Epidemics 21 (2017) 1–12

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Review

Publicly available software tools for decision-makers during an emergent epidemic—Systematic evaluation of utility and usability

David James Heslop[a,*], Abrar Ahmad Chughtai[a], Chau Minh Bui[a], C. Raina MacIntyre[a,b]

[a] School of Public Health and Community Medicine, University of New South Wales, Sydney, Australia
[b] College of Community Solutions and Public Affairs, Arizona State University, USA

ARTICLE INFO

Article history:
Received 12 October 2016
Received in revised form 24 January 2017
Accepted 19 April 2017
Available online 26 April 2017

Keywords:
Epidemic modelling
Software evaluation
Decision support
Outbreak response
Systematic review

ABSTRACT

Epidemics and emerging infectious diseases are becoming an increasing threat to global populations—challenging public health practitioners, decision makers and researchers to plan, prepare, identify and respond to outbreaks in near real-timeframes. The aim of this research is to evaluate the range of public domain and freely available software epidemic modelling tools. Twenty freely utilisable software tools underwent assessment of software usability, utility and key functionalities. Stochastic and agent based tools were found to be highly flexible, adaptable, had high utility and many features, but low usability. Deterministic tools were highly usable with average to good levels of utility.

Contents

1. Introduction

Emerging infectious diseases (EIDs) are evolving into a significant threat to global populations (Cowling et al., 2013; Gao et al., 2013; Li et al., 2014; Lu et al., 2013; Yu et al., 2013; Assiri et al., 2013). The deepening integration and complexity of modern societies, further rise of large scale international human movement, displacement of populations both human and animal, and disruptions to ecologies and complex response systems present an extraordinary challenge to organisations whose remit is to prepare, respond and recover from epidemics (Fraser et al., 2004; Grassly and Fraser, 2008). Computational modelling and simulation is often the most appropriate and cost effective method of evaluating the ability of epidemic response systems to respond to an epidemic. Computational models have been used extensively in the past to

* Corresponding author at: Samuels Building, Room 309, School of Public Health and Community Medicine, Faculty of Medicine, University of New South Wales, Sydney, 2052, NSW, Australia.
E-mail address: d.heslop@unsw.edu.au (D.J. Heslop).

http://dx.doi.org/10.1016/j.epidem.2017.04.002
1755-4365/

explore the resilience of these response systems, evaluate the performance of existing and proposed policy, identify system failure points, and inform decision making for the optimum allocation, distribution or implementation of critical preventive or response resources. Computational methods are most valuable in situations where it is either unfeasible or unethical to conduct equivalent clinical or population studies, or where forecasting of future scenarios is required (Epstein, 2004; Roche et al., 2011).

There is currently a wide range of freely and commercially available software languages and tools that allow properly trained researchers to generate computational epidemics by directly programming published computational epidemic models – most of which are derived from partial differential (deterministic models) and probabilistic (stochastic models) equations (Anderson and May, 1991; Daley and Gani, 1999; Diekmann et al., 2013; Keeling and Rohani, 2008; McKendrick, 1926). More recently agent-based (or individual based) modelling and simulation programming languages and software development frameworks have become available also allowing researchers to generate synthetic epidemics, and incorporate sophisticated and complicated dynamics at the individual agent through to population level (National Research Council (U.S), 2010; North et al., 2013; North et al., 2010). Decision makers and health system managers outside of major institutions and academia generally do not have the expertise, resources or time to be able to generate such complex computational models in support of decision making, representing a barrier to occasional use. However, decision makers faced with an EID outbreak in their jurisdiction acutely need to understand the consequences of various courses of action, and perhaps have the greatest need for this kind of computational modelling and simulation support and perhaps without ready access to expert modelling and simulation support.

Given the dynamic and changing epidemiological patterns during an outbreak, there is an acute requirement for decision makers to revisit decisions and respond to rapidly changing and dynamic epidemic events; ideally supported by near real-time computational modelling and simulation capabilities. Bridging this gap between the highly technical and resource intensive requirements of formal academic epidemic modelling, with the pragmatic, rapid decision and policy making needs of public health and clinical services during an epidemic, are key challenges for epidemic modellers (Klepac et al., 2015; Metcalf et al., 2015).

There is a broad range of freely available software tools developed by major epidemic response and academic institutions, developed primarily for internal organisational use to assist with decision making and made public to support the wider community. Generally these tools have been simplified for operational and practical reasons, or are only fit to address a limited set of operational questions. In this review we assess and compare the range of available software tools, with particular focus on the ability of these tools to be effectively and adaptively used by decision makers, health system managers and subject matter experts during an outbreak. Additional evaluation of business and computer systems integration will be conducted. Analysing the comparative capabilities of existing modelling and simulation software tools will support the future development of more effective and capable epidemic modelling and simulation, and decision support and management tools.

The objectives of this qualitative review are to: identify the range of public domain stand-alone modelling and simulation software tools available for utilisation by the average or occasional decision maker or public health practitioner, assess the underlying modelling framework of each tool, assess the key constraints and capabilities of each tool, and to make a qualitative assessment of utility and usability, and included features of the available tools.

2. Methods

The study was conducted as a systematic internet and literature search, followed by a qualitative evaluation, and reported in accordance with the guidelines for systematic reviews and meta-analyses as laid out in the PRISMA statement (Moher et al., 2009). The review was carried out between February and July 2016. Information sources were conceptually divided into two groups; 1) internet search engines, including the websites and online resources of major stakeholders in epidemic modelling and simulation, and 2) peer-reviewed and grey scientific publications. These sources were searched in order to identify and evaluate what is the range of freely accessible computer based epidemic modelling and simulation tools currently available to end users and decision makers. This was achieved through the application of a systematic search and identification process.

1) Internet and key stakeholder data sources

Google, Bing and Yahoo search engines were used to undertake a search of online data resources. Free text searches utilising groups of key words were undertaken, utilising a customised list of search terms derived from keywords commonly associated with epidemic modelling tools. For each key word the first three web pages were reviewed.

2) Peer-reviewed and grey scientific publications

The following electronic databases were used to search the peer-reviewed scientific literature: PubMed, Medline, OvidSP, CDC and World Health Organisation (WHOLIS). A step wise search protocol was used, starting initially with MeSH (Medical Subject Heading) terms (PubMed and OvidSP), coupled with Boolean logical operations. Results were combined in EndNote, and duplicates removed through a second round of revisions. Publications were then analysed to extract the data analysis method and tool used for that research. Internet searches for the websites providing the tools further identified related publications, or user manuals, and were cross referenced with the results of the findings of the internet searches.

