NOV 22, 2023

Protocol status: Working
We use this protocol and it's working

🌐 EpiPopSynth

km.zhu[1], Kang Liu[1], Junli Liu[2], Yepeng Shi[1], Xuan Li[1], Hongyang Zou[3], Huibin Du[3], Ling Yin[1]

[1]Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China;
[2]Hangzhou Institute of Technology, Xidian University, Hangzhou, China;
[3]College of Management and Economics, Tianjin University, Tianjin 300072, China

Ling Yin: Corresponding author;

 km.zhu

ABSTRACT

Agent-based models have gained traction in exploring the intricate processes governing the spread of infectious diseases, particularly due to their proficiency in capturing nonlinear interaction dynamics. The fidelity of agent-based models in replicating real-world epidemic scenarios hinges on the accurate portrayal of both population-wide and individual-level interactions. In situations where comprehensive population data are lacking, synthetic populations serve as a vital input to agent-based models, approximating real-world demographic structures. While some current population synthesizers consider the structural relationships among agents from the same household, there remains room for refinement in this domain, which could potentially introduce biases in subsequent disease transmission simulations. In response, this study unveils a novel methodology for generating synthetic populations tailored for infectious disease transmission simulations. By integrating insights from microsample-derived household structures, we employ a heuristic combinatorial optimizer to recalibrate these structures, subsequently yielding synthetic populations that faithfully represent agent structural relationships. Implementing this technique, we successfully generated a spatially-explicit synthetic population encompassing over 17 million agents for Shenzhen, China. The findings affirm the method's efficacy in delineating the inherent statistical structural relationship patterns, aligning well with demographic benchmarks at both city and subzone tiers. Moreover, when assessed against a stochastic agent-based Susceptible-Exposed-Infectious-Recovered model, our results pinpointed that variations in population synthesizers can notably alter epidemic projections, influencing both the peak incidence rate and its onset.

Created: Nov 22, 2023

Last Modified: Nov 22, 2023

PROTOCOL integer ID: 91327

Keywords: population synthesis, agent-based model, epidemic model

GUIDELINES

plese just follow the steps.

MATERIALS

High-performance computing environments with more than 500 CPU cores, the MPI (Message Passing Interface) framework is recommended.

SAFETY WARNINGS

> ⛔ There is no particularly dangerous operation in this experiment.

Population Generation

1 Process micro household survey data, re-group individuals into age groups, and remove irrelevant fields and non-family households.

2 Recode members with the same household ID into family structure strings to generate a pool of family structures.

3 Conduct a logistic regression test on the distribution of family structures in the pool, determine the number of family structures needed to cover a given proportion α of the population, and select the top k types of families as family motifs.

Combinatorial Optimization

4 Initialize the combinatorial optimization iterator with the distribution of motifs in the pool as the initial guess, with the decision variable being a k-dimensional vector.

5 Use the number of households/people of different household sizes, age groups, and genders in regional demographic data as the least squares optimization target function, optimize based on the trf algorithm, and output the optimal decision variable upon reaching the termination condition.

6 Generate the synthetic population within the region based on the optimal decision variable values as the proportion of each family motif.

Population Generation

7 Execute steps 4-6 in parallel to generate synthetic populations for multiple subzones within the city and merge them into the final synthetic population.

Epidemic Simulation

8 Run the infectious disease agent-based model using the above synthetic population as the carrier and test the results.