Epidemiol. Infect. (2017), **145**, 1069–1094. © Cambridge University Press 2017
doi:10.1017/S0950268817000164

REVIEW ARTICLE
A systematic review of early modelling studies of Ebola virus disease in West Africa

Z. S. Y. WONG[1,2]*, C. M. BUI[2], A. A. CHUGHTAI[2] AND C. R. MACINTYRE[3]

[1] *Centre for Clinical Epidemiology St. Luke's International University, Chuo-ku, Tokyo, Japan*
[2] *School of Public Health and Community Medicine, The University of New South Wales, Sydney, NSW, Australia*
[3] *College of Public Service and Community Solutions, Arizona State University, Tempe, AZ, 85287, United States*

Received 21 September 2016; Final revision 29 November 2016; Accepted 5 January 2017; first published online 7 February 2017

SUMMARY

Phenomenological and mechanistic models are widely used to assist resource planning for pandemics and emerging infections. We conducted a systematic review, to compare methods and outputs of published phenomenological and mechanistic modelling studies pertaining to the 2013–2016 Ebola virus disease (EVD) epidemics in four West African countries – Sierra Leone, Liberia, Guinea and Nigeria. We searched Pubmed, Embase and Scopus databases for relevant English language publications up to December 2015. Of the 874 articles identified, 41 met our inclusion criteria. We evaluated these selected studies based on: the sources of the case data used, and modelling approaches, compartments used, population mixing assumptions, model fitting and calibration approaches, sensitivity analysis used and data bias considerations. We synthesised results of the estimated epidemiological parameters: basic reproductive number (R_0), serial interval, latent period, infectious period and case fatality rate, and examined their relationships. The median of the estimated mean R_0 values were between 1·30 and 1·84 in Sierra Leone, Liberia and Guinea. Much higher R_0 value of 9·01 was described for Nigeria. We investigated several issues with uncertainty around EVD modes of transmission, and unknown observation biases from early reported case data. We found that epidemic models offered R_0 mean estimates which are country-specific, but these estimates are not associating with the use of several key disease parameters within the plausible ranges. We find simple models generally yielded similar estimates of R_0 compared with more complex models. Models that accounted for data uncertainty issues have offered a higher case forecast compared with actual case observation. Simple model which offers transparency to public health policy makers could play a critical role for advising rapid policy decisions under an epidemic emergency.

Key words: Ebola virus, infectious disease, modelling.

INTRODUCTION

The largest Ebola virus disease (EVD) epidemic in history began in Guinea in December 2013. As of 30 March 2016, the EVD epidemic resulted in over 28 600 cases and 11 300 fatalities mainly in Guinea, Liberia and Sierra Leone [1]. The most recent reported

* Author for correspondence: Z. S. Y. Wong, Centre for Clinical Epidemiology St. Luke's International University, OMURA Susumu & Mieko Memorial St. Luke's Center for Clinical Academia, 5/F, 3-6-2 Tsukiji, Chuo-ku, Tokyo, 104-0045, Japan. (Email: zoiesywong@gmail.com)

case was reported in Liberia in March 2016 and WHO has warned that we may find new flare-ups in the affected countries [2]. With over 17 000 survivors in West Africa and chronic persistence of virus in some survivors [3], such flare ups have already occurred and may continue to be a risk. During an epidemic emergency, it is important to provide scientific justification for rapid policy decisions. Early epidemic models can frame policy decisions by providing epidemiological characteristics to aid effective disease control – they allow policy makers and scientists to characterise disease epidemiology parameters (such as basic reproduction number (R_0), serial interval, infectious period) based on limited/early disease surveillance data, which can assist in understanding how a disease spreads during the early phase of an outbreak.

For previous EVD outbreaks in the Democratic Republic of Congo (1995) and Uganda (2000), modelling approaches consisted of both traditional Susceptible–(Exposed)–Infected–Removed models and more context-specific compartment models (which account for the hospital transmissions and post-death transmissions). Modelling these hospital and post-death transmissions can provide justification for intervention effects in different transmission contexts [4]. Choice of these modelling approaches may impact modelling outputs, for example, a study [5] revealed that estimates of R_0 tend to be underestimated if post-death transmission dynamics are neglected. For the current EVD epidemic, both traditional and context-specific approaches have also been employed; however, the impact of compartment model designs on epidemic parameters estimation and trajectory projection have not yet been systematically evaluated – this is one motivation for our review.

Another motivating factor of this research is to compare outputs of models that do or do not account for underreporting – which is an issue that is present during most outbreaks, but was thought to be a considerably significant issue during the EVD epidemic, due to a number of reasons [6]. Firstly, EVD cases can be asymptomatic [7, 8]. Secondly, West Africa is one of the poorest regions of the world, their health systems, let alone their surveillance capabilities are severely limited – in a study from the Centers of Disease Control and Prevention (CDC), timeliness of reporting was found to be particularly lacking during the early phases of the EVD outbreak [9]. Lastly, a prevailing distrust of Western medicine, particularly in more rural regions, has been thought to deter cases from presenting to health facilities [10]. This review systematically compares models, which have or have not accounted for underreporting.

There are many difficulties which are inherent in early epidemic models, such as uncertainty regarding disease epidemiology, modes of transmission and unknown rates of underreporting. In this study, we build on the work of a recently published EVD modelling review [11] – we systematically evaluate different modelling methods used to study the current EVD outbreak in West Africa and their outputs on key disease parameters, focusing on investigating the impact of using models with different compartmental structures and which account for underreporting. Our objective is to provide directions for future modelling efforts in settings where early disease outbreak data may be limited.

METHODS

A systematic review was conducted in compliance with the preferred reporting items for systematic review and meta-analyses (PRISMA) checklist (http://www.prisma-statement.org/) [12].

Search strategy

Three online databases (Pubmed, Embase and Scopus) were searched for relevant literature published between January 2014 and December 2015. We only focused on Ebola modelling studies published in about 2 years of the start of the outbreak as early publications provided insight and direction to global and national emerging diseases response and control decisions. These databases cover index international journals in the multi-disciplinary field of: public health, biomedical and pharmaceutical research, clinical and experimental research, health policy and management, and scientific, technical and social science research. For each database, we conducted a search with the following search keys: 'ebola' AND 'model*'. All searches included article title, abstract and keywords. The search was limited to studies in the English language. The detailed search strategy was slightly adjusted according to the specific database settings and was reported in Supplementary Document 1. In order to minimise the chance of missing references, we carried out hand search of key Ebola modelling papers from internet and all the included studies were cross-checked with the papers identified from our initial scoping review. The key literature search was carried out on 1 June 2016 and the final search was carried out on 18 November 2016.

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

Inclusion and exclusion criteria

Studies were included if they aimed: (1) to project the trajectory of disease outbreak or (2) to provide early epidemiological parameters estimation of Ebola (including R_0, serial interval, latent period, infectious period and case fatality rate) using modelling methods. Only those studies using the current EVD outbreak data were included. Studies were excluded if the models presented focused only on evaluating intervention strategies without offering parameter estimates or trajectory projection – evaluating such models was determined to be outside the scope of this review. Studies pertaining to EVD outbreaks prior to 2014 were also excluded. Narrative studies, response policy studies, process model studies, phylogenetic and biological or genetic modelling studies were also excluded. We included all relevant modelling studies based on the above criteria and systematically registered key features and attributes associated with modelling techniques, design and major assumptions taken for epidemic estimation and projection.

Selection of studies

We conducted a two-step screening process. In the first screening, titles and abstracts were independently screened by two reviewers – any discrepancies were resolved by discussion and/or referring to the full-text. In the second screening, remaining studies were included or excluded based on contents within the full-text.

Once all studies were identified, we collected quantitative and qualitative information pertaining to the model presented in each study. We focused on synthesising study outcomes that: (1) account for data uncertainty and (2) used hospitalisation and/or funeral compartments. We also summarised and tabulated qualitative descriptions of what questions the model aims to answer, the country of interest, type of model presented, source and date of the case data and model fitting. We quantitatively analysed estimates of key epidemiological parameters, including the R_0, serial interval, latency period, infectious period and case fatality rate. We then synthesised the estimated R_0 by country, use of compartment, consideration of underreporting; and then investigated the relationship between the R_0 estimation with other parameters. Summary statistics, boxplots, scatterplots, non-parametric significance tests (including Mood's median test, Kruskal–Wallis test and Spearman test) were used, where appropriate. All of the analysis was performed using an R version 3.3·1 64 bit platform with the packages 'plotrix' and 'zoo'.

RESULTS

Definition of modelling terms used in this study is listed in Table 1. Phenomenological models are recognised methods (such as summary statistics, regression models and predictive models), which offer computational solutions to determine values of key disease epidemiological characteristics. Mechanistic models describe the transmission of infectious diseases by categorising host populations into various stages of infection. The R_0 is an important parameter which allows epidemiologists and modellers to quantify how easily a disease can spread, and how effective interventions need to be in order to achieve disease control. R_0 is defined as the expected number of secondary cases generated by one infected individual over the course of their infection in a fully susceptible population [13] (i.e. before interventions are put in place or immunity develops).

Of the 874 studies identified through the online database search, 496 duplicates were removed. After preliminary title and abstract screening, 351 studies were excluded. Seventeen were excluded through a second screening of the full-text of articles, based on the exclusion criteria. Four additional studies were identified through hand-searching. Finally, 41 studies met the inclusion criteria and were included in the review (see Figure 1).

From the 41 studies we selected for our review, 16 studies were published in 2014 and 25 were published in 2015. Thirty-five studies aimed for EVD parameter estimation, 27 offered trajectory projection and 21 studies provided both parameter estimation and trajectory projection. There are 11 phenomenological modelling studies, 29 mechanistic modelling studies and one study employed both modelling methods. Among these 41 studies, 14 accounted for data uncertainty or data bias issues. There are five homogeneous models and 25 of the reviewed studies incorporated heterogeneous mixing assumption. Among those heterogeneous mixing studies, 17 considered a hospitalised compartment, 16 considered a funeral compartment and 12 incorporated both hospitalised and funeral compartments.