2.1. Inclusion and exclusion criteria

All tools were subject to systematic application of inclusion and exclusion criteria. The following inclusion criteria were applied: i) the purpose of the tool must be for epidemic modelling or simulation, ii) the tool must be freely available for download and use without restriction, and iii) the freely available version of the software must allow for full evaluation of its capabilities. The following exclusion criteria were applied: i) tools which could not be downloaded or freely evaluated (such as commercial tools, or where special permissions were required), ii) tools which were unable to be installed or utilised using industry standard computing platforms and common operating systems (such as PC, Mac and other modern operating systems), iii) tools where no legal utilisation, licensing or copyright information was provided, and vi) tools which require another software for analysis such as "R".

2.2. Data extraction and analysis

Data was collected by systematically using a customised set of qualitative evaluation criteria based on metrics of fitness for purpose determined by the investigators, basic software usability, sustainability and maintainability (Corporation M, 2017; Gould et al., 1991; IEEE, 1998; Ritter et al., 2014), and relevant published criteria describing good modelling and simulation design and methodology (Jakeman et al., 2006). Data were collated into spreadsheet software (MS Excel®, Microsoft Corporation), and descriptive numerical analysis consisting of generation of feature counts and averages of ratings (outlined below) and the production of simple graphs demonstrating the relative differences in usability, utility and functionality features.

A number of other data fields were also collected in the course of this research including identifying data, aims and purpose of the tool as documented by the publisher, and website reference for data handling purposes. Tools were firstly classified as either predominantly deterministic, stochastic or agent based. Seven (Fraser et al., 2004) specific qualitative measures of performance, utility and usability were applied to each tool, assessed against a five point scale. Additionally each tool was assessed for overall utility and usability separately on a five point scale. Table 1 outlines the descriptors used to assess each component, in addition to the list of key tool functionalities measured as being either present or absent.

The average rating on the five point scale was calculated across the seven individual component measures, and converted to a percentage rating. This was used to rank the tools for the purpose of graphical representation. The assessment of overall utility and usability was also combined into an average percentage rating for the purpose of comparison. The total number of functionality features present for each tool was also calculated, and a percentage of the total maximum possible features (12 possible) calculated. The functionality features assessed with descriptors is outlined in Table 2.

Each tool was evaluated independently by two researchers (DJH and CB) and the findings were then compared for discrepancies. Discrepancies between assessments of qualitative measures and overall measures, or functionality features, were resolved through independent arbitration by a third researcher (AAC).

3. Findings

Internet and stakeholder internet site searches resulted in 33 modelling tools that then underwent detailed evaluation. Searches of the peer-reviewed and grey academic literature did not reveal any additional tools to those found through searches of the internet. Of these, 13 were excluded during systematic application of the inclusion and exclusion criteria (Fig. 1); either due to unavailability for download, not evaluable due to errors in functionality or operation, or failure in installation not correctable through use of supporting documentation where available. Software products were excluded on the grounds of failure of installation only after an additional attempt on a separate computer system had failed. This resulted in a total of 20 tools that underwent full evaluation (a list of the tools analysed with associated hyperlinks is provided in Appendix A in Supplementary material). A summary of the data input features, modelling methodology employed, predicted skill level required of the tool operator and data outputs generated is presented in Table 3.

The findings of qualitative evaluations of the tools are shown in Table 4, sorted from left to right by average qualitative component evaluation score (expressed as a percentage of the maximum score possible). This ranking demonstrated two broad groups of models in relation to their type, performance, utility and usability: Group 1 with <75% on the qualitative rating score, and Group 2 with >75% on the qualitative rating score. Twelve (12/20, 60%) tools were classified as exclusively or primarily deterministic, four (4/20, 20%) were classified as primarily stochastic, and four (4/20, 20%) were classified as agent based tools.

Group 1 tools contained only deterministic or stochastic tools. These tools had less built-in functionality features and were associated with moderate to good levels of utility and usability with regard to their stated purpose (Fig. 2A). They were, however, also found to be relatively inflexible, not adaptable to a range of contexts, and could not account for complexity, incorporate transport modelling, Geospatial Information Systems support, or deal with dynamic interactions. In contrast, Group 2 tools consisted of almost exclusively of agent based or stochastic models. These tools had more added features, but were generally associated with lower degrees of usability, mainly due to a higher burden on the end user in terms of subject matter expertise, technical experience with computers and systems configuration experience. The divergence of overall ratings of usability and comparing Group 2 tools to Group 1 tools is shown graphically in Fig. 2B.

Fig. 2A shows the tools ranked and sorted according to the average of seven utility and usability component measures (expressed as a percentage of maximum possible score). Group 1 modelling tools (ratings <75%) were all deterministic or stochastic models, and Group 2 modelling tools (ratings >75%) were either agent based or stochastic models. Fig. 2B compares overall utility and usability across the range of tools. Group 1 tools were assessed as having high overall usability and utility ratings, but were found to have had relatively few added features. Group 2 tools were assessed as having high overall utility ratings but generally low overall usability ratings (due to tool complexity and operational factors), but had many added functionality features.

Compared to Group 2, Group 1 tools were found to have lower ratings of flexibility and adaptability. Individual component ratings for each tools are shown graphically in Fig. 2. These lower ratings were largely due to Group 1 tools having constraints on the types of pathogens modelled (such as influenza only), modelling only small groups of pathogens of similar characteristics (such as seasonal and pandemic influenza), or modelling in constrained population, spatial or conceptual contexts (single pathogen, in a specified environment such as a laboratory or hospital).

Group 1 tools, primarily deterministic tools, were associated with relatively lower levels of adaptability and flexibility when compared to Group 2 tools, primarily agent based and stochastic tools. Group 2 tools were also more scalable, and could be applied to more realistic and complex geographic contexts, while many Group 1 tools were single compartment (or jurisdiction) models. All tools had high ratings for time resolution and population resolution capabilities.

In contrast, Group 2 tools were significantly more flexible and adaptable (see Fig. 3). They could be applied to whole categories of agents and, for some tools, any agent or even combinations of agents circulating within a community with interactions. This flexibility and adaptability was deliberately incorporated into the design of the tools, and linked to the increased additional features characteristic of Group 2 tools. The increasing numbers of functionality features seen in Group 2 tools is shown graphically in Fig. 2A. The consequence of this increased adaptability and flexibility is a significantly greater burden on the end user in terms of subject matter expertise to design, parametrise and operate the tool, resulting in the lower ratings of usability in this group.

3.1. Group 1 utility and usability

Group 1 tools were implemented in one of three ways – MS Excel® macro enabled workbook, standalone Windows PC applications, or a web service. MS Excel® macro enabled workbook tools were generally user friendly, easy to use, predictable in their operation, and therefore required little additional expertise to operate for the end user. Many of the assessed MS Excel® tools provided detailed end-user documentation, integrated with the tool operation, to enable the end user to understand data entry requirements and the modelling rationale supporting the tool. They were uniformly visually appealing and well referenced, and easy to install

Table 1
Individual Component and Overall Assessment 5 Point Scale Descriptors.

