ABSTRACT

The recent increase in interest in Population modeling has brought up the need to define the term. Rather than a formal definition, we are collectively defining the term by population modeling examples collected from all the authors. Examples include epidemiology, Behavioral, health economics, emergency response, biology, and computational tools. A formal definition is also discussed to provide a current definition for an emerging field.

Author Keywords
Population Modeling, Definition, Multi Disciplinary.

ACM Classification Keywords
H.1 Models And Principles; J.3 Life And Medical Sciences.

INTRODUCTION

This paper originated from a discussion held at the population modeling working group at the Multistage Modeling Consortium and Interagency Modeling and Analysis Group (MSM/IMAG) meeting held at the NIH in September 2014 [1]. There was increased interest in the group, leading to the following group definition:

"Modeling a collection of entities with different levels of heterogeneity"

This is an ad hoc, quick consensus definition that fit the group that met at the NIH. However, we continued and tried to seek a wider definition. We created a SimTk portal [2] with a mailing list [3] and invited contributions of work. The raw discussions from which the paper is based on can be found in the mailing list archives [4]. The paper will reiterate the examples contributed in submission order to attempt and define the field.

EXAMPLES CONTRIBUTED:

Olaf Dammann and Benjamin Hescott, Tufts

The Tufts Population Modeling group is interested in occurrence patterns of developmental disabilities and their risk factors [5]. The idea is to simulate disease occurrence in virtual populations that can be compared to published data. The approach is named “systems epidemiology” [6]. It involves observation in populations and population modeling in order to complement systems biology, which involves biological experimentation and computational modeling of individual entities. One current focus is the question: What respective roles do oxygen exposure and neonatal infection play in the etiology of retinopathy of prematurity [7]?

Sergey Nuzhdin, University of Southern California

Social interactions can affect group size and composition, and conversely, group size and composition can affect social interactions among individuals. Individuals within societies differ from one another; for example in their likelihood of associating with, or attacking other individuals; and if they are attacked themselves, they may differ in how they adjust their own behavior based on that experience. Feedbacks between behavior at the level of individuals and behavior at the level of groups and societies must be understood in order to predict the behaviors and their key health outcomes at each of these levels. [8-10].

Jacob Barhak


1 The Authors of this manuscript are part of the population modeling mailing list and in order of contribution they are: Olaf Dammann, Tufts, USA; Sergey Nuzhdin, University of Southern California, USA; Jacob Barhak, Austin, USA; Atesmachew B Hailegiorgis, George Mason University, USA; Shweta Bansal, Georgetown University, USA; H. Stephen Leff, Human; Services Research Institute, USA; Joshua G. Behr, Old Dominion University, USA; C. Anthony Hunt, University of California San Francisco, USA; Talitha Feenstra, University of Groningen, The Netherlands; Madhav Marathe, Virginia Tech, USA; Mary Butler, University of Minnesota, USA; Bradley Davidson, University of Denver, USA; Paul Marjoram, University of Southern California, USA; Stefan Scholz, University of Bielefeld, Germany; Jonathan Karnon, University of Adelaide, Australia; Aaron Garrett, Jacksonville State University, USA; Wojciech (Al) Chrosny, TreeAge Software, USA; Samarth Swarup, Virginia Tech, USA; Naren Ramakrishnan, Virginia Tech, USA; Cristina Lanzas, North Carolina State University, USA; Amiyaal Ilany, University of Pennsylvania, USA; John Rice, Society for Simulation in Healthcare, USA.
associated with individual patient data, and therefore allows access to more modeling data. The Micro Simulation Tool (MIST) [12] is free software developed to support population modeling capabilities using Evolutionary Computation and High Performance Computing.

A spatial agent based model is used to explore the climate driven outbreak of cholera in refugee camps. The interaction between humans (host) and their environment is modeled and the spread of the epidemic using Susceptible-Exposed-Infected-Recovered (SEIR) model. Results show that seasonal rains caused the emergence of cholera outbreaks, and more importantly agents’ social behavior and movements aggravated the spread of cholera to other camps where water sources were relatively safe [13].

Shweta Bansal, Georgetown University
The field of network epidemiology is a branch of infectious disease modeling which focuses on disease-independent heterogeneity in host contact rates [14]. Incorporation of individual-level contact heterogeneity in population modeling of infectious disease spread has led to an understanding of super-spreading phenomenon [15], of the preferential impact of past epidemics on future disease dynamics [16], and the design of targeted intervention studies that can effectively control disease outbreaks [17]. The field of network epidemiology advanced in recent years, however, many challenges still remain [18].