Overview of EVD modelling studies

Tables 2–4 summarised our descriptive results by research aim. Thirty-one out of the 41 included studies

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

Table 1. *Modelling terms definitions*

Term	Definition
Phenomenological models	Mathematical/statistical expressions that relate different epidemic observations to each other, where the relationship seeks to best describe the data. Descriptive statistics or regression-based models are some of the examples
Mechanistic models	The nature of disease spread relationship is compartmentalised by observed biological processes that are thought to have given rise to the data. The biological possesses could be parameterised and that could be inferred independently from observational outbreak data. These models can be implemented by system-level perspective, such as ordinary differential equations (ODEs), or stochastic models or agent-based models
Homogenous mixing	Homogeneous model assumes that all hosts have identical rate of disease-causing contacts. The simple susceptible-(exposed)-infected-removed model without consideration of additional population heterogeneity are considered as homogeneous mixing in this study
Heterogeneous mixing	Heterogeneous model sub-divides population into different groups, depending upon characteristics that may influence the risk of receiving and transmitting an infection [14]. Models considering any host heterogeneities or included additional compartments (such as hospital and/or funeral) are considered as heterogeneous mixing here
Basic reproduction number	Basic reproduction number is the expected number of secondary cases generated by one infected individual over the course of their infection in a fully susceptible population [13] (i.e. before interventions are put in place or immunity develops)
Serial interval	Serial interval is defined as the time between illness onset in the primary case to illness onset in the secondary case [15]. Understanding serial interval and their moment generating function would help shaping the relationship between epidemic growth rates and reproductive numbers [16]
Latency period	Latency period is the time between infected individual to become infectious. This metric can be converted to the rate at which an exposed becomes infective as a modelling parameter for those models considering exposed stage
Infectious period	Infectious period is defined as the period that an infected person transmits disease to a susceptible person. The reciprocal of infectious period determines the removal/recovery rate for epidemic modelling
Case fatality rate	Case fatality rate is the proportion of deaths of cases over the course of the disease
Underreporting	Underreporting is interpreted as surveillance systems that fail to reflect all infection cases in a given population. Most of the modern surveillance systems are affected by a degree of underestimation of the true incidence of disease [6] due to asymptomatic cases, inability of disease recognition and detection
Compartment model	Compartment model is a type of modelling methods used for mimicking the way how a disease is transmitted among stages of a population system

have specified referencing case data from World Health Organization (WHO) [1]. The WHO source provided cumulative numbers of reported confirmed, probable and suspected cases and deaths. Other commonly used data include: National Ministry of Health data from the affected countries (such as [17]), CDC Morbidity and Mortality Weekly Reports [18] and/or synthesised data from public data repositories (such as Caitlin Rivers of Virginia Polytechnic Institute [19] and Virology Down Under blog [20]). Demographic and population data were taken from sources such as the Central Intelligence Agency (CIA) factbook [21], United Nation (UN) reports [22] and national census. In many studies, model parameters were derived from models of past Ebola outbreaks [23, 24] and/or early published models of the current Ebola epidemic [9, 25]. For instance, Althaus' study in 2014 [26] adopted estimates of incubation and infectious periods from the 1995 EVD outbreak in Congo, and the team published a study for Nigeria in 2015 [27] that incorporated the estimated incubation period from a 1976 EVD outbreak in Zaire.

When developing an epidemic model with more in-depth detailed compartment structures, many state parameters (i.e. parameters defining the transition between infection stages) will be required and calibrated. For example, Barbarossa *et al*. [28] created a model with seven compartments (including hospitalised and buried) and estimated state parameters based on the outcomes of a range of earlier EVD modelling studies [24, 29–31]. It is noted that model parameters should be estimated with caution as they are prone to biases, and the intended prediction outcomes can be

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

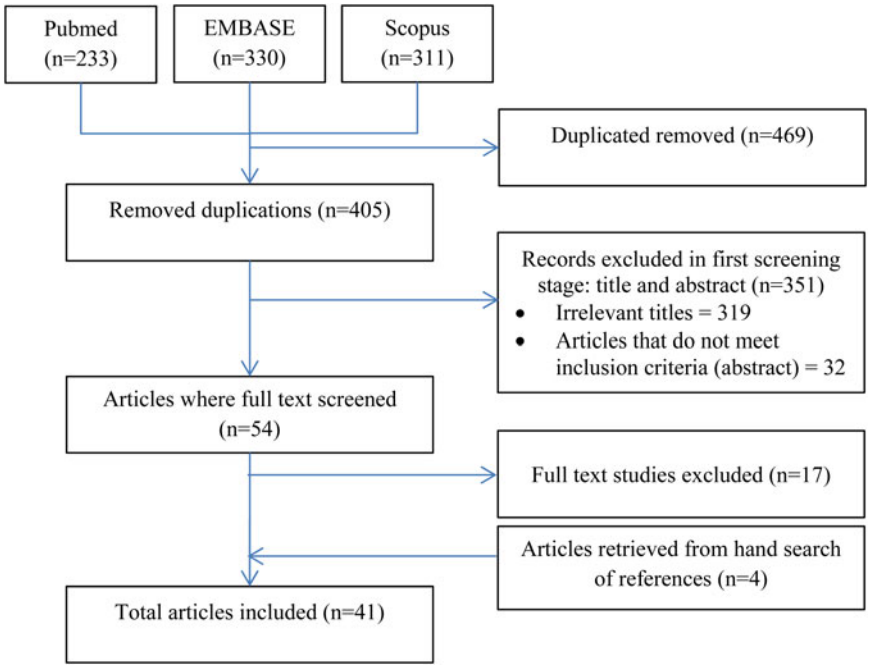

Fig. 1. PRISMA flow diagram of the selection process for including studies in review.

highly sensitive to small changes in some parameters. Twenty-eight of the review studies have carried out sensitivity analysis (so-called stress tests) to examine the robustness of modelling outcomes.

Synthesised results: estimates of epidemiological parameters

Thirty-five studies offer epidemiological parameter estimation and 29 of them estimated the R_0. We evaluated the estimated mean of R_0 by country, compartment consideration and accounting for underreporting, as shown in Figure 2 and reported the details values in Supplementary Document 2 Tables S1–S3. The median of the R_0 mean estimate for the ongoing epidemic (overall) is 1·78 (interquartile range: 1·44, 1·80), 1·30 (interquartile range: 1·24, 1·51) for Guinea, 1·84 (interquartile range: 1·69, 2·10) for Liberia, 1·70 (interquartile range: 1·34, 2·05) for Sierra Leone and 9·01 for Nigeria. Kruskal–Wallis non-parametric test result showed that the estimated R_0 do not have identical data distributions across the included countries (P value $\leqslant 0{\cdot}05$) when we considered Nigerian estimates – which had much higher R_0 estimate comparing with other countries. We performed additional Kruskal–Wallis test (without considering Nigeria estimate) and showed that there is an identical data distributions across other included countries (We cannot reject null hypothesis). When we performed additional pairwise Mood's median tests between each pair of countries (without Nigeria), we found the median values of estimated R_0 are generally not significantly different (apart from the pair between Guinea and Liberia). The results of additional Kruskal–Wallis test and pairwise tests between each pair of countries were reported in Supplementary Document 1 Figure S3. A trend line indicates the potential temporal patterns of the reported R_0 estimation from the bottom-right panel of Figure 2 and there is a slight increasing trend of R_0 when more recent data were used in each study.

The median of R_0 values of 1·90 (interquartile range: 1·90, 2·10) is found for those models accounting for underreporting and 1·71 (interquartile range: 1·40, 1·96) for those not accounting for underreporting. The median of the estimated mean R_0 values is reported as 1·49, 1·80, 1·78 and 1·73 for those models, which considered compartments with hospital, funeral, hospital and funeral, and without hospital and funeral stages, respectively. The Kruskal–Wallis and Mood's median test results revealed that we cannot reject the null hypotheses – i.e. the estimated mean of R_0 remains insignificantly different, regardless of the model's consideration of compartments (with hospital and/or funeral) or accounting for underreporting. The above results yield the same when we do not consider the Nigerian study data. The results were reported in Supplementary Document 1 Figure S3.

We also synthesised the values of other key modelling parameters, including serial interval, latency

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

Table 2. *An overview of modelling studies of Ebola and study designs (Research aim: Parameter estimation)*

References	Dataset/factors	Description of modelling approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Agusto *et al.* [32]	Case data form WHO reports [1]. Population data from Guinea census	Compartmental mathematical model which stratifies population into those in the community and those in healthcare facilities and incorporates disease transmission features in such units	Susceptible–exposed–symptomatic–recovered–deceased–cremated (SEIRDC)	Model divided population into individuals from the community and individuals who work in healthcare settings and subdivided individuals from the community into those who visit healthcare settings and the rest of the public	Some model parameters were fitted using data from other study. Threshold quantity R_0 is estimated using the Ebola data for Guinea and the demographic parameter	R_0 was the response function of sensitivity analysis of other model parameters. partial-rank correlation coefficients (PRCCs) was used	NA
Ajelli *et al.* [33]	Data consisted of routine health data and medical records of the outbreak in the Pukehun District. Also analysed registers of two Ebola holding centres, contact tracing forms and interviewed healthcare workers	Reported summary statistics of key epidemiological parameters, generate time series plots and developed transmission chain for the outbreak using outbreak data	NA	NA	Fitted distribution number of secondary cases using negative binomial model	NA	NA
Althaus [26]	Case data from WHO reports [1]. Parameters from previous Ebola outbreaks	Transmission model with a set of ODEs and used maximum-likelihood estimates to determine R_0	Susceptible–exposed–infectious–recovered (SEIR)	Assumed homogeneous mixing	Fitted the model to the reported data of infected cases and deaths in Guinea, Sierra Leone and Liberia and provided the maximum-likelihood estimates of R_0	NA	NA

Table 2 (*cont.*)

References	Dataset/factors	Description of modelling approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Browne *et al.* [34]	Case data from WHO reports [1]	Deterministic model that traces back to transmissions, and incorporate disease traits and control together	Susceptible–exposed–infectious (hospitalised/reported)– infectious (not hospitalised and unreported)–cumulative hospitalised/reported	Modelled key features of contact tracing and hospitalisation	NA	Studied the effects of model parameters on the effective reproduction number using LHS and PRCC	NA
Gomes *et al.* [29]	Case data from WHO reports [1]. Population data from 'Gridded Population of the World'. Air travel data from the International Air Transport Association and Official Airline Guide. Mobility data from administrative regions	Used the Global Epidemic and Mobility Model to generate stochastic, individual based simulations of EVD spread worldwide	Susceptible–exposed–infectious–hospitalised–death but not yet buried–removed (SEIHFR)	Model included hospitalised and funeral compartment and accounted for different population sizes around the world, including different traffic flows	Multi-model inference approach was used to calibrate on data from official WHO data. Also performed Latin hypercube sampling (LHS) of the parameter space	Performed sensitivity analysis assuming 80% airline traffic reduction from and to West African countries affected by EVD and 50% underreporting	Tested on 50% underreporting assumption
Hsieh [35]	Case data from WHO [1]	Developed mathematical model, the Richards model, to predict the cumulative number of reported cases of infections using variables, such as final case number, per capita growth rate, the exponent of deviation of the cumulative case curve and turning point of the epidemic	NA	NA	Fitted the model to the Ebola data during various time intervals	NA	NA