Measure Rating	**Flexibility** Can the tool tolerate a variety of input parameters?	**Adaptability** Can the tool be applied to multiple contexts. To what degree have assumptions been introduced	**Scalability** Can the tool be scaled up or down to any population or other metric size?	**Geographic Resolution** What is the geographic resolution of the model?	**Time Resolution** What is the time resolution of the model?	**Population Resolution** What is the population resolution of the model?	**Computer Platform Compatibility**	**Overall Rating – Utility** Overall assessment of the utility of this tool in relation to its stated purpose	**Overall Rating – Usability** Overall assessment of the usability of this tool
5	Complete flexibility in selecting parametrizations formats and types and data	Adaptable to any context and any situation	User defined scaling to any population or parameter size	User defined, without pre-set limitation, up to and including individual level GIS continuous	Minutes	1	Any	Very Good	Very Good
4	Moderate flexibility in selecting parametrizations formats and types and data	Can be adapted to pathogens across classes, or species, in many contexts	Scaling allowed within the assumptions of the modelling technique used	GIS with defined areas (e.g. Census district)	Days	2–99	3 Platforms, Plus Mobile Devices	Good	Good
3	Some flexibility in selecting parametrizations formats and types and data	Can be adapted to pathogens in one class only, in limited contexts	Scaling within some contexts, according to rules or constraints	Sub jurisdictional divisions allowed or supported	Weeks	100–999	3 Platforms Supported	Moderate	Moderate
2	Limited to little flexibility in selecting parametrizations formats and types and data	Can be adapted to one or a narrow range of pathogens in one context only	Can scale within narrow limits or in narrow context	Multi-area or multi-compartment	Months	1000–9999	2 Platforms Supported	Poor	Poor
1	No flexibility in selecting parametrizations formats and types and data	One pathogen, in one context only.	No scalability features	Single jurisdiction or compartment	Years	>10000	1 Platform Supported	Very Poor	Very Poor

Table 2
Functionality features assessed and their definition.

Feature	Definition
Transport Modelling	Able to model movement of individuals between geographic areas
Demographic Modelling	Able to incorporate divisions of population based on age and other demographic features
Complex Systems Features	Able to incorporate complex adaptive system modelling, such as feedback, agent-environment interaction or recursive modelling
HPC Support	Can utilise High Performance Computing (i.e. supercomputing) to enable larger scale or more detailed modelling tasks
Real Time Capabilities	The end user can employ the tool to support decision making in real-time, within minutes to hours.
Retrospective Analysis	The tool supports entry of data that provides the ability to retrospectively analyse an historical event
Prospective Analysis	The tool generates snapshot or time-series data that allows for prediction of features of an epidemic based on input parameters and assumptions
Decision Support	The tool is designed to directly support decision making, or addresses a specific question related to epidemic response decision making
Risk Management Support	The tool is designed to generate data that supports the development of risk analyses directly
Machine Learning Support	The tool has features that allow the system to adapt to previous scenarios or results, potentially without user direction or input.

on common operating systems and platforms. Minimal to no additional training was required.

Similarly, the web service based tools were also highly user friendly, and offered significant advantages over MS Excel® based tools in terms of accessibility and portability, and thus usability.

The remaining Group 1 tools were provided as Windows PC based stand-alone applications. These required direct installation onto the end-users operating system, the insoluble failure of which was the primary reason that some tools were excluded from this study. These tools were either text based (console) applications, or Windows form based applications. These tools were all specifically designed to assist decision makers in select scenarios – generally geographically based (e.g. Texas or Sweden) or accepted only conceptual/abstract inputs (e.g. nucleotide sequence).

In general, while Group 1 tools operationally met their stated aims and objectives they were generally constrained by the modelling techniques used, data input formats, and available data and graphical output options. Group 1 deterministic tools were assessed as being unable to account for complex systems interactions, emergent phenomenon, or incorporate transport modelling.

3.2. Group 2 utility and usability

Group 2 tools shared common characteristics, and differed significantly from Group 1 tools. While Group 2 tools had higher qualitative component ratings (Fig. 2), and had significantly more additional features (Fig. 2A), they were associated with lower overall usability (Fig. 2B). This was primarily due to the significant additional requirements placed on end users in terms of setup, training, expertise (computer systems, computer programming and epidemic modelling and simulation) and time. Except for the web based tools, all Group 2 tools required computer configuration and installation support to build environments capable of successful operation, storage of results and the use of necessary ancillary applications for specialised data visualisations (particularly GIS analysis and visualisation).

The Epidemiological Modelling (EMOD) software had the highest overall rating of the assessed tools. EMOD was found to offer near complete modelling flexibility and adaptability, and integration of a wide variety of contexts and abstract modelling concepts such as social-technical system modelling, demographic and transport model situation. EMOD utilises a console (text) based modelling system, incorporating many additional features such as Geospatial Information Systems (mapping, linkage and network visualisation for example). Utilisation of these features, however, required the end user to have significant programming, configuration and modelling design expertise, have experience and understanding of deterministic, stochastic and agent based model parametrisation, and how to properly implement this within EMOD. Nevertheless EMOD is able to account for complex phenomena and uncertainty, and supports the widest range of potential decision support needs. Additionally it was assessed as capable of supporting both retrospective and prospective analyses.

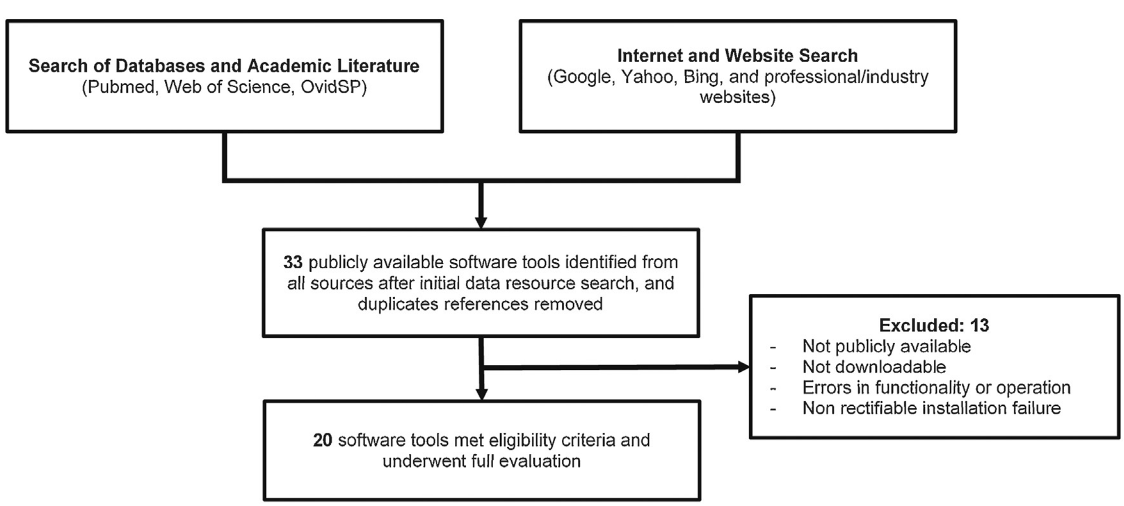

Fig. 1. Identification, Screening and Eligibility Flow Diagram.

