

On the reliability of predictions on Covid-19 dynamics: A systematic and critical review of modelling techniques

Janyce Eunice Gnanvi, Kolawolé Valère Salako, Gaëtan Brezesky Kotanmi, Romain Glèlè Kakai PhD *

Laboratoire de Biomathématiques et d'Estimations Forestières, Université d'Abomey-Calavi, 04 BP 1525, Cotonou, Benin

ARTICLE INFO

Article history:

Received 16 September 2020

Received in revised form 29 December 2020

Accepted 29 December 2020

Available online 12 January 2021

Handling editor: Dr. J Wu

Keywords:

Predictions

Accuracy

Precision

SARS-CoV-2

Pandemic

Ratio

ABSTRACT

Since the emergence of the novel 2019 coronavirus pandemic in December 2019 (COVID-19), numerous modellers have used diverse techniques to assess the dynamics of transmission of the disease, predict its future course and determine the impact of different control measures. In this study, we conducted a global systematic literature review to summarize trends in the modelling techniques used for Covid-19 from January 1st 2020 to November 30th 2020. We further examined the accuracy and precision of predictions by comparing predicted and observed values for cumulative cases and deaths as well as uncertainties of these predictions. From an initial 4311 peer-reviewed articles and pre-prints found with our defined keywords, 242 were fully analysed. Most studies were done on Asian (78.93%) and European (59.09%) countries. Most of them used compartmental models (namely SIR and SEIR) (46.1%) and statistical models (growth models and time series) (31.8%) while few used artificial intelligence (6.7%), Bayesian approach (4.7%), Network models (2.3%) and Agent-based models (1.3%). For the number of cumulative cases, the ratio of the predicted over the observed values and the ratio of the amplitude of confidence interval (CI) or credibility interval (CrI) of predictions and the central value were on average larger than 1 indicating cases of inaccurate and imprecise predictions, and large variation across predictions. There was no clear difference among models used for these two ratios. In 75% of predictions that provided CI or CrI, observed values fall within the 95% CI or CrI of the cumulative cases predicted. Only 3.7% of the studies predicted the cumulative number of deaths. For 70% of the predictions, the ratio of predicted over observed cumulative deaths was less or close to 1. Also, the Bayesian model made predictions closer to reality than classical statistical models, although these differences are only suggestive due to the small number of predictions within our dataset (9 in total). In addition, we found a significant negative correlation ($\rho = -0.56$, $p = 0.021$) between this ratio and the length (in days) of the period covered by the modelling, suggesting that the longer the period covered by the model the likely more accurate the estimates tend to be. Our findings suggest that while predictions made by the different models are useful to understand the pandemic course and guide policy-making, some were relatively accurate and precise while other not.

© 2021 The Authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail address: romain.glelekakai@fsc.uac.bj (R. Glèlè Kakai).

Peer review under responsibility of KeAi Communications Co., Ltd.

1. Introduction

The current outbreak of the novel coronavirus SARS-CoV-2 which original epicenter was Wuhan city, has spread to many other countries (Velavan and Meyer, 2020) causing devastating public health impacts across the world. The novel coronavirus spilled over from the non-human primate population into humans on the Huanan seafood market in Wuhan, China. Since March 2020, while new cases in China appears to have settled down, the number of cases is exponentially growing in the rest of the world (Toda, 2020) with Africa the least affected continent.

As with the two other coronaviruses that caused major outbreaks in humans in recent years (namely, Severe Acute Respiratory Syndrome and the Middle Eastern Respiratory Syndrome (WHO, 2020; Yin & Wunderink, 2018), Covid-19 is transmitted from human-to-human through direct contact with contaminated objects or surfaces and through inhalation of respiratory droplets from both symptomatic and a-symptomatic infectious humans (Bai, Yao, et al., 2020).

In the absence of a safe and effective vaccine or antivirals, strategies for controlling and mitigating the burden of the pandemic are focused on Non-Pharmaceutical Interventions (NPI), such as social-distancing, contact-tracing, quarantine, isolation, and the use of face-masks in public (Ngonghala et al., 2020). Though many countries rely on those mitigation measures which help to slow the spread of the pandemic (Taboe et al., 2020), researchers, across various medical, public health and modelling disciplines, are actively engaged in efforts to understand the epidemiology of the disease. Modelling novel coronavirus disease has then become of extreme importance. Many researchers around the world have studied the patterns of Covid-19 pandemic and several mathematical, computational, clinical and examination studies have been put forward for modelling, prediction, treatment and control of the disease (Ngonghala et al., 2020; Taboe et al., 2020; Achoki et al., 2020; Bartolomeo et al., 2020; Cao et al., 2020; Ceylan, 2020; Kim et al., 2020; Pasayat et al., 2020; Shaikh et al., 2020; Tang, Bragazzi, et al., 2020; Xiao et al., 2020; Zhao, Gao, et al., 2020; Ziauddeen et al., 2020). This growing interest of scientists has resulted in a deluge of studies predicting the dynamics of Covid-19, and summarizing trends in these studies is necessary. Some studies (e.g. Achoki et al., 2020; Roda et al., 2020) reported the difficulty of current models to accurately predict the Covid-19 pandemic. Roda et al. (2020) showed that non-identifiability in model calibration using data on confirmed cases is a main source of large variation in model predictions. Other studies (Achoki et al., 2020; Bai, Gong, et al., 2020; Jewell et al., 2020) have raised the issues of data quality that is necessary for accurate predictions. The type of models used could also affect the accuracy of predictions (Ceylan, 2020; Wu, Darcet, Wang, & Sornette, 2020).

