Swiss TPH
Swiss Tropical and Public Health Institute
Schweizerisches Tropen- und Public Health-Institut
Institut Tropical et de Santé Publique Suisse

Biostatistics & Computational Sciences
Dept. Epidemiology & Public Health

malariacontrol.net: status update

Nicolas Maire

BOINC Workshop, August 31th 2010

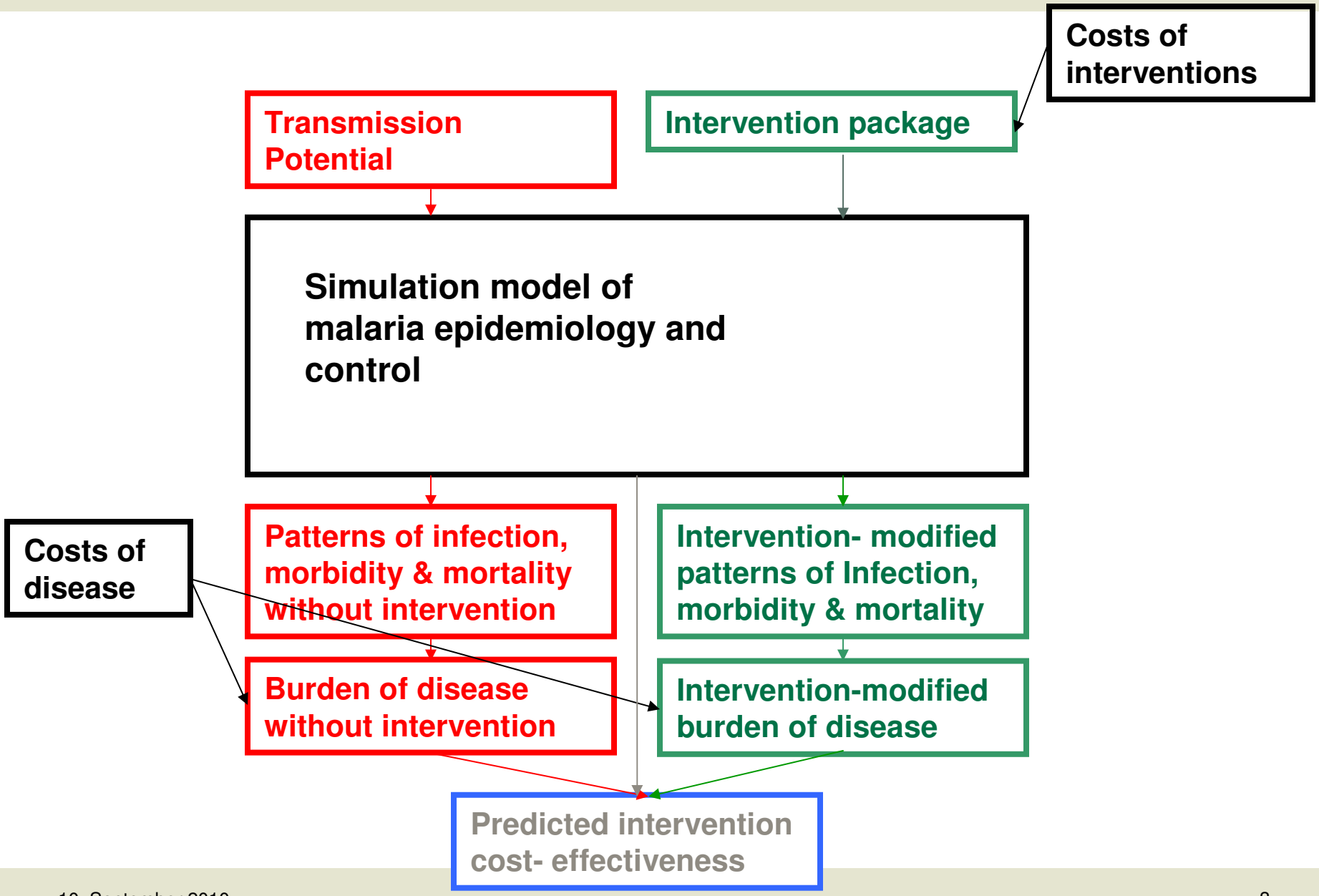
malariacontrol.net:

Computing resource for malaria modeling projects at the Swiss Tropical and Public Health Institute since 2005

July 2003: Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of *Plasmodium falciparum* malaria

June 2006: Simulation modeling of the epidemiological impact and cost-effectiveness of malaria interventions