Table 3
Features of assessed modelling tools.

Tool Name	Pathogens Modelled	Modelling Methodology	Data Input Requirements	Data Outputs Generated	Required Computer Programming Skills
SimFlu1.0	Influenza A (H1N1, N3N2, H5N1)	Deterministic • genetic sequence mutation model • retrospective	Text input of influenza genetic sequences	FASTA sequence text file	Familiarity with command line program execution
VacStockPile	Vaccine preventable childhood diseases	Deterministic • logistic requirements model • prospective snapshot	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs	Graphical interface No programming skills required
GER	Ebola Virus	Deterministic • logistic requirements model • prospective snapshot	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs	Graphical interface No programming skills required
FluAid2.0	Influenza (generic)	Deterministic • logistic requirements model • prospective time series	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs	Graphical interface No programming skills required
Community Flu	Influenza (generic)	Stochastic • probabilistic parametrisation • multiple iteration individual based model • prospective time series	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs	Graphical interface No programming skills required
StatFlu	Influenza (except H5N1)	Stochastic • logistic demand model • prospective time series	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Data tables and graphs within custom program interface	Graphical interface No programming skills required
FluSurge2.0	Influenza (generic)	Deterministic • logistic requirements model • prospective time series	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs	Graphical interface No programming skills required
FluWorkLoss	Influenza (generic)	Deterministic • logistic requirements model • prospective time series	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs	Graphical interface No programming skills required
Asia Flu Cap	Influenza (generic)	Deterministic • state transition model • Coupled logistic model • prospective time series	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs, GIS system linkable	Graphical interface No programming skills required
FluLabSurge	Influenza (generic)	Deterministic • logistic requirements model • prospective time series	Defaults provided. Advanced use requires familiarity with and user input of numerical disease parameters and assumptions	Excel data tables and graphs	Graphical interface No programming skills required
CPID	Influenza (generic)	Deterministic • state transition model • prospective time series	Web based entry of estimated model parameters and probabilities, with some default values provided	Interactive maps and charts of influenza cases linked to GIS data	Graphical interface No programming skills required
GleamViz	Any	Stochastic • probabilistic state transition model • individual based transport model • prospective time series	Web based entry of estimated model parameters and probabilities, with some default values provided. Customised compartment models and transitions programmable using graphical interface.	Three dimensional, two dimensional visualisations of transmission, transport and distribution of compartment data.	Graphical interface Familiarity with compartment models represented as statecharts
Texas Pandemic	Influenza (generic)	Deterministic • logistic model • state transition model • prospective time series	Direct input of model parameters into custom program interface	Data tables, graphs and GIS output within custom program interface	Graphical interface No programming skills required

Table 3 (*Continued*)

Tool Name	Pathogens Modelled	Modelling Methodology	Data Input Requirements	Data Outputs Generated	Required Computer Programming Skills
BERM	Any	Stochastic • probabilistic state transition model • logistic model (discrete event) • prospective time series	Web based entry of estimated model parameters and probabilities, with some default values provided	Data tables and graphs	Graphical interface No programming skills required
FluTE	Influenza (H2N2 and H1N1)	Agent Based • state transition model • prospective time series	Direct coding of model parameters into settings files prior to execution	Text based data files	Familiarity with generating parameter files in custom format, and command line program execution. C++ experience required for advanced use.
SISpread	Any	Deterministic • state transition model • user defined contact network • prospective time series	Network topology programmable by user. User can adjust limited model parameters through command line switches	Text based data files	Familiarity with command line program execution Network topology data formats
EpiGrass	Any	Stochastic • probabilistic state transition model • user defined contact network • prospective time series	Specific data files specifying model parameters, network topology and model structure	Text based data files. GIS linked visualisation as an option. Database output.	Familiarity with generating parameter files in custom format, and command line program execution
FRED	Any	Agent Based • state transition model • prospective time series	Direct input of model parameters into custom program interface	Data tables, graphs and GIS output within custom program interface	Familiarity with command line program execution
Modelling4All	Any	Agent Based • state or probabilistic transition model • prospective time series	User has complete control and must design epidemic model from scratch using graphical model composer	Graphical visualisation of model progress. Model generated data can be reported and saved for analysis	User must learn bespoke model design rules. Netlogo familiarity required for advanced modelling
EMOD	Any	Agent Based • state or probabilistic transition model • prospective time series	Model design and parameters defined in specialised data file format	Graphical, GIS linked and various text based data formats exportable. Customised outputs.	Familiarity with generating parameter files in custom format, and command line program execution. Python experience required for advanced use.

3.3. Real time capabilities

The tools were assessed for their ability to undertake real-time analysis, defined as the ability to update modelling parameters and inputs and generate updated results rapidly. We found that 19 of 20 tools were able to provide the end user with some form of rapid predictive analysis within the constraints and limitations of the modelling technique and software. In contrast, there were significant differences between Group 1 and 2 tools in terms of their ability to be used “off the shelf”. Group 1 deterministic modelling tools were found to be best suited to “off the shelf” use to support immediate decision support. Stochastic and agent based tools (Group 2) required the end user to first construct a model prior to generation of results. Group 2 tools were assessed as being unable to be used for the first time “off the shelf”, but once configured were assessed as able to provide near real-time results similar to Group 1 tools.

3.4. Cost, maintenance and support

In addition to being freely utilisable for initial use, no tools were associated with a requirement to pay royalties, maintenance or support costs. Most tools provided some form of installation and operational manual mainly in online or pdf format, with certain tools offering very detailed supporting documentation. In general, however, Group 1 tools did not offer additional support options such as formal email contact, while Group 2 tools generally provided additional support options, such as email or telephone support, to end users.

4. Discussion

This study investigates key usability and utility features of freely available software tools which have been developed for use by public health practitioners and decision makers in planning, preparing for and responding to epidemics. We found that the assessed tools could be divided into two distinct groups with common features.