Here we conducted a systematic and critical review of studies published between January 1st and November 30th 2020 on Covid-19 to (1) summarize trends in the modelling techniques used to predict Covid-19 cases and deaths, and (2) assess the reliability of predictions of Covid-19 cases and deaths. The overarching goal is to determine and discuss to what extent studies accurately predict Covid-19 cases and deaths and whether some differences exist among modelling techniques.

2. Methods

2.1. Article search and selection

Relevant scientific databases, such as Pubmed, medRxiv and Google Scholar were used to search for models for COVID-19 transmission. The following keywords were used: “Coronavirus”, “Covid-19”, “Corona”, “SARS viruses”, “Sars-CoV-2”, OR “2019-nCoV” and in combination (i.e. AND) with “Model/modelling” “Prediction/Predicting”, “Dynamics”, “Estimates/Estimations/Estimating” OR “Forecast/Forecasting”. The time period covered was from January 1st 2020 to November 30th 2020. The bibliographies of retrieved studies as well as bibliographies of current reviews and texts were searched for additional relevant studies. From an initial list of 4291 articles, 242 were finally included in the systematic review. Selection of articles included in the systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines as illustrated on Fig. 1. The articles included were those related to Covid-19 dynamics and are model-based. All studies on covid-19 dynamics that are not model-based were excluded as were non-English language studies. Articles dealing with models for assessing risk factor to die from Covid-19 based on socio-economic factors, models for effective patient diagnosis, models on individual human behavior with regards to control measures (e.g. lockdown), etc. were excluded (see Fig. 1).

2.2. Literature synthesis and analysis

For each of the 242 papers selected (see supplementary material 1 for the full list), the data extracted were: the country for which the modelling study was conducted, the published or unpublished status of the study, the time period covered by the data (in number of days), the topics addressed in the study, the modelling techniques used, and whether the modelling was data-driven or not. We also noted whether the study accounted for asymptomatics, pre-symptomatics, both asymptomatics and pre-symptomatics, or none of these classes of individuals. This aspect was included because of the prominent role that asymptomatics and pre-symptomatics play in the transmission of the disease (He, Guo, Mao, & Zhang, 2020). When the study made predictions, we further recorded the predicted values of the cumulative number of cases, the predicted values of the cumulative number of deaths, the date at which the predicted values of the number of cumulative cases will be observed, and the uncertainty parameters around the predictions (95% Confidence Interval – CI or 95% Credibility Interval – CrI). The data analyses considered three aspects. The first aspect was related to the geographical coverage (continents and number of