March 2009: A stochastic simulation platform for predicting the effects of different malaria intervention strategies


Costs of interventions
Transmission Potential
Intervention package
Simulation model of malaria epidemiology and control
Costs of disease
Patterns of infection, morbidity & mortality without intervention
Intervention- modified patterns of Infection, morbidity & mortality
Burden of disease without intervention
Intervention-modified burden of disease
Predicted intervention cost- effectiveness

Development of the modeling project since 2003

Malaria has received more attention in recent years

Diversification of research questions

- From pre-erythrocytic vaccines to integrated control programs
- From malaria control to elimination
- More emphasis on uncertainty analysis
- Taking into account new scientific findings

Development of the modeling project since 2003

Model development and implementation

Interdisciplinary research team

Increasing number of end-users

Increasing demand for computing power

Current Research Team

Applied Mathematics
Melissa Penny (Swiss TPH)
Nakul Chitnis (Swiss TPH)

Epidem./Public Health
Allan Schapira (Swiss TPH)
Blaise Genton (Swiss TPH)
Don de Savigny (Swiss TPH)
Marcel Tanner (Swiss TPH)

Quantitative biology
Ian Hastings (LSTM)
Katherine Winter (LSTM)
Michael Bretscher (Swiss TPH)

Databases
Konstantina Boutsika (Swiss TPH)

Statistics
Amanda Ross (Swiss TPH)
Tom Smith (Swiss TPH)

Computer Science
Diggory Hardy (Swiss TPH)
Aurelio di Pasquale (Swiss TPH)
Guillaume Gnaegi (Swiss TPH)
Nicolas Maire (Swiss TPH)
Tiago Antão (LSTM)
Henning Mortveit (VBI)

Health Economics
Fabrizio Tediosi (Milan)
Josh Yukich (Swiss TPH)
Lesong Conteh (LSHTM)
Valerie Crowell (Swiss TPH)

Financial support
Bill & Melinda Gates Foundation, PATH-MACEPA,
Swiss National Science Foundation

Recent focus (computing)

Simulator software design and extension

Parameter estimation

Computing platform for design and analysis of simulation experiments

Component modules of the model 0 of natural history

Acute illness

An episode of acute morbidity occurs in individual i, at time t, with probability

$$P_m(i,t) = \frac{Y_{\max}(i,t)}{Y^*(i,t) + Y_{\max}(i,t)} \quad (22)$$

where Y^* is the pyrogenic threshold and $Y_{\max}$ is the maximum density of five daily densities sampled during the five-day time interval t. The pyrogenic threshold evolves over time via:

$$\frac{dY^*(i,t)}{dt} = \frac{\alpha Y(i,t)}{(Y_1^* + Y(i,t))(Y_2^* + Y^*(i,t))} - \varpi Y^*(i,t) \quad (23)$$

with the initial condition $Y^*(i,0) = Y_0^*$ at the birth of the host and α, ϖ, Y_1^*, and Y_2^* are constants.

Parasite densities

Each new infection j, initiated in individual i at time t_0 is assigned a duration of $\tau_{\max}$, sampled from

$$\ln(\tau_{\max}(i,j)) \sim \text{Normal}(5.13, 0.80) \quad (5)$$

The log density in the absence of previous exposure at each time point, $\tau = 0,1,\ldots,\tau_{\max}(i,j)$ of the infection j in host i is then normally distributed with expectation

$$\ln(y_0(i,j,\tau)) = \ln d(i) + \ln(y_G(\tau, \tau_{\max})) \quad (6)$$

where $y_G(\tau,\tau_{\max})$ is an empirical description of malariatherapy patients from the Georgia hospital and $d(i)$ represents between-host variation drawn from a log-normal distribution with variance σ_y^2.

We measure exposure to asexual blood stages with

$$X_y(i,j,t) = \int_{t-a}^{t} Y(i,\tau)\, d\tau - \int_{t_{0,j}}^{t} y(i,j,\tau) \quad d\tau \quad (7)$$

where $Y(i,\tau)$ is the total parasite density of individual i at time τ and $y(i,j,\tau,)$ is the density in individual i for infection j at time τ, and

$$X_h(i,t) = \int_{t-a}^{t} h(i,\tau) \quad d\tau - 1. \quad (8)$$

the expected log density for each concurrent infection is then

$$E(\ln(y(i,j,\tau))) = D_y D_h D_m \cdot \ln(y_0(i,j,\tau)) + \ln\left(\frac{D_x}{M(t)} + 1 - D_x\right)$$

where $M(t)$ is the total multiplicity of infection and

$$D_y = \frac{1}{1 + \frac{X_y(i,j,t)}{X_y^*}} \quad , \quad (10)$$

$$D_h = \frac{1}{1 + \frac{X_y(i,t)}{X_h^*}} \quad , \quad (11)$$

$$D_m = 1 - \alpha_m \exp\left(-\frac{0.693a}{a_m^*}\right) \quad (12)$$

and X_y^*, X_h^*, D_x, a_m^*, and α_m, are further constants.