The first group (Group 1) comprising of mainly deterministic modelling tools were associated with lower levels of assessed flexibility, adaptability, and inbuilt features, but good levels of overall usability and utility. Deterministic tools were generally less able to be flexibly deployed across a range of computer platforms, generally were associated with more limited end user support options (such as documentation, or additional forms of utilisation support), and had fewer features such as Geospatial Information Support (GIS) support, and demographic modelling or transport

Table 4
Table of results.

Computational Tool Name	SimFlu1.0	VacStockPile	GER	FluAid2.0	Community Flu	StatFlu	FluSurge2.0	FluWorkLoss	Asia Flu Cap	FluLabSurge	CPID	GleamViz	Texas Pandemic	BERM	FluTE	SISpread	EpiGrass	FRED	Modelling4All	EMOD
Primary Modelling Method*	D	D	D	D	S	D	D	D	D	D	D	S	D	S	AB	D	S	AB	AB	AB
Flexibility	3	2	4	2	2	3	2	2	1	2	1	4	3	4	5	3	5	5	5	5
Adaptability	1	2	1	2	3	1	2	2	2	2	3	4	1	4	5	5	5	4	5	5
Scalability	1	4	2	4	3	2	4	4	4	4	1	1	1	5	5	5	5	4	5	5
Geographic Resolution	1	1	1	1	1	3	1	1	2	3	3	4	4	1	4	5	5	5	5	5
Time Resolution	1	1	4	4	4	4	4	4	4	4	3	4	4	5	4	5	4	4	5	5
Population Resolution	3	5	5	5	5	5	5	5	5	5	5	2	5	5	5	5	5	5	5	5
Platform Compatibility	3	1	1	1	1	1	1	1	1	1	5	3	5	5	1	2	2	5	5	5
Overall Utility Estimate	3	4	4	3	4	4	4	4	4	4	5	4	4	3	5	4	5	5	4	5
Overall Usability Estimate	2	4	4	4	4	4	4	4	4	4	4	4	4	4	2	2	2	5	5	2
Transport Modelling												•			•		•	•	•	•
Demographic Modelling			•	•	•	•	•	•		•		•	•	•	•		•	•	•	•
Complex Systems Features																•			•	•
HPC Support												•	•		•			•		•
Real Time Capabilities	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Retrospective Analysis	•				•				•		•				•		•			•
Prospective Analysis		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Decision Support		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Risk Management Support																				•
Machine Learning Support																				
GIS Support Capabilities									•		•	•	•		•		•	•		•
Multimodal User Support												•				•		•	•	•
	Group 1													Group 2						

[a] The primary modelling method was recorded as either deterministic (D), stochastic (S), or agent based (AB). The ratings on the 5 point scale are shown for the individual and overall measures of utility and usability. The presence of a functionality measure is indicated with • (Tool details are outlined in Appendix A).

modelling capabilities, and simplified prospective modelling capabilities. Nevertheless they were the most easy to install and use, easy to understand, rapid to implement, had the lowest training burden, and were most able to provide real time and "off the shelf" results.

In contrast, the second group (Group 2) comprised mainly of agent based and stochastic modelling tools, and were associated with comparatively higher levels of assessed flexibility, adaptability, overall utility and inbuilt features, but relatively lower levels of assessed overall usability compared to Group 1. Most of the tools in this group placed significant training, installation, data processing and maintenance burdens on the end user. In most cases these tools were not easily usable by the average public health practitioner or health decision maker, and Group 2 tools could not be readily used "off the shelf" to provide immediate results. Therefore Group 2 tools are questionably suitable for use in the initial stages of a rapidly evolving epidemic scenario where there is not time to train end users in the use of more complex software tools. Despite these limitations, Group 2 tools possessed the greatest number of features, were able to generate highly flexible, rich and context specific outputs, and generally had sophisticated GIS, transport and demographic modelling capabilities built in. Some of these tools were also able to utilise High Performance Computing (HPC) resources in addition to operating on stand-alone computer systems.

A subgroup of modelling tools, most notably GleamViz, attempted to address the problem of software complexity impacting on usability through improved user ergonomics, sophisticated Graphical User Interface design, and by placing constraints on the interaction of the end user with the underlying software architecture. This resulted in some degradation of tool flexibility and adaptability, resulting in lower component scores, and introducing additional constraints on applicability to context for end users. This highlights the importance and fundamental challenge of carefully balancing tool complexity with end user requirements during modelling tool design.

Therefore for end users in the health or emergency response domain who require rapid and easily accessible tools to assist with rational decision making, prioritization of interventions and resource allocation, the Group 1 tools are likely to be more useful. The level of technical skill in modelling available in the first response sectors will not be as high or as comprehensive as that available in academia, and tools for front line response and real time decision making necessarily should be simple and usable. However, even though Group 1 tools were found to be the most usable, decision makers may be reluctant to employ them without some form of prior validation or familiarity (Muscatello et al., 2017). This is seen as particularly likely if a decision maker requires modelling support in the context of high consequence decision making in an unfolding epidemic scenario. This reticence has been previously noted in the literature, and identified as a key challenge for the more widespread acceptance of modelling to support rapid decision making (Metcalf et al., 2015).

At this point in time, Group 2 tools seem more applicable to academia than to public health practice at the front line, but this may change in the future with improved modelling and decision support methods. There is a trade-off between degree of model sophistication and comprehensiveness and usability in the field, the impacts and consequences of which are not currently clearly outlined to end users. These differences are summarised in Table 5.

Deterministic models generate their results through calculations based on partial differential equations and the mathematical relationships underlying these models introduce assumptions,