Table 2 (*cont.*)

References	Dataset/factors	Description of modelling approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Khan *et al.* [36]	WHO situation reports [1], corrected case data from CDC, parameters taken from other studies	Used a deterministic ODE transmission model which differentiates high-risk (e.g. health-workers) and low-risk populations. Estimated the effective contact rate using an ordinary least-squares estimation	Susceptible–exposed–infected–hospitalised–recovered (SEIHR)	Model differentiated high-risk and low-risk populations	Ordinary least-squares (OLS) estimation was used to obtain optimal value of transmission rate and then estimate R_0	NA	Both raw reported data and corrected data (using corrected CDC data) were checked
Kiskowski [37]	Case data from Wikipedia, WHO reports [1]. Incubation and infectious periods based on previous modelling studies	Used a stochastic network model with three levels of community structure (households and communities of households within a country population) to model SEIR transmission dynamics for the EVD spread	Susceptible–exposed–infectious–recovered (SEIR)	Used three levels of community structure in the stochastic model, including (households and communities of households within a country population)	A comparison of general R-square coefficients was used to identify parameter values providing a good fit to the empirical data. R was verified by changing each parameter one-at-a-time	Local sensitivity analysis was carried out to conclude that R values were locally optimised for the given choice of other parameter values	NA
Kucharski *et al.* [38]	Case data from WHO [1] and Sierra Leone Ministry of Health [17]	Developed ODE model with the consideration of hospitalisation in various healthcare units and those who are not ascertained to infection	Susceptible–exposed–infectious (ascertained)–infectious (not ascertained)–healthcare in Ebola Holding Centers (EHC)/Community Care Center (CCC)–Ebola Treatment Unit (ETU)–removed (SEIIHHR)	Assumed individuals who are ascertained initially seek healthcare in EHCs/CCCs. If no beds are available, individuals will be sent to ETUs	Fitted the model to case data reported in each district of Sierra Leone using Bayesian approach	Sensitivity analysis were performed to examine the effect of varying percentage of case ascertainment on infection cases	Incorporated a compartment that consider a proportion of infection cases are not ascertained

Table 2 (*cont.*)

References	Dataset/factors	Description of modelling approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Li *et al.* [39]	Case data from WHO reports [1]. Population data from WHO website. Birth data from CIA factbook [21] and Wikipedia. Death rate from Wikipedia	Developed differential equations model, using least-squares method for parameters estimation. Used partial rank correlation coefficients for uncertainty and sensitivity analysis of R_0	Susceptible–exposed–infectious–treated/recovered (SEIR)	Assumed homogeneous mixing	Fitted the observed variables with onset and death data of EVD by least-squares method	Sensitivity analysis of R_0 were performed by examining seven parameters with PRCCs	NA
Merler *et al.* [40]	Demographic Health Survey data, Case data from Liberian Ministry of Health & Social Welfare Situation Reports and WHO situation reports [1], Household size data from Demographic Health Survey data. Population density Population of the World. Locations of hospitals and clinics from OpenStreetMap. Parameters from other studies	A spatial agent-based model that matched population density estimates id developed. Markov chain Monte Carlo is used to calibrate the model	Susceptible–exposed–infectious–funeral–recovered (SEIFR)	Model accounts for differences in transmission in households, general community, hospitals and funerals	Matching an early report from WHO, Markov chain Monte Carlo approach is used to explore the likelihood of the recorded number of deaths in healthcare workers and in the general population based on official reports. Random-walk Metropolis-Hastings sampling was used to explore the parameter space	Sensitivity analyses were carried out with respect to main epidemiological parameters, on the transmission in the general community and on transmission in hospitals	Considered a under-reporting scenario in which the study still assumed 100% reporting in healthcare workers but a 50% reporting in the general population

Table 2 (*cont.*)

References	Dataset/factors	Description of modelling approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Valdez *et al.* [41]	Case data from WHO [1]	Developed stochastic model with 10 compartmental states to consider various hospitalisation and dead groups using Gillespie algorithm	Susceptible–exposed–infected but not infectious–infected–hospitalised–recovered-dead (SEIHRF)	Considered infected and hospitalised individuals according to their fate. For instance, those who are infected, will be hospitalised, and will die as one class and those who are infected, won't be hospitalised, and will die as class another	Calibrated the model with the data by the maximum-likelihood method. Computed the least-square values of the model outcomes based on a set of parameters generated using LHS from plausible parameter space	Sensitivity analyses were performed to examine the robustness of the estimated values of the transmission coefficients when parameters change	NA
Weitz and Dushoff [5]	Case data from the Caitlin Rivers github website [19]	Transmission model using ensemble adjustment Kalman filter (EAKF) to model a population of deceased infectious individuals	Susceptible–exposed–infected–contaminated deceased–isolated infectious–removed (SEICIIR)	Modelling containment and isolation compartments	Parameters estimated using a least-squares curve fitting algorithm to obtain a choice of parameters with relatively accurate fit	NA	NA
Yamin *et al.* [42]	Model parameters from other studies, Liberia Institute of Statistics and Geo-Information Services. Incidence and case fatality reports, contact tracing data obtained from Ministry of Health and Social Welfare, Republic of Liberia	Transmission model that considers three sequential phases including incubation, early symptomatic, and late symptomatic and accounts for viral load differences in survivors and non-survivors	Incubation-early symptomatic-late symptomatic	Accounted for viral load differences in survivors and non-survivors	Sampling possible ranges of epidemiological parameters to generate contact distribution for contact data. The number of secondary cases arose was calculated	NA	NA

Table 3. *An overview of modelling studies of Ebola and study designs (Research aim: Trajectory prediction)*

References	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Area *et al.* [43]	Case data from WHO [1]	Developed classical differential equations and fractional SEIR model to predict the outbreak trajectory	Susceptible–exposed–infectious–removed (SEIR)	Assumed homogeneous mixing	Fitted models with the real data and obtained parameters values using l^2 norm	NA	NA
Bellan *et al.* [44]	Unspecified	Compared the projections of two simple models (with and without asymptomatic infection) based on the Ebola epidemic in Liberia	NA	NA	NA	NA	Modelled scenarios that does not account for asymptomatic infections and account for the presence of asymoptomic infection
Dong *et al.* [45]	Case data from WHO [1]	Ebola case prediction was obtained by a SEIIRF model with the consideration of post-death and hospitalisation compartments	Susceptible–exposed–infected–recovered–hospitalised–buried (SEIIRHF)	Considered those who have died and are in the process of being buried and number of people in hospital who develop to first stage or second stage of infection	Fitted case prediction curve with real data both in total cases and death cases	NA	NA
Drake *et al.* [46]	Case data from WHO [1], Liberia Ministry of Health or United Nations Office for the Coordination of Humanitarian Affairs (UN-OCHA)	Developed a multi-type branching process model that incorporates key heterogeneities and time-varying parameters to mimic changing human behaviour and controls in Liberia. Focused on the numbers of new infections caused by each case considering offspring distributions	Two generations of infection in a multi-type branching process model	Accounted for subpopulation differences, including hospital treatment *vs.* community care, transmission at funerals, and scenario-dependent transmission risk differences during care-giving	Fitted ensemble model from plausible parameters with the case data in Liberia and also fitted outcome to infection generations in HCWs and in the community	Investigated the sensitivity of the model to parameters using LHS over a much wider range of the least-squares fit for each parameter	Assumed under-reporting by a factor of 2·5

Table 3 (*cont.*)

References	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Fast *et al*. [47]	Case data from the Liberian Ministry of Health and Social welfare ad local county health offices. ETU admissions data from Github. Social mobilisation data from WHO and UNICEF reports. Parameters from UNICEF reports, WHO reports [1]. Population data from Liberian Institute of Statistics and Geo-information Services	Transmission model based on contact network. Accounts for change in attitudes that underlie behaviour change by simulating the progress of the EVD epidemic with and without population behaviour change, then comparing scenarios with observed data	Susceptible-exposed-infected-hospitalised-buried-unburied fatality-recovered (SEIHFCFBR)	Contact network model considered individuals as nodes and disease-spreading contacts as edges. Hospitalisation and fatality components were included	Model was fitted to weekly cases in Lofa County by considering the fitting metric of mean absolute error	NA	The model considered only cases that sought treatment or were safely buried, but modelled all cases
White *et al*. [48]	Case data from WHO [1] and daily counts from the public press releases from the Sierra Leonean Ministry of Health [17]	Developed compartmental model with six compartments to describe the outbreaks in Sierra Leone	Susceptible–exposed–infectious (not yet reported)–treated (reported)–dead (unreported)–recovered (reported)–dead (reported) (SEOTRDTRTD)	Considered infected and hospitalised individuals according to their reported status	Implemented an ensemble trajectory model and generated a matrix of plausible parameter values to fit the model for the first 56 days	NA	Used 2·5 correction factor estimated by the CDC to correct for underreporting

Table 4. *An overview of modelling studies of Ebola and study designs (Research aims: Parameter estimation and Trajectory prediction)*