Variation within individual hosts is quantified by a term $\sigma_y^2(i,j,\tau)$, where

$$\sigma_y^2(i,j,\tau) = \frac{\sigma_0^2}{1 + \frac{X_h(i,t)}{X_v^*}} \quad (13)$$

and σ_0^2 and X_v^* are constants (Table 1). The simulated densities are specified using:

$$\ln(y(i,j,\tau)) \sim \text{Normal}(E(\ln(y(i,j,\tau))), \sigma_y^2(i,j,\tau)) \quad (14)$$

Infection of mosquitoes

Let

$$\Upsilon(i,t) = \beta_1 Y(i,t-2) + \beta_2 Y(i,t-3) + \beta_3 Y(i,t-4) \quad (16)$$

where t is in 5-day units, and

$$\ln(y_g(i,t)) \sim \text{Normal}(\ln(\rho\Upsilon(i,t)), \sigma_g^2) \quad (17)$$

where $\beta_1,\beta_2,\beta_3,\rho,\sigma_g^2$ are constants (Table 1). Define

$$\Pr(y_g(i,t) > y_g^*) = \Phi\left[\frac{\ln(\rho\Upsilon(i,t)) - \ln(y_g^*)}{\sigma_g}\right] = \Phi\left[\frac{\ln(\Upsilon(i,t))}{\sigma_g} + \rho^*\right] \quad (18)$$

where Φ is the cumulative normal distribution, y_g^* is the density of female gametocytes necessary for infection of the mosquito, and $\rho^* = (\ln(\rho) - \ln(y_g^*))/\sigma_g$. Then the proportion of mosquitoes that are infected feeding on individual i at time t is

$$I_m(i,t) = [\Pr(y_g(i,t) > y_g^*)]^2 \quad (19)$$

and the probability that a mosquito becomes infected at any feed is:

$$\kappa_u(t) = \eta \frac{\sum_i (A(a(i,t))I_m(i,t))}{\sum_i A(a(i,t))} \quad (20)$$

where η is a constant scale factor.

Define $\kappa_u^{(0)}(t)$ as the value of $\kappa_u(t)$ in the simulation of an equilibrium scenario to which an intervention has been applied. Let $E_{\max}^{(0)}(t + l_v)$ be the corresponding entomologic inoculation rate. $\kappa_u^{(1)}(t)$ and $E_{\max}^{(1)}(t + l_v)$ are the corresponding values for the intervention scenario. Then

$$E_{\max}^{(1)}(t + l_v) = \frac{E_{\max}^{(0)}(t + l_v)\, \kappa_u^{(1)}(t)}{\kappa_u^{(0)}(t)} \quad (21)$$

where l_v corresponds to the duration of the sporogonic cycle in the vector, which we approximate with two time steps (10 days). $(E_{\max}^{(0)}(t + l_v)/\kappa_u^{(0)}(t)$ is the total vectorial capacity).

Infection of humans

$E_a(i,t)$, the age-adjusted entomologic inoculation rate (EIR) for individual i at time t, is given by

$$E_a(i,t) = E_{\max}(t) \frac{A(a(i,t))}{A_{\max}} \quad (1)$$

where, $A(a(i,t))$ is the average body surface area estimated for an individual of age $a(i,t)$ and $A_{\max}$ is the average surface area of people ≥ 20 years of age in the same population. $E_{\max}(t)$ refers to the usual measure of the EIR computed from human bait collections. The force of infection is then

$$\lambda(i,t) = E_a(i,t)\left(S_\infty + \frac{1 - S_\infty}{1 + \frac{E_a(i,t)}{E^*}}\right)\left(S_{imm} + \frac{1 - S_{imm}}{1 + \left(\frac{X_\rho(i,t)}{X_\rho^*}\right)^{\gamma_\rho}}\right) \quad (2)$$