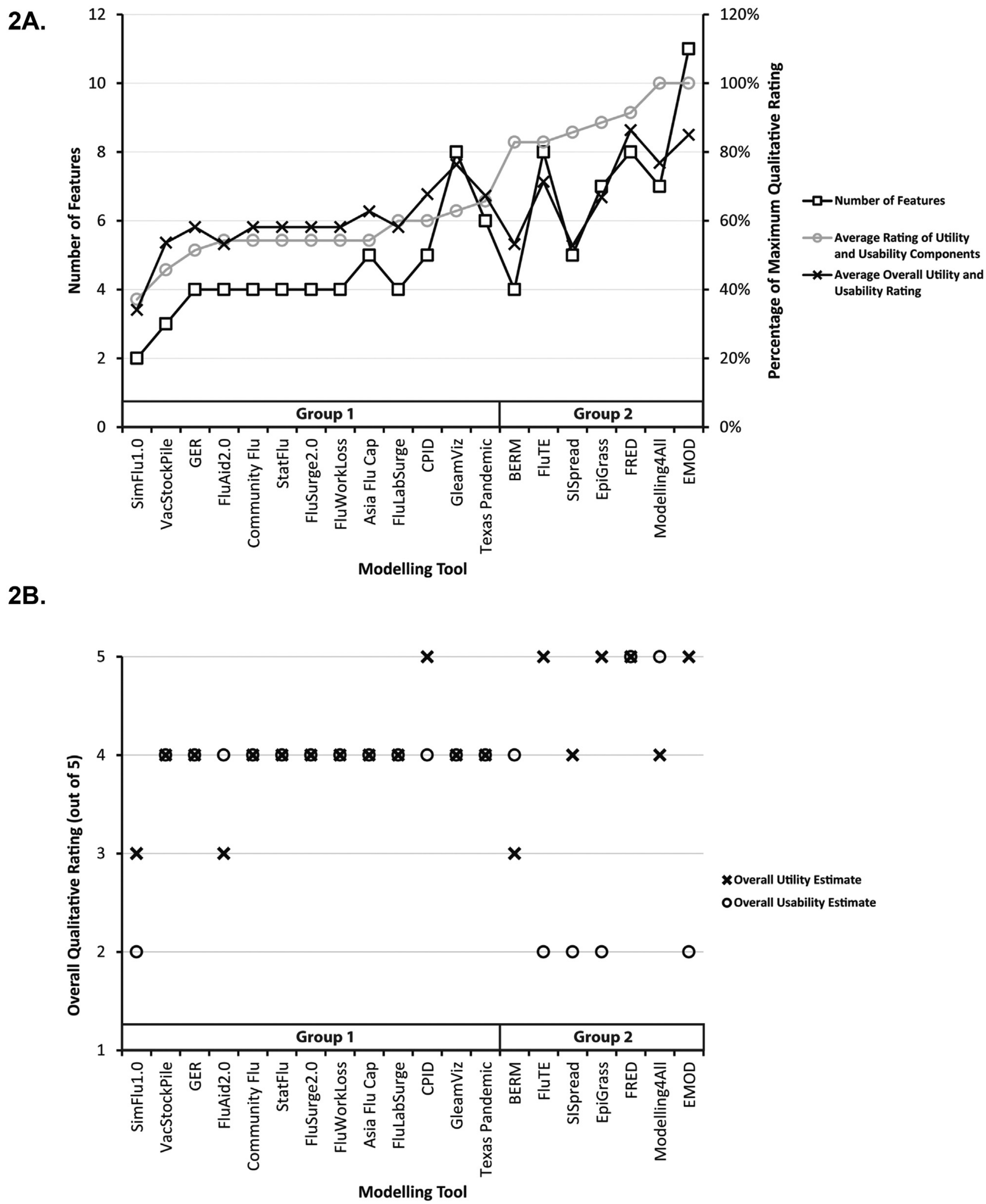

Fig. 2. Comparison of component and overall ratings of utility and usability.

such as homogenous mixing within compartments, instantaneous contacts between individuals and various behavioural simplifications (Metcalf et al., 2015). A risk associated with any form of modelling method, but acknowledged particularly for deterministic tools, is that the novice or average public health practitioner user of a deterministic tool may choose an inappropriate context to apply their model, with the potential for erroneous or inaccurate results to influence decision making, or false reassurance (Roberts et al., 2015).

Compared to traditional deterministic and stochastic models, agent based models have emerged recently as a potentially viable alternative in support of public health practitioners and deci-

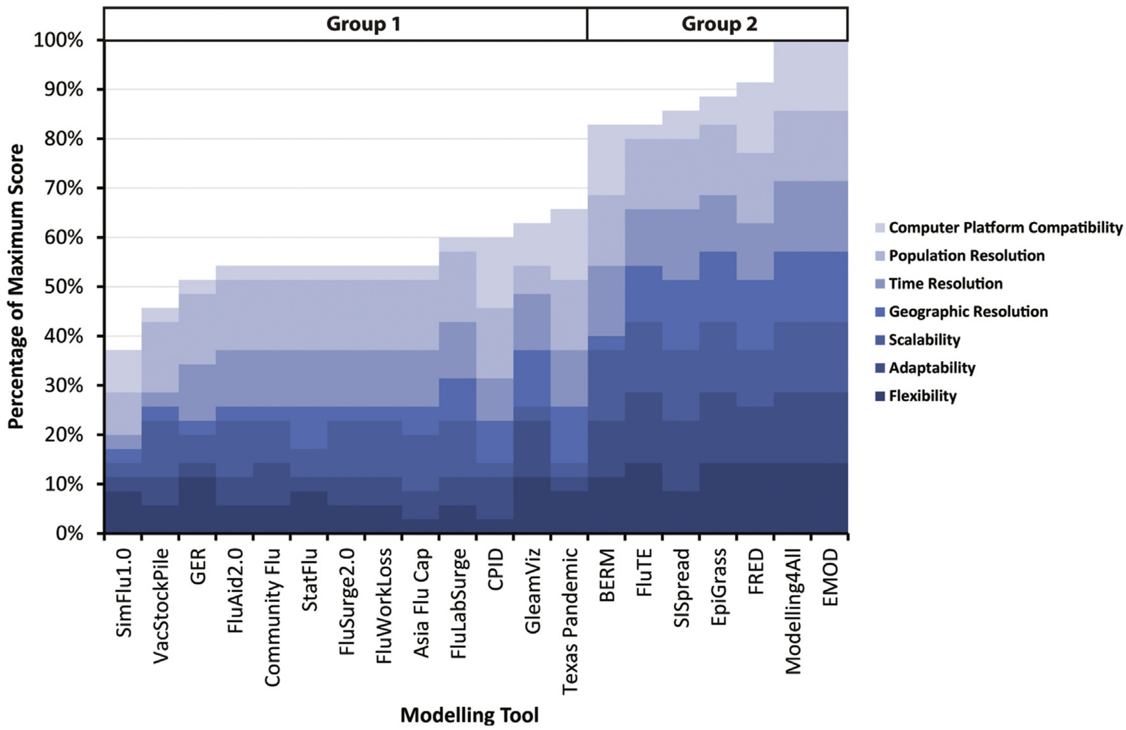

Fig. 3. Ratings by tool for seven Utility and Usability components.

sion makers (North and Macal, 2007; Railsback and Grimm, 2012; Hadžikadicí et al., 2013). While historically this approach has its roots in Von Neumann cellular automata (North and Macal, 2007; Hadžikadicí et al., 2013), recent expansion of ABM approaches have been driven by the development and availability of sophisticated hybrid modelling languages, middleware and specialist tools, the decreased cost of HPC hardware and the availability of low cost Graphics Processing Units (GPU) solutions for HPC, and demonstrations of the validity and applicability of agent based modelling and simulation approaches to complex epidemic modelling problems (National Research Council (U.S), 2010). The future of such models for front line applicability by policy makers and responders will depend on improved accessibility and addressing the currently negative relationship between model complexity and usability.