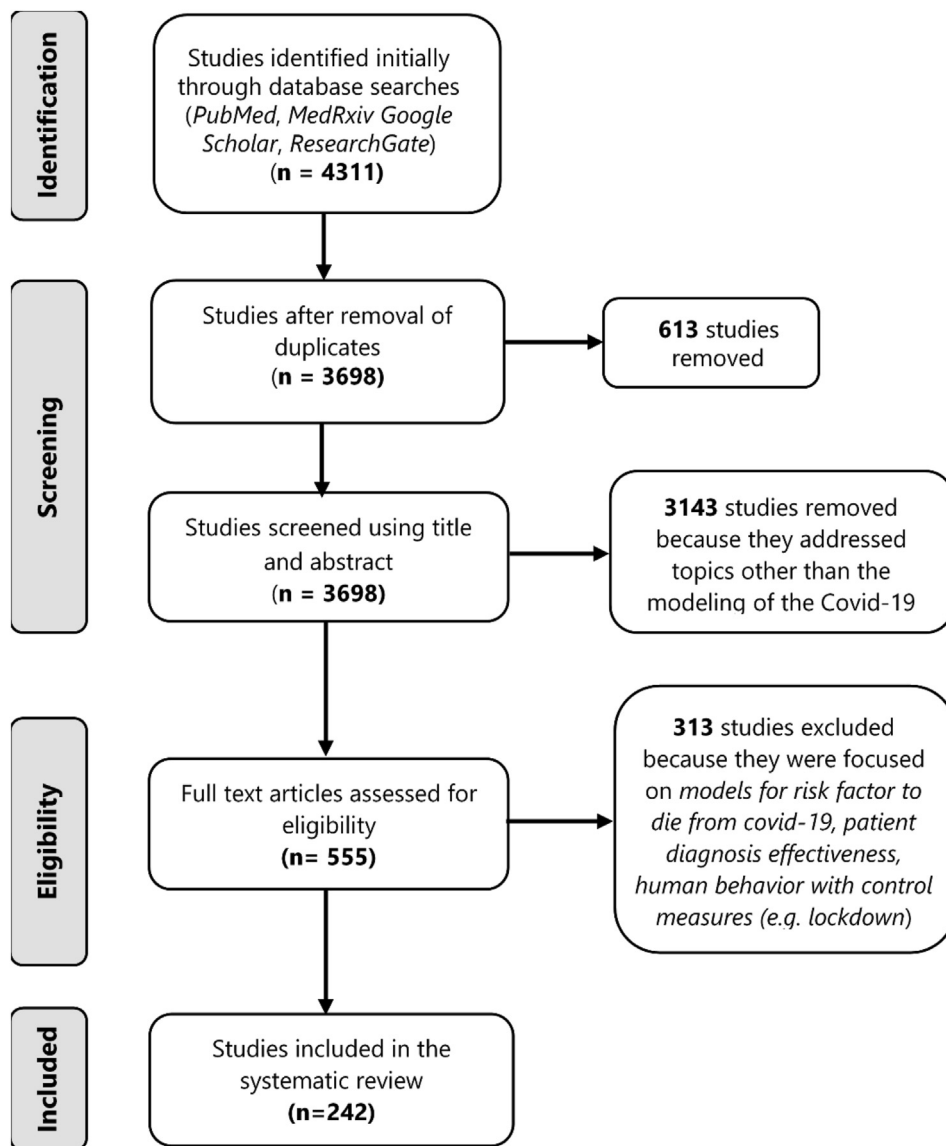


Fig. 1. PRISMA flow diagram of the selection process of the 242 studies included in the systematic review.

countries covered per continent) and topics addressed in the studies, whether the modelling was data-driven and include symptomatics, pre-symptomatics, or not; we used count and relative frequencies to describe this trend. The second aspect was related to the modelling techniques and was also addressed using count and relative frequencies after grouping modelling techniques in relatively similar groups. The third aspect was the accuracy and precision of the predictions made by studies. Accuracy refers to how close a prediction is to the true value but precision refers to how certain is the prediction (Stallings & Gillmore, 1971). For this, we used three parameters; the first is the ratio between the value predicted and the value actually observed on the day on which the prediction was made. This ratio is a measure of accuracy of the prediction. A value close to 1 indicates that the prediction was accurate. Values less or larger than one indicates underestimation or overestimation, respectively. The second parameter which is a measure of precision was the ratio between the amplitude of the uncertainty parameter (95% CI or 95% CrI) and the central value. For studies that used statistical methods, the uncertainty parameter is the 95% CI. For studies which used Bayesian methods, the 95% CrI is the uncertainty parameter. The uncertainty parameters indicate that given the observed data, the prediction has 95% probability of falling within this range. This ratio is an estimate of the accuracy of the predictions. A value of 1 for this ratio indicates that the amplitude is larger as the central value. Smaller values indicate more accurate prediction (i.e. prediction with low uncertainty). The values of these ratios were plotted against studies, and type of models used. We additionally plotted these ratios against the number of days covered by the studies. We expected that the longer the period of data considered, the smaller the values of these ratios. The third

parameter was whether the value actually observed in a study was within or out of the 95% CI or CrI of the predictions; this was used to compute the proportion of predictions for which the number of cumulative cases or deaths actually observed are within the 95% CI or CrI of the predictions.

3. Results

3.1. Characteristics of the papers selected: publication status, geographical coverage and topics addressed

Of the 242 papers reviewed, 33.88% were preprints. The largest part of the studies focused on Asia (78.93%), especially on China (51) and on India (24) (Fig. 2a). However, the coverage (percentage of countries in a continent where a study was carried out) was higher in Europe where we found studies conducted in 81.81% (36 countries out of the 44) of European countries with Italy (39 studies), France (25 studies) and Spain (22 studies) being the countries where more studies were done. From our sampled studies, 35 focused on African countries either at country level (9 in Nigeria, 7 in South Africa) or region level (i.e. west, east, north, south, or central), or the whole continent level. Some studies did not focus on a specific country but on an entire continent or part of continent (e.g. Taboe et al., 2020). The selected studies covered 18 out of 35 countries in the America continent, with the United States (31 studies) and Brazil (18 studies) as the countries on which most of the studies were carried out (Fig. 2b, see Supplementary material 2 for the list of countries per continent).

The studies addressed diverse topics which can be classified into four groups. The studies were primarily designed for three main purposes: study the dynamics of the transmission of Covid-19 with (43.80%) or without (20.66%) attempting to predict the course of the pandemic (cumulative cases, deaths, hospitalized cases etc.); estimate key epidemiological parameters of the transmission of Covid-19 (12.40% of studies), and evaluate the impact of control measures on the transmission of Covid-19 (20.66% of studies). The burden of healthcare systems was also assessed by 1.24% of the studies (Fig. 3a). With regards to the impact of control measures on the transmission of Covid-19, the following eight control measures were considered in the studies: face mask, quarantine, case isolation, contact tracing, social distancing, school closure, workplace distancing, restriction on international air travel and lockdown. Fig. 3b shows the distribution of studies that assessed the impacts of the above control measures. The most assessed measures were quarantine (19.42%), social distancing (18.60%) and lockdown (18.18%). Studies focused on six epidemiological parameters (Fig. 3c) and the most estimated parameter was the reproduction number (38.43%).

3.2. Modelling techniques

Several modelling techniques were used which we classified into five main groups, including Agent-based models, Machine Learning and Artificial Intelligence (AI) based approach, Bayesian models, Compartmental models, Network models, Statistical models, and Hybrid models (they refer to models that combine two or more approaches). Supplementary material 3 gives details of the models per group.