where S_{imm}, X_ρ^*, E^*, γ_ρ, S_∞ are constants (Table 1) and:

$$X_\rho(i,t) = \int_{t-a(i,t)}^{t} E_a(i,\tau)d\tau. \quad (3)$$

The number of infections $h(i,t)$ introduced in time step t, is distributed as

$$h(i,t) \sim \text{Poisson}(\lambda(i,t)) \quad (4)$$

Severe disease

We consider two different classes of severe episodes, B_1 and B_2. $P_{B1}(i,t)$ is the probability that an acute episode (A) is a class B_1 severe episode and is specified using

$$P_{B_1}(i,t) = \Pr(H(i,t) \in B_1 | H(i,t) \in A) = \frac{Y_{\max}(i,t)}{Y_{B_1}^* + Y_{\max}(i,t)} \quad (24)$$

where $Y_{B_1}^*$ is a constant and $H(i,t)$ is the clinical status.

The second subset of severe malaria episodes (B_2) occur when an otherwise uncomplicated malaria episode happens to coincide with some other insult, which occurs with risk

$$F(a(i,t)) = \frac{F_0}{1 + \left(\frac{a(i,t)}{a_F^*}\right)} \quad (25)$$

where F_0 is the limiting value of $F(a(i,t))$ at birth, and a_F^* is the age at which it is halved.

The probability that an episode belonging to class B_2 occurs at time t, conditional on there being a clinical episode at that time is $P_{B2}(i,t)$ where

$$P_{B_2}(i,t) = \Pr(H(i,t) \in B_2 | H(i,t) \in A) = F(a(i,t)) \quad (26)$$

The age and time specific risk of severe malaria morbidity conditional on a clinical episode is then given by

$$P_B(i,t) = P_{B_1}(i,t) + P_{B_2}(i,t) - P_{B_1}(i,t)P_{B_2}(i,t), \quad (27)$$

Mortality

Malaria deaths in hospital are a random sample of those severe malaria cases deemed to be admitted, with age-dependent sampling fraction $Q_h(a)$, the hospital case fatality rate, derived from the data of Reyburn and others.[86]

We estimate the severe malaria case fatality in the community, $Q_c(a)$ for age group a with

$$Q_c(a) = \frac{Q_h(a)\varphi_1}{1 - Q_h(a) + Q_h(a)\varphi_1}, \quad (28)$$

where φ_1, the estimated odds ratio for death in the community compared to death in in-patients, is an age-independent constant and $Q_h(a)$ is the hospital case fatality rate. Malaria mortality is the sum of the hospital and community malaria deaths.

The risk of neonatal mortality attributable to malaria (death in class D_1) in first pregnancies is set equal to $0.3\mu_{PG}$ where μ_{PG} is given by

$$\mu_{PG} = \mu_{\max}\left[1 - \exp\left(-\frac{x_{PG}}{x_{PG}^*}\right)\right], \quad (29)$$

where x_{PG} is related to x_{MG}, the prevalence in simulated individuals 20–24 years of age via

$$x_{PG} = 1 - \frac{1}{1 + \left(\frac{x_{MG}}{x_{MG}^*}\right)} \quad (30)$$

and x_{MG}^* and x_{PG}^* are constants (Table 1).

An indirect death in class D_2 is provoked at time t, conditional on there being a clinical episode at that time, with probability $P_{D2}(i,t)$ where

$P_{D_2}(i,t) = \Pr(H(i,t) \epsilon D_2 | H(i,t) \epsilon A)$ and

$$P_{D_2}(i,t) = \frac{Q_D}{1 + \left(\frac{a(i,t)}{a_F^*}\right)} \quad (31)$$

where Q_D is limiting value of $P_{D2}(i,t)$ at birth and a_F^* is a constant. Deaths in class D_2 occur 30 days (six time steps) after the provoking episodes.

Model uncertainty

How wrong is this model

Does it matter?