The two tools examined in this study that achieved a higher score for usability – Framework for Reconstructing Epidemiological Dynamics (FRED) and ModellingForAll – achieved increased usability however through the sacrificing of end user flexibility and end user customisation options. We observed that models allowing increased end user control and options were naturally associated with a requirement for more advanced end user programming skills, greater investments of time and resources, and a greater burden on the end user in relation to data processing and visualisation. Similar significant end user programming and knowledge requirements for agent based tools have been described (Railsback and

Table 5
Summary of overall study findings.

Characteristic	Group 1 Tools	Group 2 Tools
Methodology	Deterministic or Stochastic	Agent Based or Stochastic
Flexibility	Average	Very High
Adaptability	Average	Very High
Scalability	Within assumptions only	Very High
Geographic Resolution	Lower Resolution	Very High Resolution
Time Resolution	Days	Minutes/Days
Population Resolution	Individual	Individual
Platform Compatibility	Limited	Many Systems or Online
Overall Utility	Good	Excellent
Overall Usability	Good	Variable, usually poor
Transport Modelling	Not Implemented	Implemented
Demographic Modelling	Limited	Implemented
Complex Systems Features	Not Supported	Implemented
HPC Support	Not Supported	Supported
Real Time Capabilities	Supported	Supported
Retrospective Analysis	Not Supported	Supported
Prospective Analysis	Limited	Supported
Decision Support	Limited	Supported
Risk Management Support	Not Supported	Supported
Machine Learning Support	Not Supported	Limited
GIS Support Capabilities	Limited	Supported
Multimodal User Support	Limited	Multiple

Grimm, 2012; Bonabeau, 2002; Nikolai and Madey, 2009; Conte and Paolucci, 2014). This has been noted to be most acutely problematic and resource intensive for the critical tasks of parameter validation and verification, and defining attributes, interaction rules and relationships for the software agents, the variability of which has been identified as a wider problem when utilising agent based models (Hadžikadicí et al., 2013; Conte and Paolucci, 2014). Similarly there is a high degree of variability in agent based model reporting standards in the peer-reviewed literature, potentially negatively impacting wider perceptions and acceptance of agent based modelling techniques (Angus and Hassani-Mahmooei, 2015).

Echoing this, a recent study of primarily deterministic and stochastic models developed during the 2014 Ebola outbreak found there were significant barriers to decision and policy makers utilising the models due to variation in modelling techniques, standards of reporting and variable degrees of cooperation in sharing modelling parameters and output data (Chretien et al., 2015). It was noted that the average delay between model formulation and publication for all forms of modelling studies was around 3 months, delaying the critical processes of model parametrisation and adjustment, and significantly impacting the ability of the tools to be used in a rapidly evolving crisis.

One limitation of this study was the qualitative, rather than quantitative, methodology for assessing usability and utility. We chose to develop a customised qualitative tool in order to better evaluate modelling specific measures of utility and usability, rather than purely assessing industry standard measures of usability and utility such as time to failure, or software failure rates (Corporation M, 2017; Gould et al., 1991; IEEE, 1998; Ritter et al., 2014; Jakeman et al., 2006). Equally the functionality of the tools in terms of long term or repeated operation under a variety of conditions by the authors was not addressed. Studies incorporating both quantitative and qualitative measures, assessing usability and utility over a longer timeframe, and relating these measures to emerging standards in modelling and simulation are potentially valuable future research areas. Finally the differential application and interpretation of the qualitative measures by each of the assessors was identified as a possible area of bias, controlled in this study by utilising two independent assessors with third party arbitration of discrepancies. Nevertheless we found a high degree of agreement (no more than 1 point difference on all measures, for all tools) between assessors for each of the qualitative assessment points.

Based on this study we assess that agent based models are well suited to expert technical users but generally unsuited to novice or occasional users (a category in which public health and emergency responders generally fall within). Agent based models were assessed as generally unsuited to support rapid decision making, and could not be readily adjusted in real time to account for rapidly developing situations (that is, timeframes in the order of hours to a few days). Agent based tools which attempted to improve usability did so by significantly decreasing the ability of the end user to adjust or configure the model thereby decreasing utility outside the academic context. In contrast, deterministic and stochastic tools produced rapid results, could readily support real-time use (even over timeframes of minutes to hours), were generally easy to use, required comparatively little end user training or technical expertise, and could be used "off the shelf" within minutes.

A significant disadvantage of deterministic and to some extent stochastic tools was their strict limitation for use within the sometimes narrow contexts and underlying assumptions on which the underlying model was built, and the risk of misuse controlled generally through end user focussed design and extensive supporting documentation. Despite the inclusion in some tools of resources to support end users in the appropriate selection of parameters and correct model use, it is recognised that the occasional end user may be reluctant to invest time and effort to employ a tool that they have not validated, or are unfamiliar. This is particularly so where decision makers may be held to account for their decisions.

Increasing the acceptance of modelling tools for use in rapidly evolving situations could be improved through a number of mechanisms. Increasing modelling literacy amongst end users through the provision of training on computational approaches to epidemic modelling may be beneficial. This approach has been advocated by a number of organisations and researchers recently (Metcalf et al., 2015; Roberts et al., 2015; De Angelis et al., 2015). Equally, where a decision maker would benefit from the use of advanced modelling techniques such as agent based modelling, particularly incorporating transport, demographic and complex systems features, real time modelling support may best be provided through the employment of specialist modelling staff integrated into decision making teams. This approach has been used in a number of jurisdictions successfully (Metcalf et al., 2015). Both of these methods are difficult to sustain and resource intensive, but also run the risk of not reaching decision makers in certain epidemic contexts. Along with increasing the level of end user expertise, model designers need to focus effort on how to incorporate more advanced modelling techniques in a way that decision makers can trust, but also maximise usability and utility.

Available modelling tools should be viewed according to the requirements of the ultimate end user, and depending on their needs – whether they are practitioners with low levels of modelling skills, or academics with high levels of skill – the characteristics, features and limitations of the modelling tools will be different. A challenge for the epidemic modelling community is to build advanced modelling systems that can flexibly and adaptably meet those needs and support decision makers and public health practitioners directly during rapidly emerging epidemics.

Funding

This research was supported by NHMRC Project Grant APP1082524.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.epidem.2017.04.002.

References

Anderson, R.M., May, R.M., 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford; New York, viii, 757 p. p.