H. Stephen Leff, Human Services Research Institute
Planning by the Numbers (PBN) [19] is a planning and resource allocation model. The model is a web-based Markov simulation. Inputs include the population to be served, services desired, service unit costs, and predicted outcomes. The model is typically used for budget planning, settling right to treatment suits, planning to resize hospitals, and planning jail diversion programs. A comprehensive description of modeling work can be found in [20]. And [21] holds a description of a model application relevant to the current state of mental health.

Joshua G. Behr, Old Dominion University
The sheltering and evacuation "decision calculus" of individual household members when facing an impending severe storm event [22] is mapped. The factors (and their associated weights) are identified with a 'basket' of factors that individual households employ when making sheltering and evacuation decisions. Also identified are the factors and decision processes involved in choosing health care treatment venues such as emergency departments, primary care, and safety net health providers.

C. Anthony Hunt, University of California San Francisco
Hunt lab is developing M&S methods capable of representing and explaining the considerable intra- and inter-individual variability that characterize health and treatment related phenomena such as that resulting from drug-induced liver injury [23] and that observed for some but not other individuals in bioavailability of generic drug products [24]. The methods used are agent-oriented. Explanatory power is improved by making mechanisms modular [25] and imposing a strong parsimony guideline. A consequence is that mechanisms are no more fine-grained than is needed to achieve validation targets [26].

Talitha Feenstra, University of Groningen
Modeling at the Centre for Nutrition, Prevention and Health Services Research of the Dutch National Institute for Public Health and the Environment (RIVM) concerns supporting public health policy. The RIVM chronic disease model models the Dutch population, specific to age and gender. It describes the aggregate relation between several lifestyle risk factors (smoking, drinking, food intake), intermediate outcomes (blood pressure, cholesterol, BMI) and a range of chronic diseases. Outcomes include morbidity, mortality, QALYs, and costs [27-30]. Extensions exist for Chronic Obstructive Pulmonary Disease and Diabetes. The DYNAMO-HIA model [31] can be freely downloaded and is adaptable by using different sets of input data.

Madhav Marathe, Virginia Tech
The Network Dynamics and Simulation Science Laboratory (NDSSL), is a part of the Virginia Bio-informatics Institute at Virginia Tech [32]. The modeling approach is agnostic of specific populations including human populations, animal populations, cells and wireless devices. Four large bodies of work are: (i) science of networks, (ii) public health epidemiology and (iii) disaster resilience and (iv) computational immunology. See [33] for description of current work to support the Ebola response efforts. See [34] for applications, especially SIV and Granite. SIV is a visual analytics tool to visualize synthetic populations. A synthetic population for the entire US was created and being extended to the globe. Granite is a web based system to analyze large networks.

Mary Butler, University of Minnesota
The University of Minnesota's School of Public Health, holds an interdisciplinary group meeting under a "big data' label. This group grapples with issues of using health care data to answer population health questions - estimating risk factors, comparative effectiveness research, treatment heterogeneity - and how to structure the data. Additional focus included how to collect/make available more meaningful patient outcomes, and when to avoid the same biases that exist with current data sources [35].

Bradley Davidson, University of Denver
The goal is to simulate population-based randomized controlled trials (RCTs) with realistic treatment effects using efficient probabilistic techniques. Recently a probabilistic “wrapper” for OpenSim [36] was created to perform musculoskeletal simulations that account for
uncertainty and variability from multiple sources. The probabilistic interface uses traditional Monte Carlo simulations [37,38] and more efficient and innovative methods. Recent advances analyze the effects of experimental error (marker placement, movement artifacts) and parameter uncertainty (e.g. body segment parameters, muscle parameters) within patient-specific simulations [39].

Paul Marjoram, University of Southern California
The research focuses on multi-scale modeling of genetic variation in developmental networks in Drosophila [40]. Modeling investigates how populations of cells in the Drosophila embryo interact to produce patterns of gene expression that are important to development of the embryo. Additional research interests are modeling how populations of cells interact in growing tumors (the application being to colon cancer) [41,42]. Also of interest is agent-based modeling of animal behavior, exploring how variation within the population can affect behavior, or the robustness of behavior to external perturbation.

Stefan Scholz, University of Bielefeld
The SILAS-model [43], aims to simulate Sexual Infections as Large-scale Agent-based Simulation. SILAS is a demographic model close to the level of the general German population. The model is built in the FLAME-framework [44]. Each agent in SILAS calculates probability distributions in dependence to its characteristics (age, sex, sexual orientation, etc.). The behavior rules are estimated from a large panel-data set using the GAMLSS-package [45] in R.

Jonathan Karnon, University of Adelaide
Discrete Event Simulation (DES) was applied to glaucoma services at a public hospital. The effects of a range of alternative clinical pathways were evaluated (e.g. earlier use of laser in the treatment) and amendments to the organization of the glaucoma service (e.g. changing the duration of the booking cycle). Outcome improvement options were identified across the population at minimal additional cost [46]. 2) DES was used to calibrate cancer surveillance models. A simple breast cancer progression model was developed. Individuals in different prognostic groups were simulated, replicating the observed frequency and timing of surveillance. costs and QALYs of alternative surveillance strategies were evaluated [47].