A compartmental model, also broadly known as population-based model, is a model that stratifies the population into different compartments, such as different health states (e.g. Susceptible, Exposed, Infected, Quarantined, Recovered, Dead, etc.) for the modelling. Compartments are assumed to represent homogeneous sub-populations within which the entities being modelled—such as individuals or patients—have the same characteristics (Porgo et al., 2019). Compartmental models

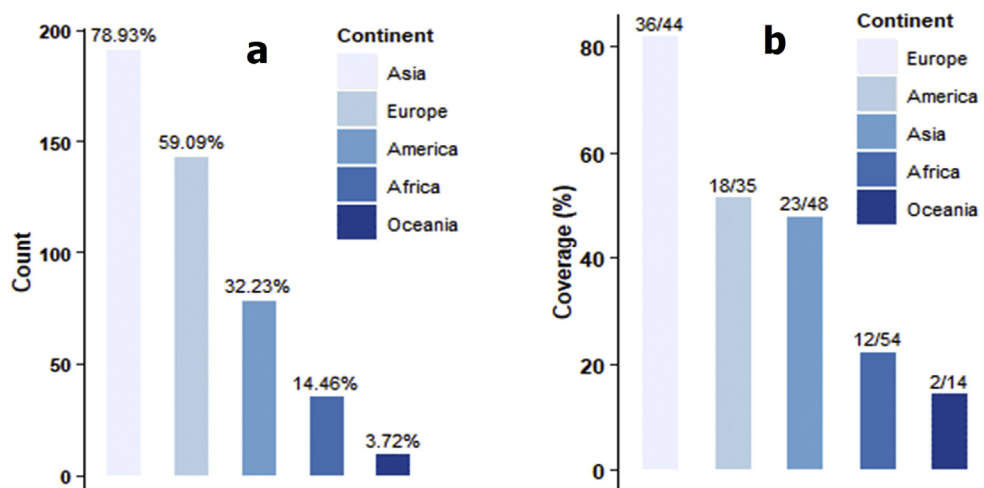


Fig. 2. Distribution of studies across continents (a) and countries coverage across continents (b).

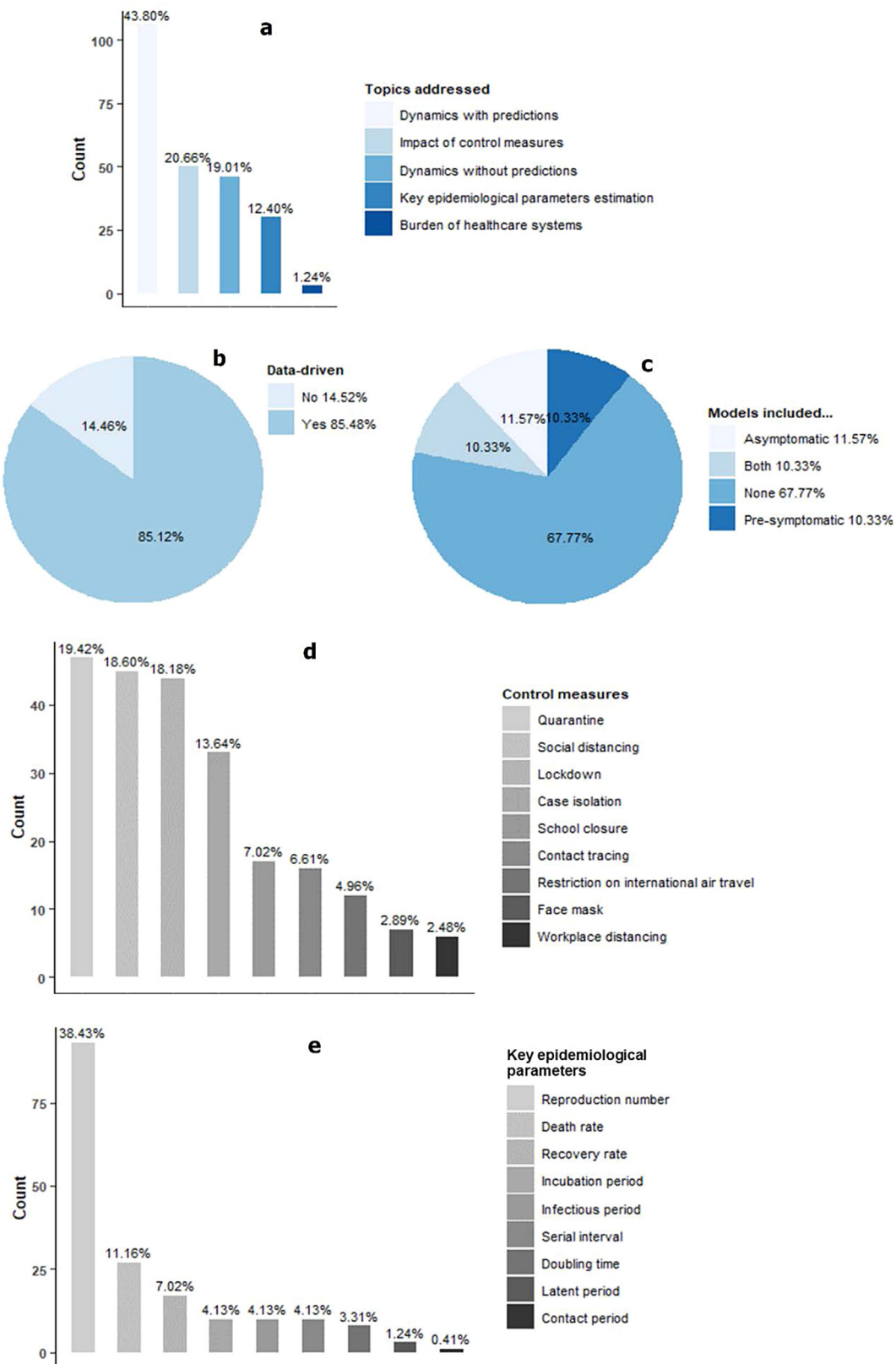


Fig. 3. Distribution of studied articles according to (a) whether the models used were data-driven or not, (b) whether the models included asymptomatic and/or pre-symptomatic individuals, (c) the topic addressed, (d) the control measures assessed, and (e) the key epidemiological parameters estimated.