Reference	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Althaus *et al.* [27]	Case data from published reports. Parameters from previous EVD outbreaks	Transmission model with a set of ODEs and used Maximum likelihood estimates to determine model parameters (baseline transmission rate, rate at which control measures reduce transmission, case fatality rate)	SEIR	Assumed homogeneous mixing	Fitted a dynamic transmission model to data about reported cases and deaths of EVD during a small urban outbreak in Nigeria using Maximum likelihood methods	NA	NA
Barbarossa *et al.* [28]	Case data from WHO reports [1]. Parameters estimated from various studies	Compartmental population model based on Legrand's study [24] that distinguishes between community and hospitalised patients, and recognises importance of deceased individuals who can still transmit the virus at burials	Susceptible–latent–infectious–hospitalised–dead–buried–removed (SEIHBDR)	Model divided infectious population into those who exist in the community and those who exist in hospitals, and considered post-death transmission	Fitted model outcomes to the WHO reports for weekly case incidence. The data were fit with piecewise exponential curves	Studied the effects of model parameters on the basic reproduction number and on the final epidemic size. LHS is used to generate a representative sample set of test parameters from the ranges. Focused mostly on the sensitivity of the time of intervention. PRCCs analysis was reported	NA
Camacho *et al.* [49]	Case data from Sierra Leone Ministry of Health and Sanitation and WHO reports [1]. Data on ETCs, EHCs, CCCs from The Humanitarian Data Exchange. Proportion of symptomatic cases from the UN for Ebola Emergency Response and the National Emergency Response Centre	Transmission model with time-dependent transmission rate, accounting for hospitalisation and delay in case reporting	Susceptible–exposed1–exposed2–infectious–infectious (hospitalisation)–removed	Modelling individuals progressed through stages, including hospitalisation compartment	Fitted model to the time series of weekly reported cases using Bayesian approach	Sensitivity analysis was performed by taking averaged posterior distribution of R over a period of Jan 2015	Assumed proportion of symptomatic cases reported at 60%, accounted for the potential variability in accuracy of reporting over time and accounted for over-dispersed delay between onset of symptoms and notification of reported cases
Chowell *et al.* [50]	Case data from WHO reports [1]	Logistic growth models fitted with the minimal amount of case data	NA	NA	Fitted logistic growth models to the cumulative number of cases using least-squares	NA	NA

Table 4 (*cont.*)

Reference	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Evans and Mammadov [51]	Case data from WHO [1]	Estimated the reproduction numbers for the total period of epidemic and for different consequent time intervals using simple linear model and considered the average infectious period as a time-dependent parameter	NA	NA	Fitted model outcome with the cumulative numbers of infected cases and deaths using global optimisation algorithm DSO	NA	NA
Fasina *et al.* [52]	Case data from WHO reports [1] and public source	Simplified version of Legrand's model [24] accounting for contribution of community and healthcare settings by adjusting baseline transmission rates, diagnostic rates, and enhancement of infection control measures	Susceptible–exposed–infectious–hospitalised–removed from isolation after recovery or death (SEIHP)	Model divided infectious individuals into groups in the community and isolation in a hospital	Assessed the timing of control interventions on the size of the EVD outbreak in Nigeria by extensive simulation runs	NA	NA
Fisman and Tuite [53]	Case data from WHO reports [1], Caitlin Rivers github website [19]	Incidence Decay with Exponential Adjustment (IDEA) model using maximum-likelihood methods to identify best-fit model parameters	NA	NA	Used maximum-likelihood methods to identify the optimal, best and worst case model parameters for the IDEA model	Sensitivity analysis were performed by varying vaccine efficacy	NA
Fisman *et al.* [54]	Case data from Caitlin Rivers github website [19], Virologically-confirmed case counts by date from the Virology Down Under blog [20]	IDEA model, a two parameter mathematical model describes exponential growth simultaneous decay, was used to project epidemic curve accounting for incidence decay	NA	NA	Model was fitted to time series data iteratively, using a progressively increasing number of outbreak generations. Best fit parameter values are estimated by fitting; – objective function: minimise the root mean-squared distance between model estimates and empirical data	Checked separate models to epidemic curves derived from reported deaths, curves based only on virologically confirmed cases, as well as curves based on varying assumptions about case-underreporting	Fitted separate models using assumptions about case underreporting (50% and 100%).

Table 4 (*cont.*)

Reference	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Lewnard *et al.* [55]	Montserrado case data, Number of Beds in Ebola Treatment Centres from Ministry of Health and Social Welfare, Liberia Ebola situation reports. Total population of Montserrado from Republic of Liberia 2008 Population and Housing Census. Time to burial, relative transmission rate from previous study. International SOS Hospital response and isolation/treatment centres also used to determine Number of Beds in ETCs	Developed a differential equation model which includes a latent population (L), two types of recovery populations (R), differentiation of ascertained (A) cases (which do not contribute to community transmission)	SEIR	Assumed homogeneous mixing	Modelled cumulative cases and mortality as a Poisson-distributed random variable. Model calibrated by sampling via Markov Chain Monte Carlo using a Metropolis–Hastings acceptance rule	NA	Addressed underreporting or delays in reporting by fitting model to estimate the delay between the beginning of the infectious period and time of ascertainment
Liu *et al.* [56]	Case data from WHO [1]	Logistic, Gompertz, Rosenzweg and Richards models were developed and compared	NA	NA	Fitted the model using precise estimates by Bayes factors and obtained model parameters	NA	NA
Meltzer *et al.* [9]	Infectious period from The World Bank website, CDC website. Likelihood of a patient going to an ETU and the number of days that a patient in each patient category would spend in the hospital. Actual number of beds in use – from expert opinion	EbolaResponse, a Markov chain model, categorises patient setting (i) hospitalised facility, (ii) home/ community where there is reduced disease transmission, (iii) home with no effective isolation. Model also accounts for under-reporting of cases, and allows for imported cases or cases with no known contacts	Susceptible–infected–incubation–infectious–infectious (burial)–recovered	Considered those who die but whose burial provides risk for onward transmission. Patients were categorised into hospitalised in an Ebola treatment unit (ETU) or medical care facility, home or in a community setting and home with no effective isolation	Model parameters were altered to produce a matched outcome with the reported cases to date using goodness-of-fit test	NA	A underreporting correction factor of 2·5 was used to estimate future total cases

Table 4 (*cont.*)

Reference	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Nishiura and Chowell [57]	Case data form WHO reports [1]	Mathematical modelling considering time- and country-specific incidence data to estimate reproductive numbers, using likelihood-based method for real-time parameter estimation. Estimated daily incidence curves by fitting smoothing spline to county-specific cumulative curves of cases	NA	NA	Fitted a smoothing spline to cumulative reported cases by country and adjusted the spline function based on the daily incidence time series	Sensitivity analyses of reproduction number were carried out by varying the mean generation time	NA
Pandey *et al.* [30]	Demographic data from the 2008 National, Housing Census of Liberia. Case data from Liberian Ministry of Health and Social Welfare	Transmission model takes into account transmission within and between the community, hospitals and funerals	Susceptible–latent–infected–deceased–recovered–buried (SEIFRD)	Stratified epidemiological class into compartments that correspond to the general community, hospitals and funerals	Used weighted least-squares to fit the model to case data. Converted event-based stochastic model to a discrete-time difference equation model. Then evaluated the difference equation model, fitted the output to the data, and calculated the best-fit estimates of certain parameters by minimising the weighted least-squares difference between the model output and the data with the Quasi-Newton algorithm	Done extensive sensitivity analysis and elasticity analysis on intervention effectiveness to variation in epidemiological parameters. PRCCs was used	Refitted model to account for a range of plausible underreporting. Varied the levels of under-reporting of Ebola cases in both community and hospitals, as well as only communities
Rivers *et al.* [58]	Case data from WHO reports [1] and Ministry of Health of Liberia and Sierra Leone (available online from the Caitlin Rivers github website) [19]	Compartmental model adapted from Legrand's study [24]; stochastic model was implemented using Gillespie's algorithm	Susceptible–exposed–infectious–hospitalised–funeral–removed (SEIHFR)	Model accounted for those who are hospitalised and deceased individuals who can still transmit virus	A deterministic version of the model was fit and validated to the current outbreak data using least-squares optimisation. The last 15 days of reported cases were given one -quarter of the weight in the model to preferentially fit the most recent data	NA	NA

Table 4 (*cont.*)

Reference	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Shaman *et al.* [59]	Case data from WHO reports [1]	An ensemble Susceptible–exposed–infectious–recovered-X (SEIRX)–EAKF framework using observations, dynamic modelling and Bayesian inference to generate simulations	SEIRX	Additional X compartments were introduced to describe assimilation of mortality and case fatality rate	Fitted the model variables and parameters with weekly observations and allowed for time varying of variables and R_t parameters	Sensitivity analysis were performed by changing model structure, and varying population size and initial parameter ranges	NA
Shen *et al.* [60]	Case data from WHO [1]	Developed mathematical ODE model to study Ebola infection with isolation, media impact, post-death transmission and vaccination	Susceptible–vaccinated–latent (undetectable)–latent (detectable)–infectious with symptoms–isolated individuals–dead but have not been buried–recovered	Considered those who have died and are in the process of being buried	Fitted the model to epidemiological data of reported cumulative numbers of infected cases and deaths	Examined the most sensitive parameters to the R_0 and the final epidemic size. LHS and PRCC methods were used	NA
Siettos *et al.* [61]	Time series count data, including cumulative incidence, from WHO reports. Cumulative deaths data from Wikipedia and WHO case reports [1]. Demographic data from United Nations [22]	Agent-based model using a small-world network constructed using the Watts & Strogatz algorithm	Susceptible–exposed–infected–dead (but not yet buried)–dead (safely buried)–removed	Agent-based model considered individuals interact through a small-world network	Fitted two network characteristics and four epidemic rates with reported outbreak data	NA	NA
Towers *et al.* [62]	Case data from HealthMap, WHO reports [1]	Fitted piecewise exponential curves along the data time series to estimate the evolving rate of exponential rise (or decline) in cases. SEIR model developed to estimate the temporal patterns of the effective reproduction number for the outbreak in each country	SEIR	Assumed homogeneous mixing	Fitted piecewise exponential function to estimate rates of exponential rise from the average daily EVD incidence data	Sensitivity analyses were performed using LHS and PRCC methods to test the robustness of result with respect to the exact number of contiguous points used for the fits	NA

Table 4 (*cont.*)

Reference	Dataset/Factors	Description of Modelling Approaches	Compartments, if applicable	Assumption of population mixing, if applicable	Model fitting and calibration approach, if applicable	Sensitivity analysis, if applicable	Description of account for data bias, if applicable
Webb *et al.* [63]	Case data from WHO reports [1]. Parameters from various studies	Differential equations with compartments of the epidemic population that simulated forward projection of epidemic using Continuous Time Markov Chain. Model incorporates contact tracing of infectious cases. Both deterministic and stochastic models were run	SEICIIR	Modelling containment and isolation compartments	Parameters estimated using a least-squares curve fitting algorithm to obtain a choice of parameters with relatively accurate fit	Sensitivity analyses were performed in two contact tracing parameters. The contact tracing parameters were tested in sensitivity analysis	A ratio of unreported case of 1·78 is used for estimation
WHO [25]	Case data from investigation forms from confirmed, probable and suspected EVD cases identified in Guinea, Liberia, Nigeria, and Sierra Leone; also from informal case reports; data from diagnostic laboratories; data from burials	Projected case numbers using two methods: (i) regression method and (ii) stochastic branching process model	NA	NA	Epidemiological parameters were fitted with gamma probability distributions	Sensitivity analyses were performed by assuming different mean serial intervals of 11 and 13 and including suspected as well as confirmed and probable cases in the analysis	NA
WHO [64]	Case data from viral haemorrhagic fever data collection forms, treatment facilities, contact tracing forms	Reported summary statistics and prediction outcomes through simple model fitted to data. Used weighted average to estimate the duration reported of the observed means of the distributions of durations from hospitalisation to discharge and hospitalisation to death	NA	NA	Gamma distributions were fitted to confirmed and probable cases	NA	NA

period, infectious period and case fatality rate, as shown in Figure 3 (and Supplementary Document 1 Figure S4). The median of the mean serial interval is 14·35 days (interquartile range: 12·28, 16·35), latency period is 9·70 days (interquartile range: 8·80, 10·38), the infectious period is 7 days (interquartile range: 4·00, 10·00) and case fatality rate is 0·68 (interquartile range: 0·48, 0·71) (the distributions by country are shown in Figure 3).