Plug-in model components based on different assumptions

Towards model ensembles

Original implementation

Core model in Fortran

Somewhat modular implementation of sub-models

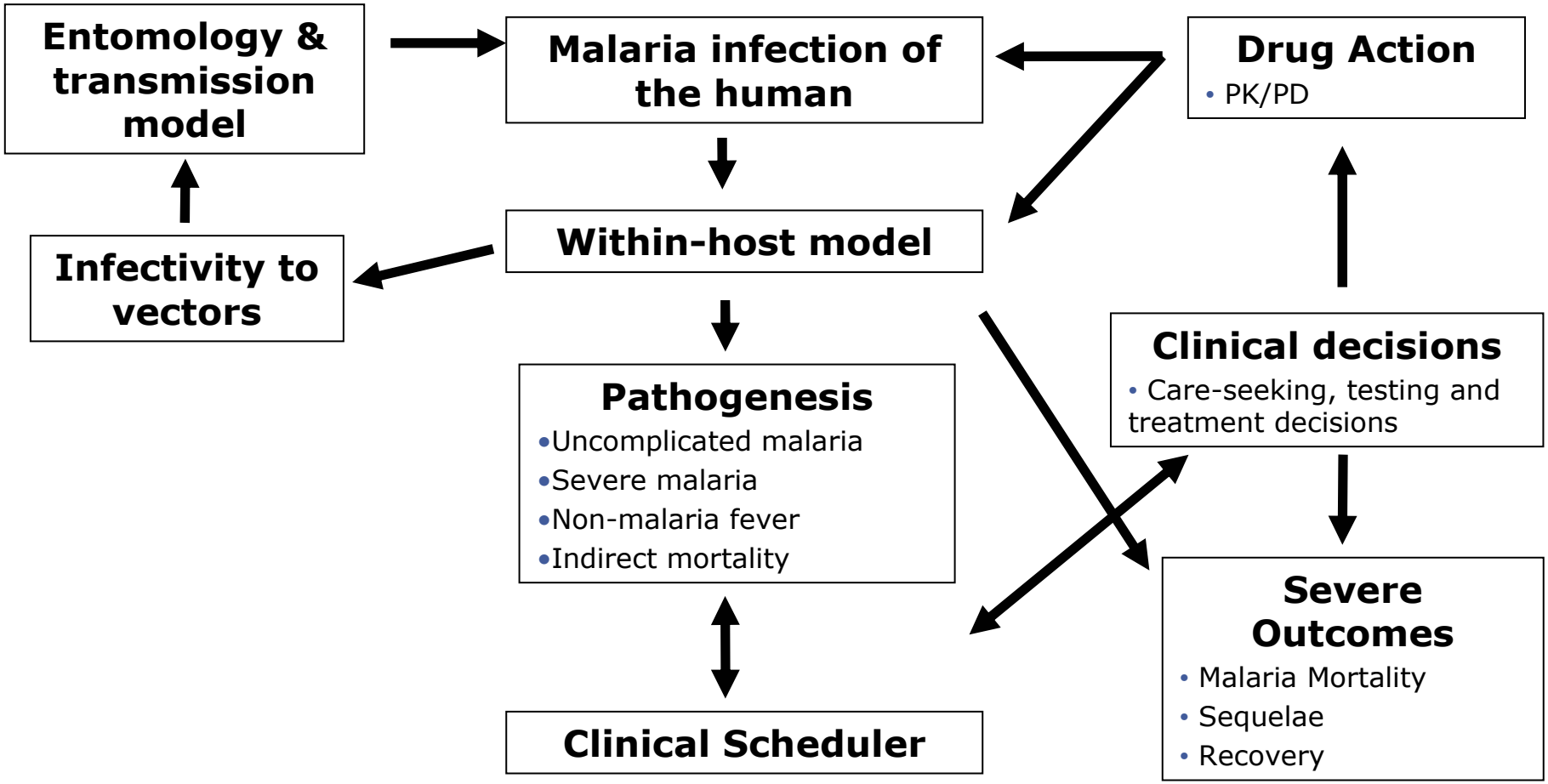
Increasingly difficult to extend and maintain