were the most used (46.15%) regardless the topic addressed (Fig. 4a). The classical SEIR model was considered in many studies (e.g. Nyabadza et al., 2020; Taboe et al., 2020; Wang et al., 2020) (22.30%). Some studies however adopted an improved SEIR model. Considering the traditional SEIR model as not realistic, Zhao, Li, et al. (2020) considered an improved SEIR model by introducing both quarantine status and intervention measures. Liu et al. (2020) incorporated three important elements of Covid-19 to the classical SEIR model to estimate epidemiological parameters of the disease in South Korea, Italy, and Spain: (1) the number of asymptomatic infectious individuals (with very mild or no symptoms), (2) the number of symptomatic reported infectious individuals (with severe symptoms), and (3) the number of symptomatic unreported infectious individuals (with less severe symptoms). The SEDQIR model based on SEIR model was established by Cao et al. (2020) with D, the suspected cases of infection or potential victims and Q, the diagnosed and quarantined. Moreover, de Camino-Beck (2020) introduced a compartment C and then developed the SEICR model. The compartment C was for confined individuals, that was,

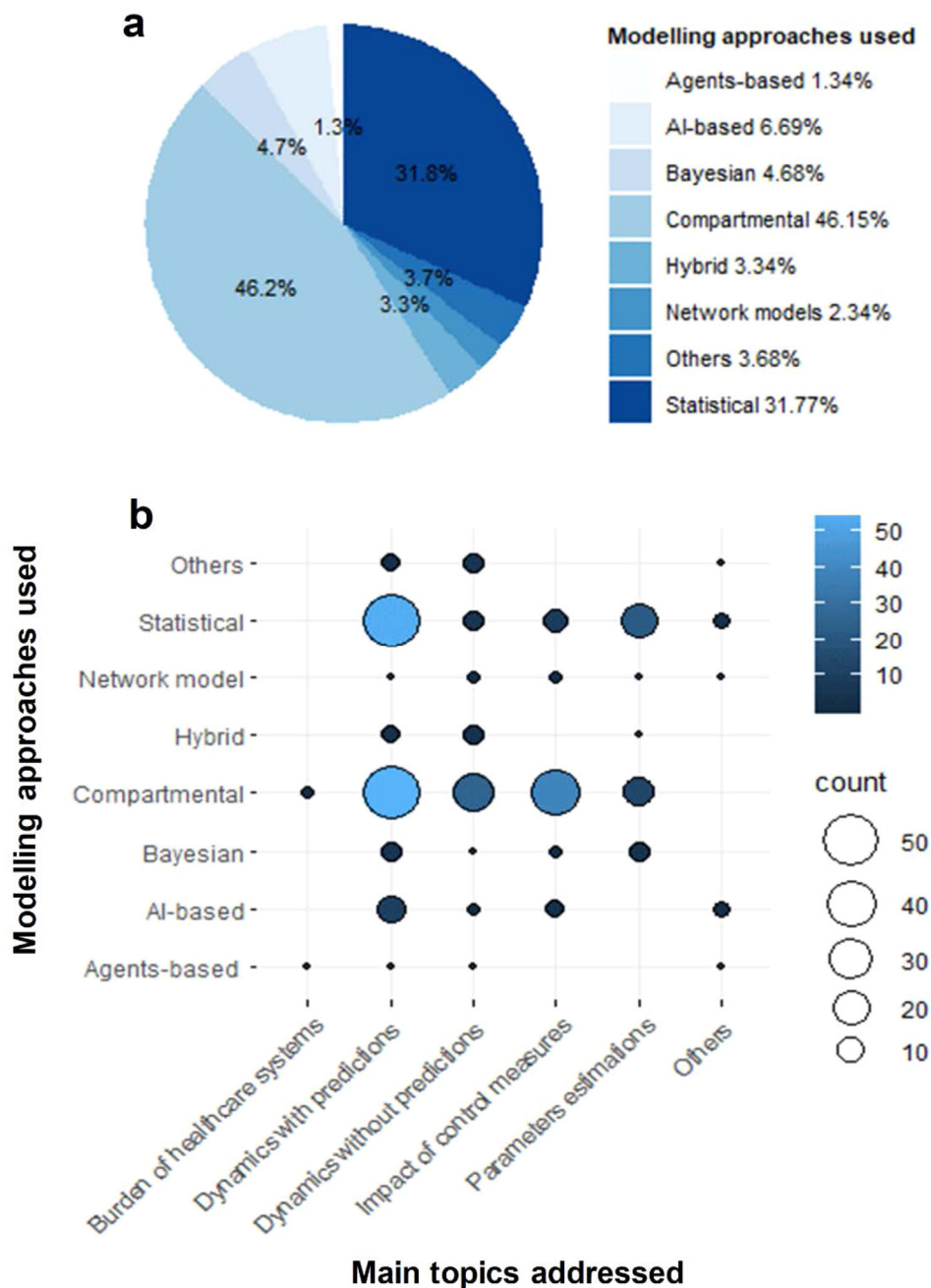


Fig. 4. Diversity of modelling techniques used for Covid-19 (a), and topics addressed with the modelling techniques (b).

individuals whose movement are restricted and effectively removed from the susceptible population by strong NPI, like lock downs, closure of retail and entertainment, parks, and vehicular restrictions (de Camino-Beck, 2020).