We also studied the relationship between estimated R_0 and the used or estimated epidemiology parameters. Different studies may estimate/fine tune these parameters or use previously published values when producing R_0 estimates. Figure 4 demonstrates the relationship between the estimated R_0 with different serial interval, latency period, infectious period and fatality rates (We also reported the same figure without considering the Nigerian data in Supplementary Document 1 Figure S4). Based on the results, we do not observe any obvious trend of R_0 estimates from studies using/estimating various mean values of serial interval, incubation period and infectious period. We carried out correlation tests between these values (for pairwise complete observations) and the Spearman test results were insignificant (on overall and per-country basis). The synthesised values of reported epidemiological parameters in terms of R_0, serial interval, latency period, infectious period and case fatality can be found in the Supplementary Document 2 Table S2.

Synthesised results and figures stratified by modelling types, approaches, mixing assumptions and sensitivity analyses are provided in Supplementary Documents 1 Figures S5–S12 and Supplementary Documents 2. We do not observe significant difference in the distributions of R_0 mean estimate from those included studies did or did not carry out sensitivity analysis – as most sensitivity analyses are typically conducted in orthogonal manner to estimate related parameters. Furthermore, R_0 is also sometime being used as a response function of sensitivity analysis of other model parameters. Additionally, 27 studies offered epidemic trajectory projection and provided estimated number of cases (without additional intervention). Figure 5 shows the relationship between model prediction to WHO case observation ratio (matched with forecast target date) and account for underreporting and consideration of compartment. The median values of ratio between prediction and observation are significantly different (P value $\leqslant 0{\cdot}1$) in the pair between do and do not account for underreporting [median ratio of those do account for underreporting is 4·35 (interquartile range: 1·52, 15·03) and do not account for underreporting is 1·19 (interquartile range: 0·98, 1·53)]. However, we do not observe a significant relationship in the pairs of different compartments used. We also paired up models that offered model prediction with and without considering underreporting and carried out the same set of analysis. The matched models analysis results are provided in Supplementary Document 1 Figure S12. The synthesised results of epidemic projection with indication of accounting for data uncertainty and used of hospitalisation and/or funeral compartments are also provided in the Supplementary Document 2 Table S3.

DISCUSSION

We systematically reviewed 2014–2015 Ebola modelling studies, which provided epidemiological insights to the current Ebola and future outbreaks. We evaluated the selected studies based on the sources of the case data used, and modelling approaches by modelling aim and we further synthesised the reported R_0 results and the distributions of key epidemiological parameters based on compartment designs, and consideration of underreporting. We found that epidemic models offered R_0 mean estimates for this EVD are country-specific, but these are not associating with several key disease parameters, compartment designs and accounting for underreporting.

In this EVD outbreak, we noticed a significantly different relationship between the contexts (i.e. the country of interest) with estimations in R_0 (particularly with regards to Nigeria, which had a much higher R_0 estimate). We only observed differences in the values of estimated/used serial interval, latency period, infectious period and case fatality rates by country but the median differences are not statistically significant.

We generally did not observe any apparent systematic pattern in the distribution of estimated R_0 when specifying different compartments. This may be due to one fundamental issue that models are generally fitted based on observed epidemiology data from the same original sources and other associated model parameters within the model could be calibrated altogether to achieve model fitting. This also coincides with our finding – models that utilised different mean serial intervals, incubation periods and infectious periods within plausible ranges yielded similar estimates of R_0.

R_0 can be expressed as the product of transmission probability per contact, number of contacts per time

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

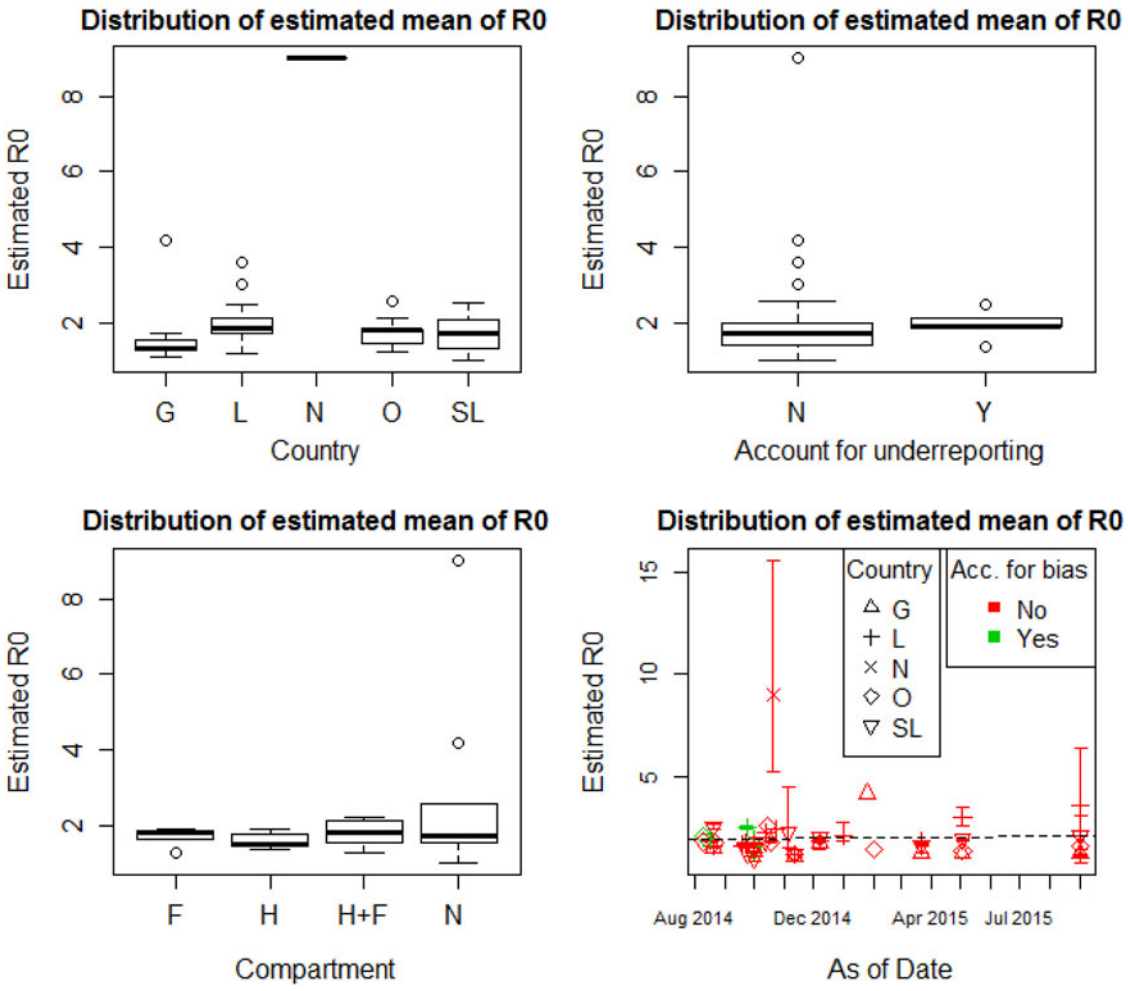

Fig. 2. Summary of estimated basic reproduction numbers (topleft) by West African countries (G, Guinea; L, Liberia; N, Nigeria; O, Overall; and SL, Sierra Leone), (topright) by account for underreporting, (bottomleft) consideration of compartment (F, funeral; H, hospitalisation; H+F, both hospitalisation and funeral; N, not considered), (bottomright) last updated data used (with trend line of R_0 estimation). Kruskal–Wallis non-parametric test result showed that the differences between the medians of the estimated R_0 mean by country (excluding Nigeria) are statistically insignificant (P value $>$ 0·05). Furthermore, we cannot find any significant relationship to reject the null hypotheses for those accounting for underreporting and different compartments used (by both Kruskal–Wallis and Mood's median tests).

unit and duration of infectious period; however, these quantities are usually difficult to parameterise directly from observing a outbreak [14]. In a simplified epidemic model assumption, the lengths (and distributions) of serial interval [16] and infectious period [65] are influential to R_0 estimation. Furthermore, inclusion of a latency period into a model would result in a slower epidemic growth rate after pathogen invasion due to individuals needing to pass through the exposed class before they can contribute to the transmission process [14].

Although serial interval, latency period and infectious period are computationally related to the estimate of R_0 (differences account for compartmental contact heterogeneity and epidemic model assumptions), we observed that the use of different mean serial intervals, incubation periods and infectious periods yield similar estimates of R_0. However, we only compared the relationship by median durations of these epidemiological parameters – using different mean distributions may produce a different effect with the same values.