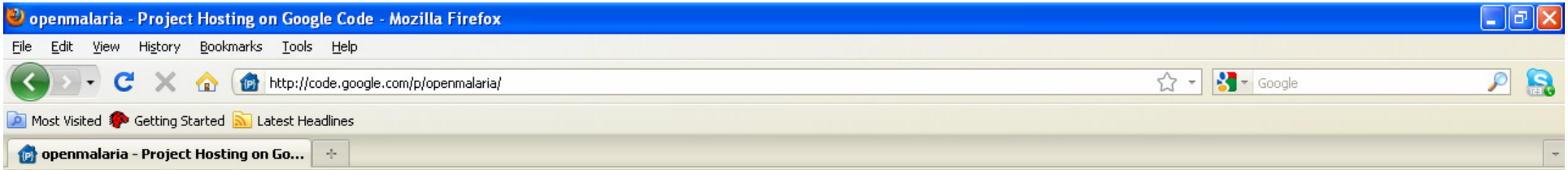
Port to C++

Definition of clean interfaces between modules

Extension with alternative model components

Module overview of simulator





openmalaria

A simulator of malaria epidemology and control

Search projects

Project Home | Downloads | Wiki | Issues | Source

Summary | Updates | People

Mathematical models are important tools for decision making in the control of infectious diseases, and malaria was one of the first infections for which such modeling was applied. However, there is still an urgent need for new models that can compare the potential impact of a comprehensive range of malaria interventions. To address this need we have developed a platform for stochastic simulations of malaria infections, nested within simulations of individuals in human populations.

The simulations of malaria infections are linked to models of interventions and health systems, epidemiology to predict the impacts of interventions on infection, morbidity, mortality, health services use and costs. We use numerous field datasets to optimise parameter estimates. By using a volunteer computing system we obtain the enormous computational power required for model fitting, sensitivity analysis, and exploration of many different intervention strategies.

The project provides a general platform for comparing, fitting, and evaluating different model structures, and for quantitative prediction of effects of different interventions and integrated control programs.

Star this project

Activity: High

Code license:
GNU General Public License v2

Labels:
malaria, epidemiology, cpp

Feeds:
Project feeds

Groups:
General discussion

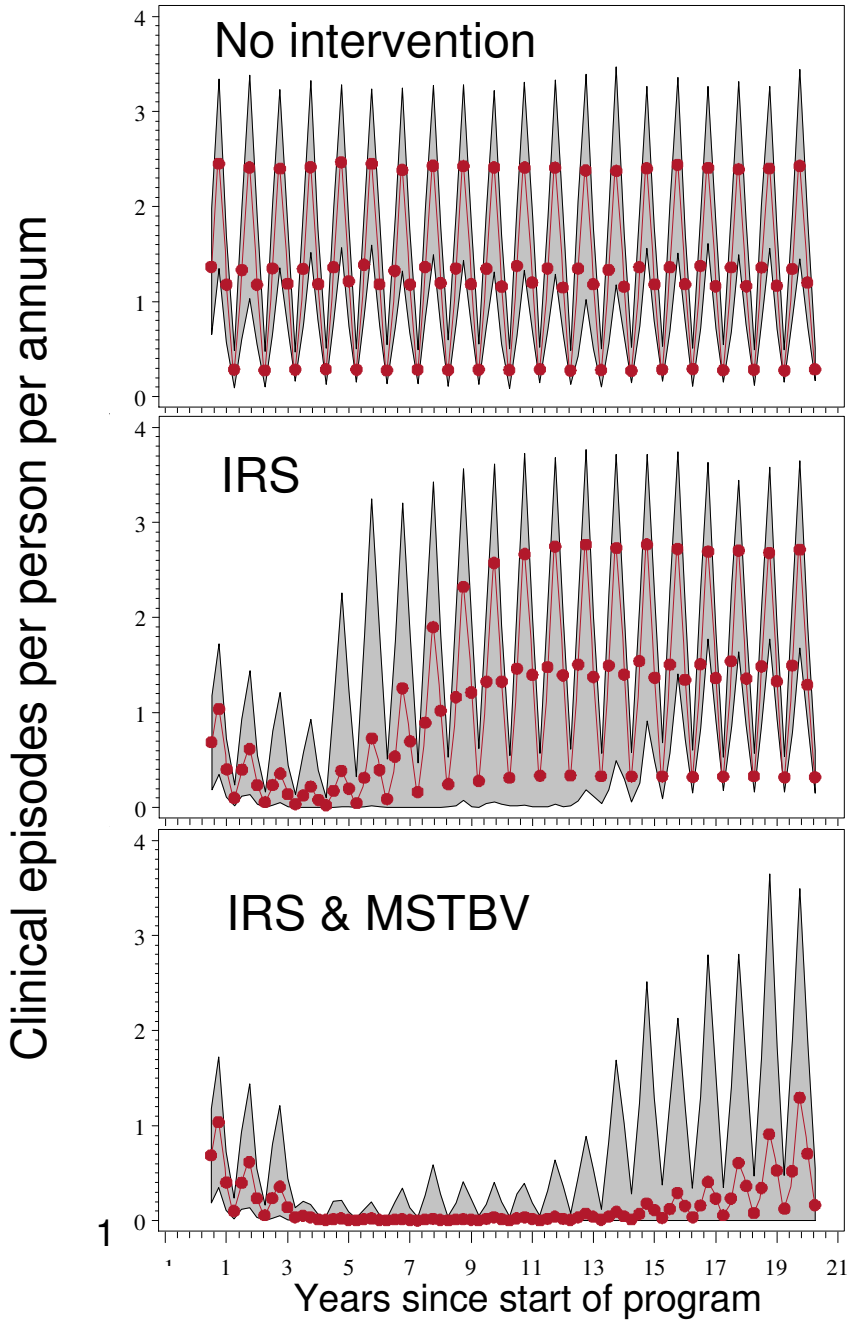
Owners:
tiagoantao, nicolas.maire, diggory.hardy, Guillaume.Gnaegi