Aaron Garrett, Jacksonville State University
Inspyred [48,49] is a Python library for computational intelligence/evolutionary computation. This is a basic tool that can aid population modelers. Other modeling work is, simulating evacuees from a structure when trying to optimize the egress locations for safety and timeliness [50].

Wojciech (Al) Chrosny, TreeAge Software
Recent work included comparison of discrete event simulation methods and Markov individual patient simulation methods. Some preliminary results of the comparison were presented at [51].

Samarth Swarup, Virginia Tech
The aftermath of a nuclear detonation was simulated with 730,000 agents, modeling transportation, communication, health, and power infrastructures. Disaster resilience results showed that relatively passive interventions like quickly partially restoring communication could have a significant effect on lives saved [52,53]. In a separate flu epidemic study, an existing synthetic population of Washington DC was augmented with a population of transients. Results showed that implementing a location-specific intervention, such as encouraging healthy behaviors (covering your cough, using hand sanitizer, etc), can have a significant impact on reducing the epidemic [54].

Naren Ramakrishnan, Virginia Tech
Recent work focus was on developing models for forecasting population-level events, e.g., disease outbreaks, civil unrest, elections. The IARPA OSI project aims to use open source information (news, blogs, tweets, and economic indicators) to develop algorithms that can identify precursors and surrogates for events, and model their progression. Examples of work includes epidemiology, civil unrest [55-57].

Cristina Lanzas, North Carolina State University
The focus is on the epidemiology and ecology of infectious diseases in animal and human populations. Data, epidemiological analysis and mathematical models are combined to study transmission mechanisms. The emphasis is on the role that environment plays on transmission and the dissemination of antimicrobial resistant pathogens [58]. Models that capture more realistic exposure patterns and include spatial features of the pathogen transmission are required [59]. Currently mathematical models used to assess environmental transmission are being improved [60].

Amiyaal Ilany, University of Pennsylvania
Research focuses on social networks and on principles of animal communication. Concepts and analytical tools integrate biology, sociology, and network science. Social network analysis provides metrics to quantify social structure at different levels of organization. Social interactions in a wild rock hyrax population are studied. A general agent-based model demonstrates how social stability is achieved when cooperation is practiced in cohesive clusters of individuals. [61-63]

DISCUSSION AND TECHNICALITIES
The examples provided describe population models applied to multiple fields: behavioral, biology, epidemiology, health economics, and emergency response. Some modeling tools were introduced. A discussion followed and is publicly accessible in [4]. The main issues addressed are:
Population modeling for things other than humans including cells, forests, wireless devices, animals etc.

What is the place of cohort models such as Markov Models within population modeling that ignore heterogeneity?

The group will address these issues in the future to provide a better definition for Population Modeling.

This paper is a cumulative effort of all contributors who responded to the population modeler call. Each contributor sent text to the mailing list. Jacob Barhak assembled and edited the paper. Readers are welcome to read the actual archives and join this discussion at the mailing list [3].

REFERENCES

1. IMAG Population Modeling Working Group, Online: http://www.imagwiki.nibib.nih.gov/content/population-modeling-working-group
2. SimTk: Population Modeling Workgroup Project, Online: https://simtk.org/home/popmodwkgrpimag
4. The PopModWkGrpIMAG-news Archives, Online: https://simtk.org/pipermail/popmodwkgrpimag-news/
19. Planning by the Numbers Free, Online: http://model.planningbythenumbers.org/


27. RIVM Chronic disease model, Online, http://www.nationaalkompas.nl/algemeen/meta-informatie/modellen/cdn/


of-an-innovative-shared-diabetes-care-model-methods/

31. DYNAMO-HIA online: http://www.dynamo-hia.eu/

32. The Network Dynamics and Simulation Science Laboratory. Online http://www.vbi.vt.edu/ndssl

33. The Network Dynamics and Simulation Science Laboratory. Our Ebola Research, Online: http://www.vbi.vt.edu/ndssl/ebola

34. The Network Dynamics and Simulation Science Laboratory. Application Page, Online: http://ndssl.vbi.vt.edu/apps/


39. C. A. Myers, P.J. Laz, K. B. Shelburne, B. S. Davidson, A probabilistic approach to quantify the impact of uncertainty propagation in musculoskeletal simulations. Annals of Biomedical Engineering http://dx.doi.org/10.1007/s10439-014-1181-7


48. inspyred 1.0 documentation. Online: https://pythonhosted.org/inspyred/overview.html