Due to changes in reporting systems and/or public awareness of disease over time, there may have been unknown observation biases and errors over time (i.e. there may be higher rates of disease reporting due to greater public awareness of the disease). The included studies that considered underreporting issues generally offered similar R_0 estimates to those without incorporated underreporting and offered a larger case

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

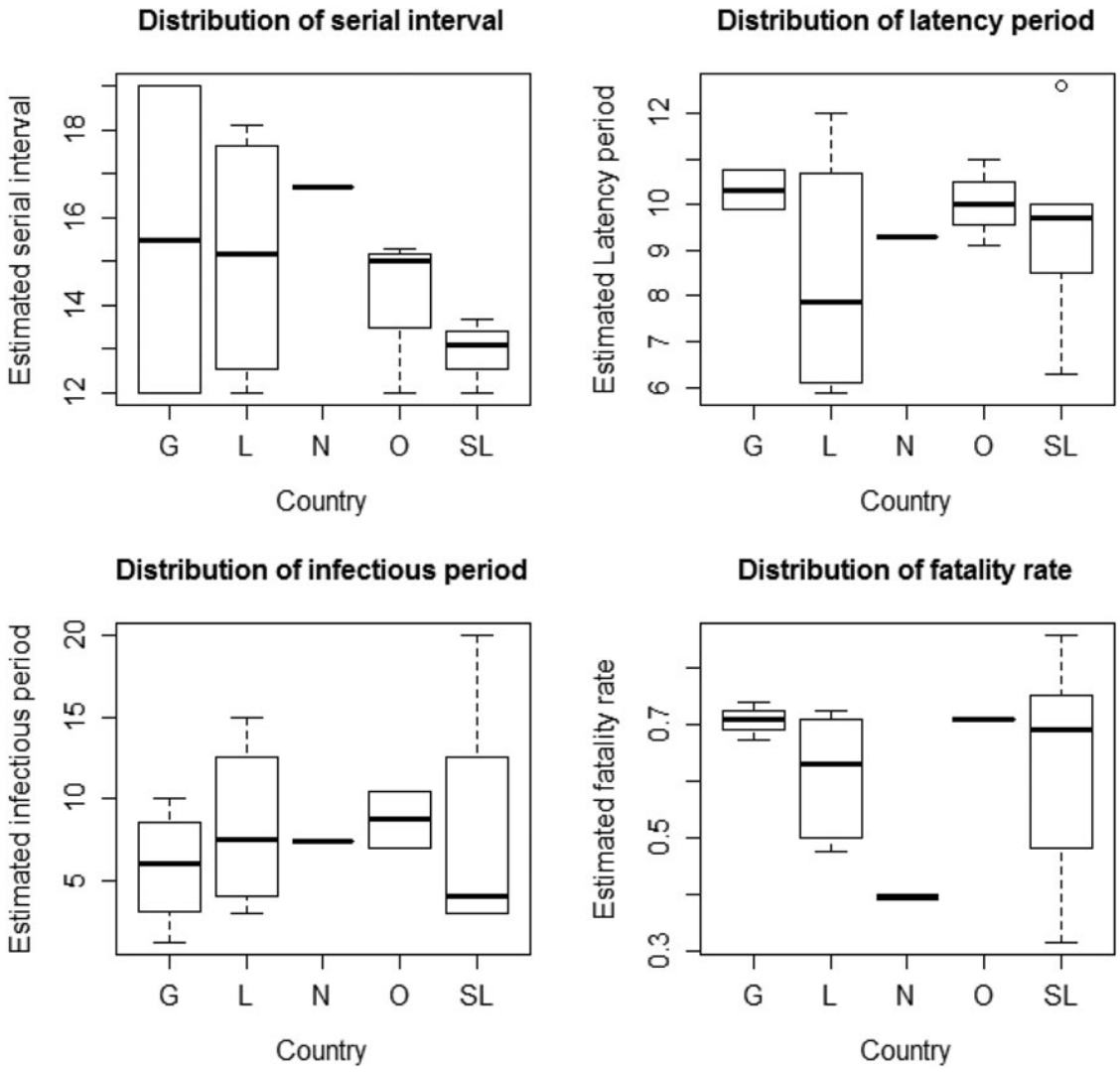

Fig. 3. Summary of estimated epidemiology parameters by country (G, Guinea; L, Liberia; N, Nigeria; O, Overall; SL, Sierra Leone) for (topleft) serial interval, (topright) incubation period, (bottomleft) infectious period, (bottomright) fatality rate. Kruskal–Wallis non-parametric test results showed that we cannot reject the null hypotheses – i.e. the mean values of these epidemiology parameters have identical data distributions from the included countries.

forecast compared with actual case observation at the forecast target date. Without high-quality and reliable data, it is challenging to accurately estimate the impact of time-dependent changing factors and incorporate them into a model. We recommend health authorities endeavour to share detailed epidemiologic attributes related to reported cases, including geographical locations of cases, information on contact networks and date of symptom onset. The availability of ongoing case data can help to recognise hidden changes of disease patterns over time. Health authorities could consider providing time-dependent associated correction factors according to the practical underreporting situations and force of intervention strategies. Furthermore, incorporating surveillance outcomes from phylogenetic [66] and serological [67] studies would potentially useful for advising the underlying emerging disease transmission characteristics and identify undetected cases. These would be useful to modellers for accurately calibrating epidemiological models based on actual outbreak situations, which can then feedback meaningfully into decision support during the outbreak.

Furthermore, we observed that many included models in this review have inferred data about disease behaviour using disease parameters calculated from previous EVD outbreaks. It is noted that this Ebola outbreak has similar parameters estimates with the previously EVD outbreak parameters given by Drake *et al.* [4]. Furthermore, it will be useful to develop a

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

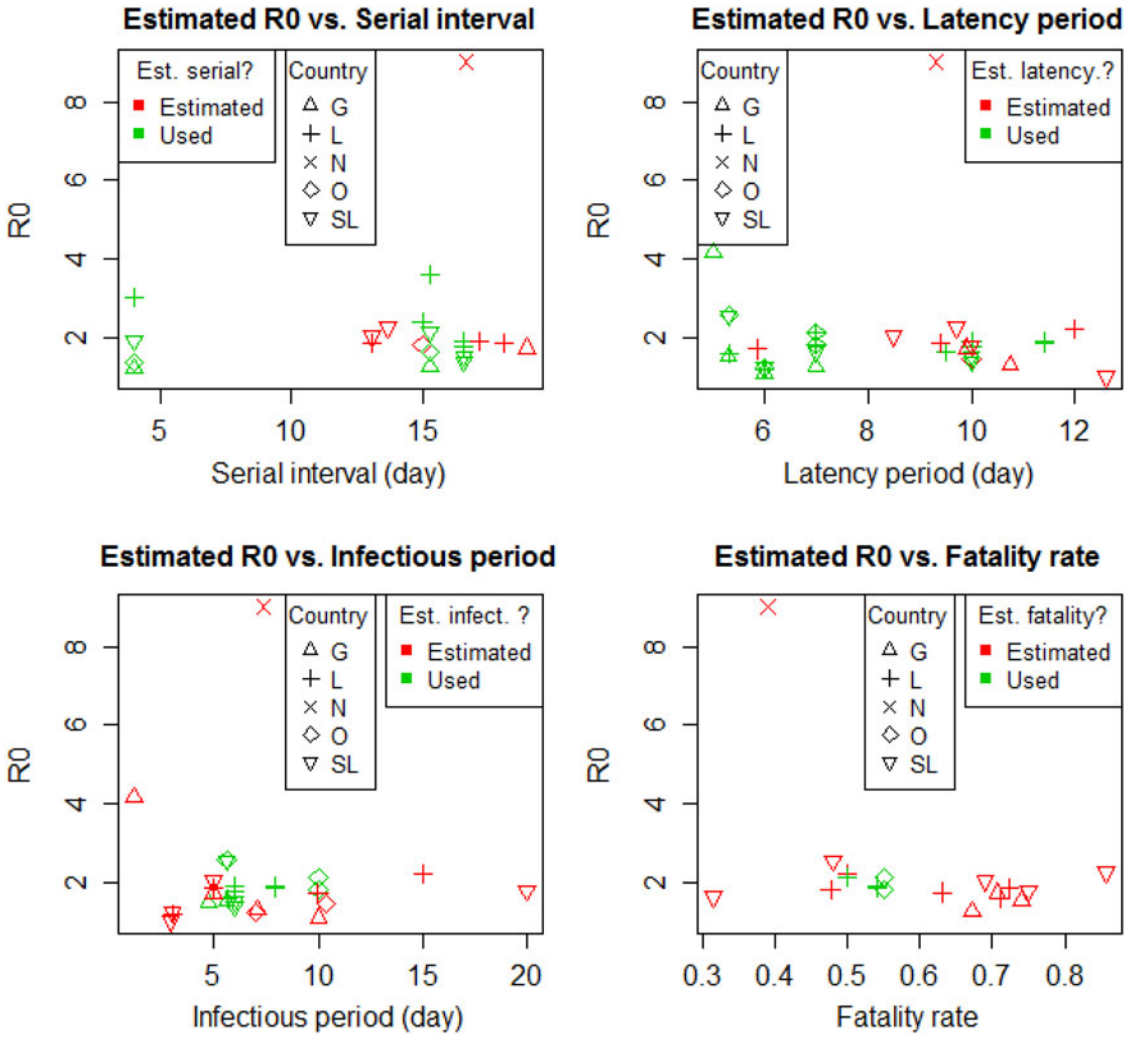

Fig. 4. Relationship between estimated R_0 and epidemiology parameters. (topleft) serial interval, (topright) incubation period, (bottomleft) infectious period, (bottomright) fatality rate. (G, Guinea; L, Liberia; N, Nigeria; O, Overall; SL, Sierra Leone) (Spearman tests among these pairs (for pairwise complete observations) are all insignificant). Only complete pairs between R_0 and epidemiology parameters are shown in this figure.

centralised reference data platform, which allows for sharing of epidemiological parameters. This would enable modellers to use/calibrate disease parameters based on comparable metrics. Together with other EVD modelling review outcomes [4, 11], this study outcome laid groundwork for such reference.

During the later stage of the EVD epidemic, new evidence emerged that EVD can survive in various body fluids during convalescence [68, 69] and may result in transmission of infection [70–75]. WHO has recently highlighted the potential of the occurrence of EVD flare-ups and disease re-introduction [2]. At the post-EVD outbreak stage, accounting for risk of transmission from 'recovered' EVD patients [76, 77] should be a priority for future EVD modelling.