Committers:
amanda.ross114, nakul7, thomas-A.Smith@unibas.ch, melissa.code, vccrowell, hastings@liverpool.ac.uk, kwinter@liverpool.ac.uk, Henning.Mortveit, aurdipas

Contributors:
erin.stuckey

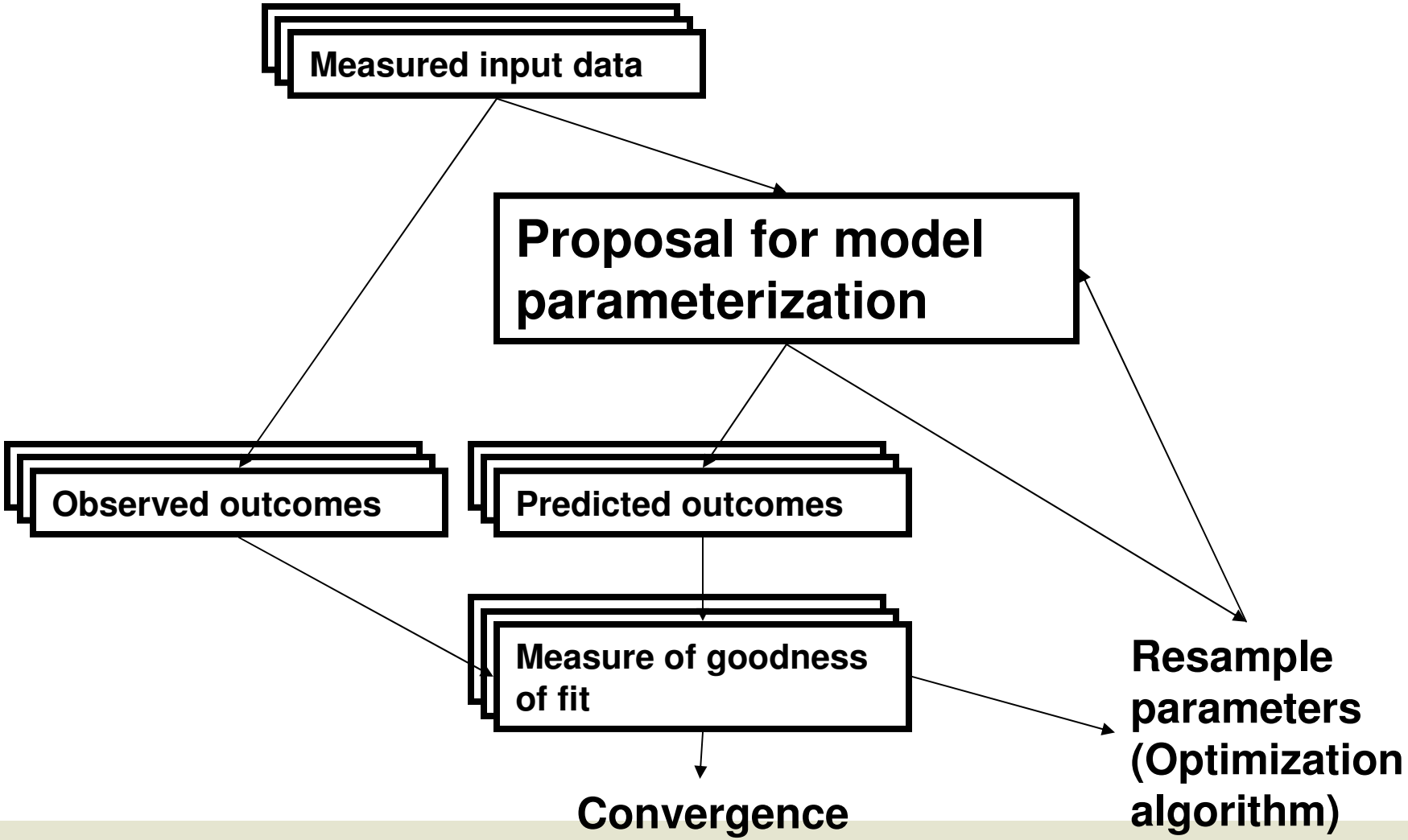
People details »