Angus, S.D., Hassani-Mahmooei, B., 2015. Anarchy reigns: a quantitative analysis of agent-based modelling publication practices in JASSS, 2001–2012. Jasss-J. Artif. Soc. S. 18 (4).

Assiri, A., McGeer, A., Perl, T.M., Price, C.S., Al Rabeeah, A.A., Cummings, D.A., et al., 2013. Hospital outbreak of Middle East respiratory syndrome coronavirus. N. Engl. J. Med. 369 (5), 407–416.

Bonabeau, E., 2002. Agent-based modeling: methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. U. S. A. 99, 7280–7287.

Chretien, J.P., Riley, S., George, D.B., 2015. Mathematical modeling of the West Africa ebola epidemic. Elife 4.

Conte, R., Paolucci, M., 2014. On agent-based modeling and computational social science. Front. Psychol., 5.

Corporation M, 2017. Usability in Software Design. MSDN Technical Paper [Internet], 2000 05 Aug 16. Available from: https://msdn.microsoft.com/en-us/library/ms997577.aspx.

Cowling, B.J., Jin, L., Lau, E.H., Liao, Q., Wu, P., Jiang, H., et al., 2013. Comparative epidemiology of human infections with avian influenza A H7N9 and H5N1 viruses in China: a population-based study of laboratory-confirmed cases. Lancet 382 (9887), 129–137.

Daley, D.J., Gani, J.M., 1999. Epidemic Modelling: An Introduction. Cambridge University Press, Cambridge; New York, xii, 213 p. p.

De Angelis, D., Presanis, A.M., Birrell, P.J., Tomba, G.S., House, T., 2015. Four key challenges in infectious disease modelling using data from multiple sources. Epidemics 10, 83–87.

Diekmann, O., Heesterbeek, H., Britton, T., 2013. Mathematical Tools for Understanding Infectious Disease Dynamics. Princeton University Press.
Epstein, J.M., 2004. Toward a Containment Strategy for Smallpox Bioterror: an Individual-based Computational Approach. Brookings Institution Press, Washington, D.C, v, 55 pp.
Fraser, C., Riley, S., Anderson, R.M., Ferguson, N.M., 2004. Factors that make an infectious disease outbreak controllable. Proc. Natl. Acad. Sci. U. S. A. 101 (16), 6146–6151.
Gao, R., Cao, B., Hu, Y., Feng, Z., Wang, D., Hu, W., et al., 2013. Human infection with a novel avian-origin influenza A (H7N9) virus. N. Engl. J. Med. 368 (20), 1888–1897.
Gould, J.D., Boies, S.J., Lewis, C., 1991. Making usable, useful, productivity—enhancing computer-applications. Commun. ACM 34 (1), 74–85.
Grassly, N.C., Fraser, C., 2008. Mathematical models of infectious disease transmission. Nat. Rev. Microbiol. 6 (6), 477–487.
Hadžikadicí, M., O'Brien, S., Khouja, M., 2013. Managing Complexity: Practical Considerations in the Development and Application of ABMs to Contemporary Policy Challenges. Springer, Heidelberg, viii, 171 pages p.
IEEE, 1998. IEEE Standard for a Software Quality Metrics Methodology. IEEE, New York, USA.
Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environ. Modell. Softw 21 (5), 602–614.
Keeling, M.J., Rohani, P., 2008. Modeling Infectious Diseases in Humans and Animals. Princeton University Press, Princeton, xi, 366 p. p.
Klepac, P., Funk, S., Hollingsworth, T.D., Metcalf, C.J., Hampson, K., 2015. Six challenges in the eradication of infectious diseases. Epidemics 10, 97–101.
Li, Q., Zhou, L., Zhou, M., Chen, Z., Li, F., Wu, H., et al., 2014. Epidemiology of human infections with avian influenza A(H7N9) virus in China. N. Engl. J. Med. 370 (6), 520–532.
Lu, L., Liu, Q., Du, L., Jiang, S., 2013. Middle East respiratory syndrome coronavirus (MERS-CoV): challenges in identifying its source and controlling its spread. Microbes Infect. 15 (8–9), 625–629.
McKendrick, A.G., 1926. Applications of mathematics to medical problems. Proc. Edinb. Math. Soc. 14, 98–130.
Metcalf, C.J., Edmunds, W.J., Lessler, J., 2015. Six challenges in modelling for public health policy. Epidemics 10, 93–96.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Group, P., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6 (7), e1000097.
Muscatello, D.J., Chughtai, A.A., Heywood, A., Gardner, L.M., Heslop, D.J., MacIntyre, C.R., 2017. Translation of real-time infectious disease modeling into routine public health practice. Emerg. Infect. Dis. (May), http://dx.doi.org/10.3201/eid2305.161720.
National Research Council (U.S), 2010. Committee on Modeling Simulation and Games. The Rise of Games and High-performance Computing for Modeling and Simulation. National Academies Press, Washington, D.C, xiii, 116 p. p.
Nikolai, C., Madey, G., 2009. Tools of the trade: a survey of various agent based modeling platforms. Jasss-J. Artif. Soc. S. 12 (2).
North, M.J., Macal, C.M., 2007. Managing Business Complexity: Discovering Strategic Solutions with Agent-based Modeling and Simulation. Oxford University Press, Oxford; New York, xi, 313 p. p.
North, M., Macal, C., St Aubin, J., Thimmapuram, P., Bragen, M., Hahn, J., et al., 2010. Multiscale agent-based consumer market modeling. Complexity 15 (5), 37–47.
North, M., Collier, N., Ozik, J., Tatara, E., Macal, C., Bragen, M., et al., 2013. Complex adaptive systems modeling with repast simphony. Complex Adapt. Syst. Model. 1 (1), 3.
Railsback, S.F., Grimm, V., 2012. Agent-Based and Individual-Based Modeling: A Practical Introduction. Princeton University Press, Princeton, xviii, 329 p. p.
Ritter, F.E., Baxter, G.D., Churchill, E.F., 2014. Foundations for Designing User-Centered Systems: What System Designers Need to Know About People. Springer, London, xxx, 442 pages p.
Roberts, M., Andreasen, V., Lloyd, A., Pellis, L., 2015. Nine challenges for deterministic epidemic models. Epidemics 10, 49–53.
Roche, B., Drake, J.M., Rohani, P., 2011. An Agent-Based Model to study the epidemiological and evolutionary dynamics of Influenza viruses. BMC Bioinf., 12.
Yu, H., Cowling, B.J., Feng, L., Lau, E.H., Liao, Q., Tsang, T.K., et al., 2013. Human infection with avian influenza A H7N9 virus: an assessment of clinical severity. Lancet 382 (9887), 138–145.