Accuracy, transparency and flexibility are the major considerations when formulating models for infectious diseases [14]. The EbolaRepsonse tool created by the CDC [9], which allows users to project the number of Ebola cases in Liberia and Sierra Leone using a simple model implemented on a Microsoft Excel Worksheet. The WHO Response Team [31] study was one of the first studies providing a detailed epidemiological description of the EVD epidemic using primary data collected from hospitals and patients – the study also provided short-term projections of the epidemic for Guinea, Liberia and Sierra Leone. These early modelling studies are some of those successful examples that demonstrate how timely and effectively use of phenomenological and mechanistic

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

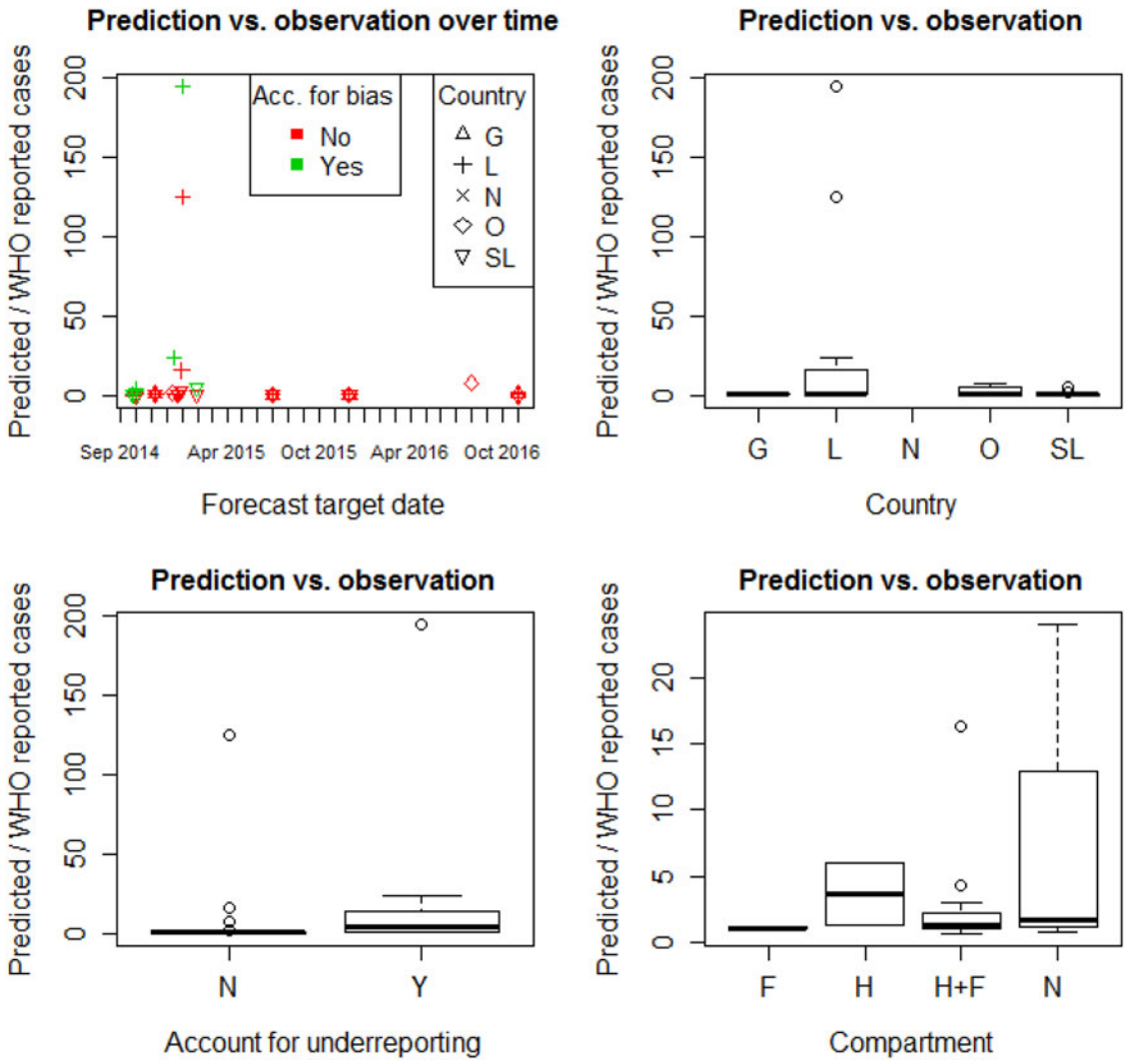

Fig. 5. Summary of ratio between predicted cases and WHO reported cases (matched with forecast target date). (topleft) forecast target date, (topright) by West African countries (G, Guinea; L, Liberia; N, Nigeria; O, Overall; and SL, Sierra Leone), (bottomleft) by account for underreporting, (bottomright) consideration of compartment (F, funeral; H, hospitalisation; H + F, both hospitalisation and funeral; N, not considered). The Mood's two sample median test from the pair between do and do not account for underreporting shows that the median values of ratio between prediction and observation are significantly different (P value $\leqslant$ 0·1). We cannot reject the null hypotheses of the pairs of different compartments used (by both Kruskal–Wallis and Mood's tests).

modelling methods support emerging disease understanding and public health responses.

Modelling outcomes can be different depending on the epidemiological characteristics of the emerging diseases and the interplay among pathogen, host and environment. For this EVD outbreak, simpler models generally yielded similar estimates of R_0 regardless of the consideration of additional hospitalisation and/or funeral compartments and underreporting. However, context-specific models, which mimic the way a disease is transmitted among stages realistically and meaningfully, allow for justification for intervention effects under different transmission contexts. Under an epidemic emergency, a "simple enough but not simpler" model, which allows for understanding of key epidemic dynamics (i.e. offer transparency to public health policy makers), may play a critical role for advising rapid policy decisions and predicting outbreak progression at the early phase of an outbreak.

CONCLUSION

Newly emerging infectious diseases are the most challenging to manage due to the uncertainties in clinical

https://doi.org/10.1017/S0950268817000164 Published online by Cambridge University Press

impact and transmissibility. Epidemiological parameters are usually difficult to observe directly from an emerging infectious disease outbreak and would require modelling methods to accurately estimate at the early phase of a disease outbreak. Epidemic modelling is a quantitative approach to understanding disease transmission within a specific population and provides indications of future trends. Such models are well-recognised tools to aid policy formulation in the early phase of epidemics.

Despite varied complexity and methods, the estimates of R_0 yielded from numerous studies were reasonably consistent in this EVD outbreak regardless of concurrent use of other associated epidemiology parameters. Different model design decision did not appear to meaningfully impact the resulting R_0 estimates but models that accounted for data uncertainty offered a larger case forecast compared with actual case observation at the forecast target date. Simple early models remain informative to reference, and provide a foundation for more complex transmission modelling to understand the progression of a disease outbreak.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at https://doi.org/10.1017/S0950268817000164.

ACKNOWLEDGEMENTS

This study was supported by National Health & Medical Research Council (NHMRC) – Project Grant (Australia) – APP1082524.

AUTHORS' CONTRIBUTION STATEMENT

The study was conceived by Z.W. and R.M. Z.W., C. B., A.C. and R.M. contributing to the study design. Research data/information was retrieved and interpreted by Z.W. and C.B. Z.W. and C.B. led the writing of the paper and all authors revised and refined the arguments. All authors approved the article.

DECLARATION OF INTEREST

None.

REFERENCES

1. **World Health Organization.** Ebola Situation Reports. World Health Organization, 2016 [cited 2016 18 Jan]. (http://apps.who.int/ebola/ebola-situation-reports).
2. **World Health Organization.** Latest Ebola outbreak over in Liberia; West Africa is at zero, but new flare-ups are likely to occur: World Health Organization; 2016. [cited 2016 18 Jan]. (http://who.int/mediacentre/news/releases/2016/ebola-zero-liberia/en/).
3. **MacIntyre CR, Chughtai AA.** Recurrence and reinfection – a new paradigm for the management of Ebola virus disease. *International Journal of Infectious Diseases* 2016; **43**: 58–61.
4. **Drake JM, *et al.*** Transmission models of historical Ebola outbreaks. *Emerging Infectious Diseases* 2015; **21**: 1447–1450.
5. **Weitz JS, Dushoff J.** Modeling post-death transmission of Ebola: challenges for inference and opportunities for control. *Scientific Reports* 2015; **5**: 1–7.
6. **Gibbons CL, *et al.*** Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health* 2014; **14**: 1–17.
7. **Heffernan RT, *et al.*** Low seroprevalence of IgG antibodies to Ebola virus in an epidemic zone: Ogooue-Ivindo region, Northeastern Gabon, 1997. *Journal of Infectious Diseases* 2005; **191**: 964–968.
8. **Leroy EM, *et al.*** Human asymptomatic Ebola infection and strong inflammatory response. *Lancet (London, England)* 2000; **355**: 2210–2215.
9. **Meltzer MI, *et al.*** Estimating the future number of cases in the Ebola epidemic – Liberia and Sierra Leone, 2014–2015. *Morbidity and Mortality Weekly Report Surveillance Summaries (Washington, DC: 2002)* 2014; **63** (Suppl. 3): 1–14.
10. **Manguvo A, Mafuvadze B.** The impact of traditional and religious practices on the spread of Ebola in West Africa: time for a strategic shift. *Pan African Medical Journal* 2015; **22** (Suppl. 1): 9.
11. **Chretien J-P, Riley S, George DB.** Mathematical modeling of the West Africa Ebola epidemic. *eLife* 2015; **4**: e09189, 1–15.
12. **Moher D, Liberati A, Tetzlaff J, Altman DG.** Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery (London, England)* 2010; **8**: 336–341.
13. **Diekmann O, Heesterbeek JA, Metz JA.** On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* 1990; **28**: 365–382.
14. **Keeling MJ.** Modeling infectious diseases in humans and animals. In: Rohani P, ed. Woodstock: Princeton University Press, 2008, 10–11,43.
15. **Fine PE.** The interval between successive cases of an infectious disease. *American Journal of Epidemiology* 2003; **158**: 1039–1047.
16. **Wallinga J, Lipsitch M.** How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* 2007; **274**: 599–604.
17. **Ministry of Health and Sanitation TRoSL.** Ebola Situation Report 2016 [cited 2016 Nov 18]. (http://health.gov.sl/?page_id=583).