Statistical modelling is an approach for developing and testing theories by way of causal explanation, prediction, and description. In many disciplines there is near-exclusive use of statistical modelling for causal explanation and the assumption that models with high explanatory power are inherently of high predictive power (Shmueli et al., 2010). Contrary to the compartmental models, statistical models often model one state (i.e. one compartment at the time) and do not often consider the flow of individuals among the different states. Statistical models were used by 31.77% of the studies. The most used statistical models were growth models which aimed to model change over time. About eighteen percent (17.6%) of studies used growth models (Exponential growth model, Generalized-growth model, Logistic growth model, Richard growth models etc.) (Wu et al., 2020; Hermanowicz, 2020; Shim et al., 2020). About tenth of the studies used time series models (9.46%). Time series models that were used included ARIMA models (Ceylan, 2020), VAR models (Silva et al., 2020), exponential smoothing models (Elmousalami & Hassanien, 2020) etc. One of the most common statistical modelling techniques used were the regression models. Statistical models, such as regression models, are typically phenomenological and describe the statistical relationship or association between different model variables (Porgo et al., 2019). Less than four percent (3.5%; 8 out of 230) of the studies used a regression model (linear regression, polynomial regression etc.) (Chauhan et al., 2020; Yang et al., 2020). Additional modelling techniques were used, including parametric distributions fitting models (Zhang, Litvinova, et al., 2020; Zhao, Gao, et al., 2020), exponential decay model (Bartolomeo et al., 2020), least square error (LSE) model (Ahmadi et al., 2020). AI-based models were used in approximately 7% of the studies whereas Bayesian approach was used in 5% of studies (Fig. 4a).

Agent-based and networks models were the least represented among the categories of models used, 1.34% and 2.34%, respectively (Fig. 4a). Contrary to population-based models, in agent-based and networks models, also known as individual-based models, individuals are considered to typically interact on a network structure and exchange infection stochastically, thus allowing to consider individual heterogeneity in the modelling process, which seems to be more realistic given the widespread heterogeneity of human individuals (Willem et al., 2017).

The different models have addressed different topics. Statistical and Bayesian methods have been more frequently used to estimate epidemiological parameters whereas compartmental models have been used more frequently to assess the dynamics of the disease (Fig. 4b). The dynamics of Covid-19 transmission have been analysed in numerous studies which have tried to predict the spread of the disease using the above models. While the predictions made by some of the models used are close to the observed reality, predictions made by other models have proven to be inaccurate. For example, on the basis of data from February 25 to March 21 and using the “Alg-Covid-19” Model, Hamidouche (2020) estimated that the number of cases in Algeria will exceed 1000 on the 35th day of the epidemic (March 31st, 2020), 5000 on the 42nd day (April 7th) and it will double and reach 10,000 on 46th day of the epidemic (April 11th) whereas until April 11th, the number of cases in the country was just 1825 (Worldometer, 2020), probably because of control measures.

3.3. Reliability of predictions on Covid-19 dynamics

3.3.1. Predictions of the number of cumulative cases

From the selected studies, 29 predicted the number of cumulative cases for the coming days and 92 predictions were made in total (Fig. 5a). The ratio between the predicted value and the observed value was calculated as a criterion for the accuracy of the predictions. Results showed that the predicted values were higher than observed values for 38.04% of the estimations (35 of the 92) and lower than the observed values for 61.96% remaining estimations (57 of the 92). Thirty-three predicted values actually departed from the values actually observed (ratio of the predicted value to the observed value less than 0.8 or greater than 1.2). There was no evidence of strong difference in the value of this ratio among the categories of models used to predict the future values of the cumulative cases (Fig. 6b). Relatively large variation of the ratio was also observed among predictions within most of the categories of models (Fig. 5b). Models that accounted for pre-symptomatics seem to have lower ratio, but data were not enough for a statistical significance test across (Fig. 5c). There was a large variation of the ratio among models that were parameterized data (Fig. 5d). The regression line between the time periods (expressed as number of days) covered by the data used in the selected studies and the ratio ($\beta = 0.004$; $p\text{-value} = 0.100$) was not significant, indicating no statistical evidence of more accurate estimation with longer time periods (Fig. 5e).