Combination of mosquito stage transmission blocking vaccines with indoor residual spraying



- Simulated population of 1000 people
  - 15 models/parameterisations
  - 10 seeds for each
- IRS,
  - 95% coverage,
  - 1 round each of 1st 3 years
- MSTBV,
  - 90% efficacy,
  - 10 year half-life of effect,
  - 95% coverage,
  - 1 dose at year 3.
- Imported infections:
  - 2.9/1000 hosts per annum
  - evenly spread over the year

Estimating model parameters from field data



Challenge of fitting these models

Objective functions non-differentiable

Loss function values are not reproducible because of stochasticity

High-dimensional parameter space

Computationally expensive

Takes a long time

Potential to reduce time-to-convergence (wall time)?

Characteristics of the volunteer computing platform

- Latency/Unpredictability/Validation overhead
- Return on in investment in scale-up

Choice of optimization algorithm

- Recently more interest in massively parallel asynchronous optimization

BOINC scheduler improvements

Reliable host scheduling

- Significantly reduces latency

Adaptive replication for validation

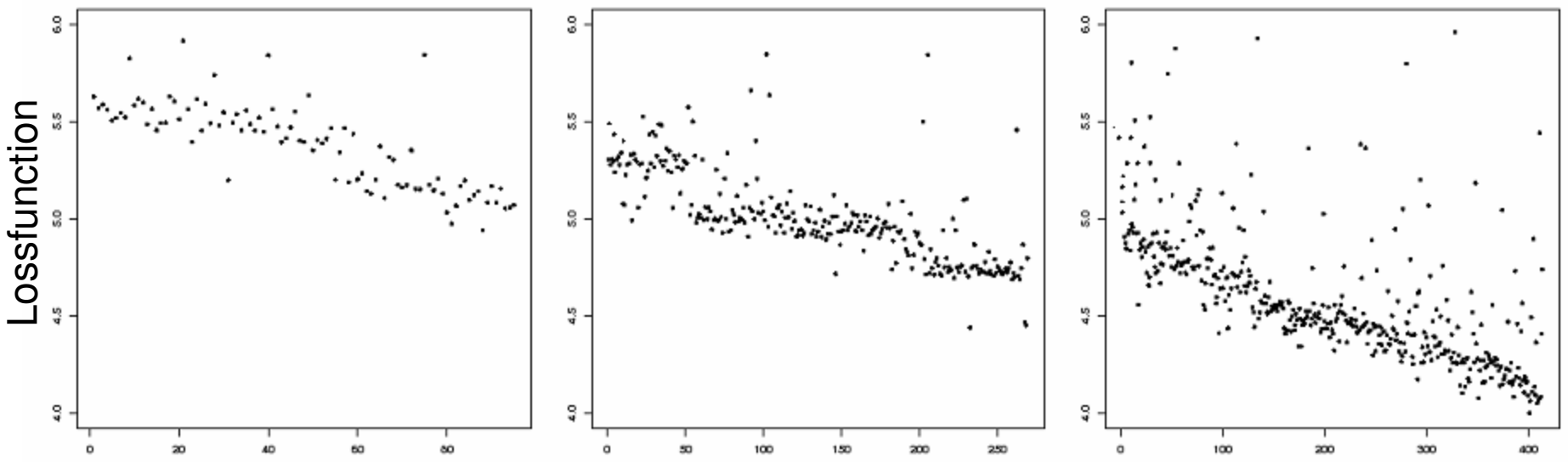
- Significantly increases throughput

Convergence rate (wall time) by investement

18 samples/eval

36 samples/eval

72 samples/eval



Active recruitment of volunteers

Progress Through Processors

Facebook application developed by gridrepublic.org, sponsored by Intel

15‘000 new users

Alternative optimization algorithms

Travis Desell et al., milkyway@home:

Asynchronous Global Optimization for Massive-Scale Computing. PhD thesis, Rensselaer Polytechnic Institute, December 2009

Outlook

Performance optimiziation

- Memory
- GPUs/Multicore

Computing platform/User interface