18. **(MMWR).** MaMWR. Current weekly 2016 [cited 2016 Nov 18]. (http://www.cdc.gov/mmwr/index.html).
19. Data for the 2014 ebola outbeak in West Africa 2016 [cited 2016 Nov 18]. (https://github.com/cmrivers/ebola).
20. blog Vs. Ebola virus disease (EVD) 2014 West African outbreak 2016 [cited 2016 Nov 18]. (http://virologydown under.blogspot.jp/2014/07/ebola-virus-disease-evd-2014-west.html).
21. **Agency CI.** The World Factbook 2016 [cited 2016 Nov 18]. (https://www.cia.gov/library/publications/the-world-factbook/).
22. Nations U. Population Division 2016 [cited 2016 Nov 18]. (http://www.un.org/en/development/desa/population/).
23. **Chowell G, *et al.*** The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *Journal of Theoretical Biology* 2004; **229**: 119–126.
24. **Legrand J, *et al.*** Understanding the dynamics of Ebola epidemics. *Epidemiology and Infection* 2007; **135**: 610–621.
25. **WHO Ebola Response Team.** Ebola virus disease in West Africa – the first 9 months of the epidemic and forward projections. *New England Journal of Medicine* 2014; **371**: 1481–1495.
26. **Althaus CL.** Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa. *PLoS Currents* 2014; **6**.
27. **Althaus CL, *et al.*** Ebola virus disease outbreak in Nigeria: transmission dynamics and rapid control. *Epidemics* 2015; **11**: 80–84.
28. **Barbarossa MV, *et al.*** Transmission dynamics and inal epidemic size of Ebola virus disease outbreaks with varying interventions. *PLoS ONE* 2015; **10**: e0131398.
29. **Gomes MF, *et al.*** Assessing the international spreading risk associated with the 2014 West African Ebola outbreak. *PLoS Currents* 2014; **6**.
30. **Pandey A, *et al.*** Strategies for containing Ebola in West Africa. *Science (New York, NY)* 2014; **346**: 991–995.
31. **Hsieh YH, *et al.*** Impact of visitors and hospital staff on nosocomial transmission and spread to community. *Journal of Theoretical Biology* 2014; **356**: 20–29.
32. **Agusto FB, Teboh-Ewungkem MI, Gumel AB.** Mathematical assessment of the effect of traditional beliefs and customs on the transmission dynamics of the 2014 Ebola outbreaks. *BMC Medicine* 2015; **13**: 96–112.
33. **Ajelli M, *et al.*** The 2014 Ebola virus disease outbreak in Pujehun, Sierra Leone: epidemiology and impact of interventions. *BMC Medicine* 2015; **13** (no pagination)(281).
34. **Browne C, Gulbudak H, Webb G.** Modeling contact tracing in outbreaks with application to Ebola. *Journal of Theoretical Biology* 2015; **384**: 33–49.
35. **Hsieh YH.** Temporal course of 2014 Ebola virus disease (EVD) outbreak in West Africa elucidated through morbidity and mortality data: a tale of three countries. *PLoS ONE* 2015; **10**: e0140810.
36. **Khan A, *et al.*** Estimating the basic reproductive ratio for the Ebola outbreak in Liberia and Sierra Leone. *Infectious Diseases of Poverty* 2015; **4**: 13.
37. **Kiskowski MA.** A three-scale network model for the early growth dynamics of 2014 west Africa Ebola epidemic. *PLoS Currents* 2014; **6**.
38. **Kucharski AJ, *et al.*** Measuring the impact of Ebola control measures in Sierra Leone. *Proceedings of the National Academy of Sciences of the United States of America* 2015; **112**: 14366–14371.
39. **Li Z, *et al.*** Dynamical analysis of an SEIT Epidemic model with application to Ebola virus transmission in Guinea. *Computational and Mathematical Methods in Medicine* 2015; **2015**: 582625.
40. **Merler S, *et al.*** Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *Lancet Infectious Diseases* 2015; **15**: 204–211.
41. **Valdez LD, *et al.*** Predicting the extinction of Ebola spreading in Liberia due to mitigation strategies. *Scientific Reports* 2015; **5**: 12172.
42. **Yamin D, *et al.*** Effect of Ebola progression on transmission and control in Liberia. *Annals of Internal Medicine* 2015; **162**: 11–17.
43. **Area I, *et al.*** On a fractional order Ebola epidemic model. *Advances in Difference Equations* 2015; **2015**: 278–289.
44. **Bellan SE, *et al.*** Ebola control: effect of asymptomatic infection and acquired immunity. *Lancet* 2014; **384**: 1499–1500.
45. **Dong F, *et al.*** Evaluation of Ebola spreading in West Africa and decision of optimal medicine delivery strategies based on mathematical models. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 2015; **36**: 35–40.
46. **Drake JM, *et al.*** Ebola cases and health system demand in Liberia. *PLoS Biology* 2015; **13**: e1002056.
47. **Fast SM, *et al.*** The role of social mobilization in controlling Ebola virus in Lofa County, Liberia. *PLoS Currents* 2015; **7**.
48. **White RA, *et al.*** Projected treatment capacity needs in Sierra Leone. *PLoS Currents* 2015; **7**.
49. **Camacho A, *et al.*** Temporal changes in Ebola transmission in Sierra Leone and implications for control requirements: a real-time modelling study. *PLoS Currents* 2015; **7**.
50. **Chowell G, *et al.*** Is West Africa approaching a Catastrophic phase or is the 2014 Ebola Epidemic slowing down? Different models yield different answers for Liberia. *PLoS Currents* 2014; **6**.
51. **Evans RJ, Mammadov M.** Dynamics of Ebola epidemics in West Africa 2014. *F1000Research* 2014; **3**: 319.
52. **Fasina FO, *et al.*** Transmission dynamics and control of Ebola virus disease outbreak in Nigeria, July to September 2014. *Euro Surveillance: Bulletin Europeen sur les maladies transmissibles = European Communicable Disease Bulletin* 2014; **19**: 20920.
53. **Fisman D, Tuite A.** Projected impact of vaccination timing and dose availability on the course of the 2014 West African Ebola Epidemic. *PLoS Currents* 2014; **6**.
54. **Fisman D, Khoo E, Tuite A.** Early epidemic dynamics of the West African 2014 Ebola outbreak: estimates derived with a simple two-parameter model. *PLoS Currents* 2014; **6**.

55. **Lewnard JA, *et al.*** Dynamics and control of Ebola virus transmission in Montserrado, Liberia: a mathematical modelling analysis. *Lancet Infectious Diseases* 2014; **14**: 1189–1195.
56. **Liu W, Tang S, Xiao Y.** Model selection and evaluation based on emerging infectious disease data sets including A/H1N1 and Ebola. *Computational and Mathematical Methods in Medicine* 2015; **2015**: 207105.
57. **Nishiura H, Chowell G.** Early transmission dynamics of Ebola virus disease (evd), West Africa, march to august 2014. *Eurosurveillance* 2014; **19**: 1–6.
58. **Rivers CM, *et al.*** Modeling the impact of interventions on an epidemic of Ebola in Sierra Leone and Liberia. *PLoS Currents* 2014; **6**.
59. **Shaman J, Yang W, Kandula S.** Inference and forecast of the current west African Ebola outbreak in Guinea, Sierra Leone and Liberia. *PLoS Currents* 2014; **6**.
60. **Shen M, Xiao Y, Rong L.** Modeling the effect of comprehensive interventions on Ebola virus transmission. *Scientific Reports* 2015; **5**: 1–17.
61. **Siettos C, *et al.*** Modeling the 2014 Ebola virus epidemic – agent-based simulations, temporal analysis and future predictions for Liberia and Sierra Leone. *PLoS Currents* 2015; **7**.
62. **Towers S, *et al.*** Mass media and the contagion of fear: the case of Ebola in America. *PLoS ONE* 2015; **10** (6): e129179, 1–13.
63. **Webb G, *et al.*** A model of the 2014 Ebola epidemic in West Africa with contact tracing. *PLoS Currents* 2015; **7**.
64. **WHOER team.** West African Ebola Epidemic after one year – slowing but not yet under control. *New England Journal of Medicine* 2015; **372**: 584–587.
65. **Vynnycky E, White R.** *An Introduction to Infectious Disease Modelling*. Oxford, New York: OUP, 2010.
66. **Kühnert D, Wu C-H, Drummond AJ.** Phylogenetic and epidemic modeling of rapidly evolving infectious diseases. *Infection, Genetics and Evolution* 2011; **11**: 1825–1841.
67. **Wu JT, *et al.*** Estimating infection attack rates and severity in real time during an influenza pandemic: analysis of serial cross-sectional serologic surveillance data. *PLoS Medicine* 2011; **8**: e1001103.
68. **MacIntyre CR, Chughtai AA.** Recurrence and reinfection – a new paradigm for the management of Ebola virus disease. *International Journal of Infectious Diseases* 2016; **43**: 58–61.
69. **Chughtai AA, Barnes M, Macintyre CR.** Persistence of Ebola virus in various body fluids during convalescence: evidence and implications for disease transmission and control. *Epidemiology and Infection* 2016; **144**: 1652–1660.
70. **Rodriguez LL, *et al.*** Persistence and genetic stability of Ebola virus during the outbreak in Kikwit, Democratic Republic of the Congo, 1995. *Journal of infectious diseases* 1999; **179** (Suppl. 1): S170–S176.
71. **Kreuels B, *et al.*** A case of severe Ebola virus infection complicated by Gram-negative septicemia. *New England Journal of Medicine* 2014; **371**: 2394–2401.
72. **Emond RT, *et al.*** A case of Ebola virus infection. *British Medical Journal* 1977; **2**: 541–544.
73. **Rowe AK, *et al.*** Clinical, virologic, and immunologic follow-up of convalescent Ebola hemorrhagic fever patients and their household contacts, Kikwit, Democratic Republic of the Congo. *Commission de Lutte contre les Epidemies a Kikwit. The Journal of Infectious Diseases* 1999; **179** (Suppl. 1): S28–S35.
74. **Lyon GM, *et al.*** Clinical care of two patients with Ebola virus disease in the United States. *New England Journal of Medicine* 2014; **371**: 2402–2409.
75. **Christie A D-WG, *et al.*** Possible sexual transmission of Ebola virus – Liberia, 2015. *MMWR Morbidity and Mortality Weekly Report* 2015; **64**: 479–481.
76. **Abbate JL, *et al.*** Potential impact of sexual transmission on Ebola virus epidemiology: Sierra Leone as a case study. *PLoS Neglected Tropical Diseases* 2016; **10**: e0004676.
77. **Vinson JE, *et al.*** The potential for sexual transmission to compromise control of Ebola virus outbreaks. *Biology Letters* 2016; **12**: 20151079; 1–14.