Confidence interval (CI) or credibility interval (CrI) are essential measurements of precision in parameter estimations. As an indicator of precision, the ratio of the amplitude of the 95% CI or 95% CrI and the predicted value was calculated to also assess reliability of the predicted values. Overall, very few studies have reported CI or CrI. Only 5.79% of studies (14 out of 242) have reported CI or CrI for the predicted number of cumulative cases. These 14 studies provided 20 predictions of which, one was greater than 1 (5%) and 19 were less than 1 (95%) (Fig. 6a). This ratio seems relatively lower for statistical models compared to compartmental models, indicating relatively more precise predictions for statistical models (Fig. 6b). However, this difference cannot be confirmed statistically since the compartmental, and the statistical models were used for 4, and 13 predictions respectively, that we judged not enough for a robust statistical significance test. More data would be needed to better make this comparison. Including either asymptomatics, or pre-symptomatics or none of these classes in the modelling does not seem to affect the precision of the predictions (Fig. 6c). There was not enough information in our dataset to compare

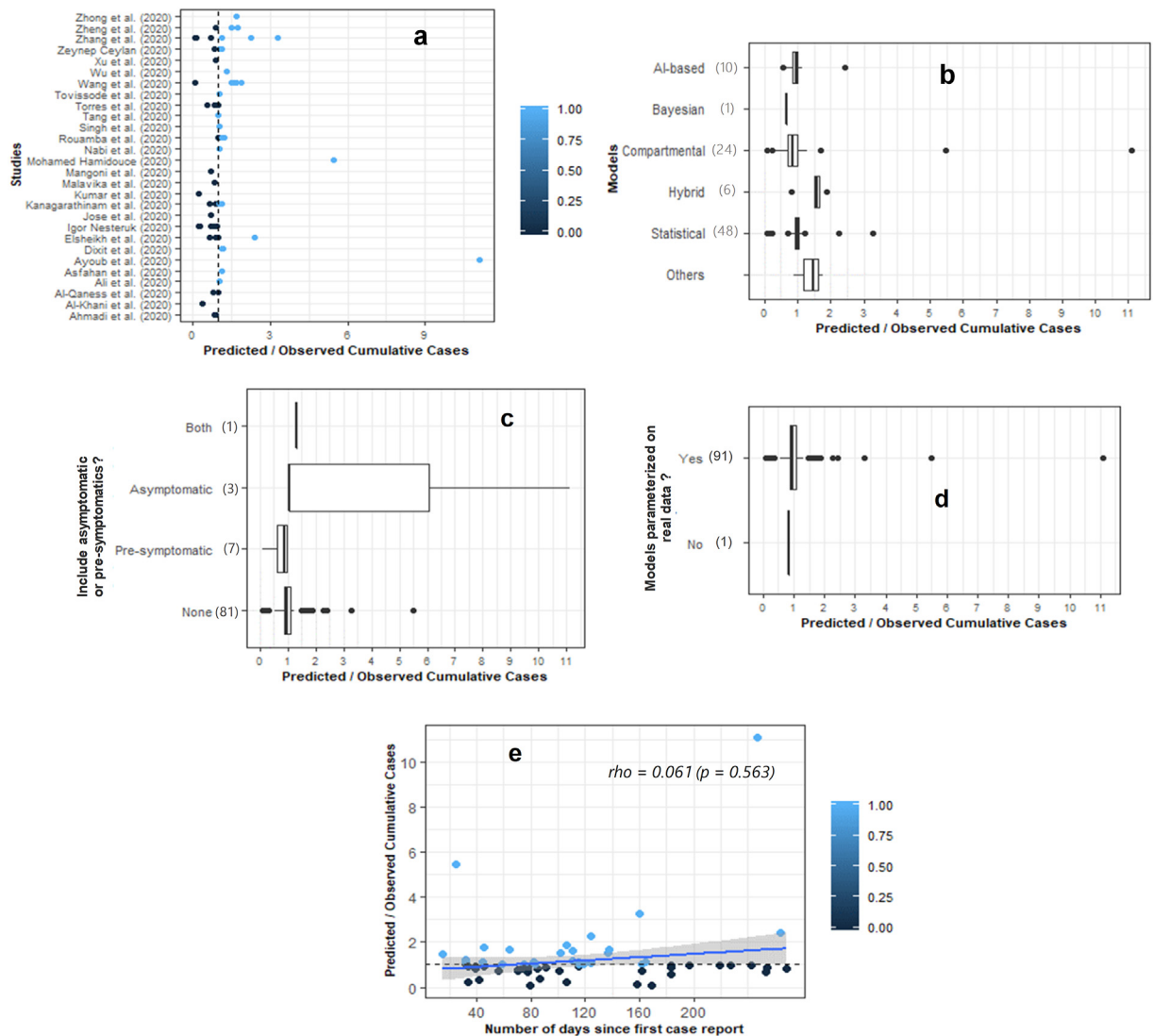


Fig. 5. Accuracy of the models' predictions: ratio of the number of cumulative cases predicted over the actual number of cumulative cases observed in 33 studies (a), across types of models (b), according to whether models included asymptomatic or pre-symptomatics (c), according to whether models were parameterized based on real data (d), and in relationships to the number of days since the first case was reported in concerned countries (e). Others (see appendix C). Values in parentheses in (b) and (d) represent the number of predictions found for each case.

this ratio between models parameterized on real data and models that were not parameterized on real data (Fig. 6d). This ratio decreases with the length of the period (in number of days) covered by the data used for the estimation, although it was not significant (linear regression analysis: $\beta = -0.001$; p -value = 0.242, Fig. 6e).

