic models	Disease specific parameter values (infectious period, contact rate, etc.)	Simulation data. Interactive views of worldwide spread of the modelled disease	Does not seem to be actively developed anymore

Transtat	http://www.epimodels.org/drupal/?q=node/14	The key to controlling a pandemic is early detection, containment, and mitigation. The TranStat tool was developed to enable field personnel and researchers to enter and revise data from local outbreaks. From these data TranStat provides a means of testing for the presence of human-to-human (or animal-to-animal) transmission. If this transmission is detected, estimates of the household-specific and neighbourhood-specific secondary attack rates and local reproductive number are provided.	Java	Resampling-based tests to detect person-to-person transmission of infectious disease		Data on symptom onsets and cluster (i.e. household) information	likelihood of person-person transmission, secondary attack rates, local reproductive numbers
ABM++ FRAMEWORK	http://www.epimodels.org/drupal/?q=node/11	The ABM++ software framework is a tool which allows the developer to implement agent based models using C++ that are to be deployed on distributed memory Linux clusters. The framework provides the necessary functionality to allow applications to run on distributed architectures. A C++ message passing API is provided which provides the ability to send MPI messages between distributed objects. The framework also provides an interface that allows objects to be serialized into message buffers, allowing them to be moved between distributed compute nodes. A synchronization method is provided, and both time-stepped and distributed discrete event time update mechanisms are provided.	C++	Agent based model, synthetic social network model	Low level framework to implement agent based models to be deployed on linux clusters	Dependent on the implemented model	Dependent on the implemented model