The third parameter of reliability was based on the 95% CI or CrI provided for each prediction of the number of cumulative cases. This parameter checked whether the true value (i.e. the value actually observed for the prediction) is within the 95% CI or CrI provided for the prediction (Fig. 7). Fig. 7 shows a graphical representation of the cross-tabulation of the number of predictions that presented a 95% CI or CrI (20 in total) and whether or not the value actually observed belongs to the 95% CI or CrI. This figure shows that 75% (15 out of 20) of the values actually observed were within the 95% CI or CrI provided for the prediction. 65% of these (13 out of 20) were predictions made based on statistical models (Fig. 7).

3.3.2. Predictions of the number of cumulative deaths

Only nine of the 242 selected studies (3.72%) made predictions of the number cumulative deaths due to covid-19 pandemic and 17 predictions were made (see Fig. 8a). For about half (52.94%) of the predictions, the ratio of predicted over the actual number of cumulative deaths was lower or close to 1 (Fig. 8a). One prediction largely exceeded (more than 6 times) the actual number of deaths (Fig. 8a). This ratio seems to be relatively lower (and also lower than 1) for Bayesian

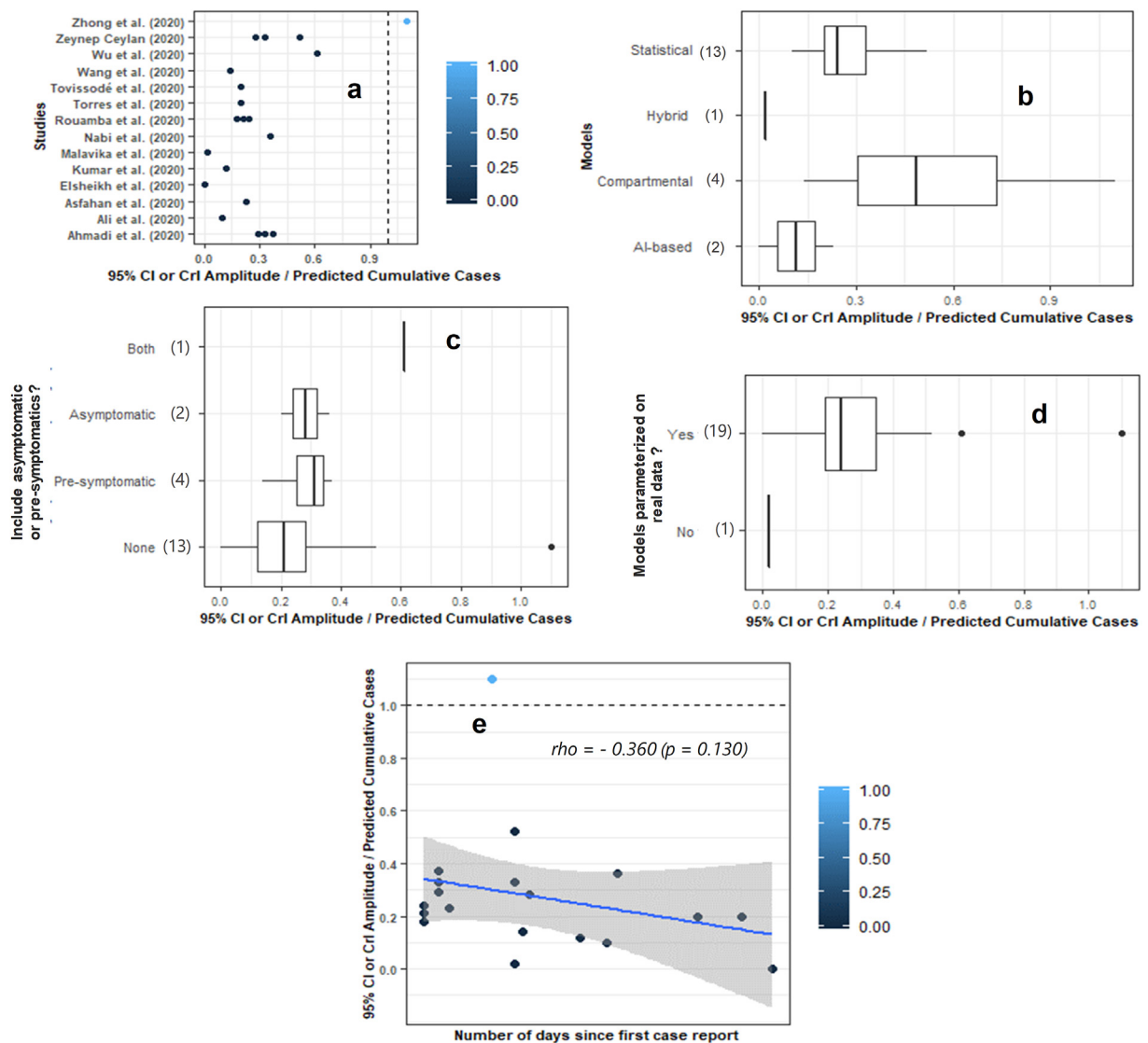


Fig. 6. Precision of the models' predictions: ratio of the amplitude of the 95%CI or 95%CrI of the predicted cumulative number of cases over the predicted cumulative number of cases in 14 studies (a), across types of models (b), according to whether models included asymptomatic or pre-symptomatics (c), according to whether models were parameterized based on real data (d), and in relationships to the number of days since the first case was reported in concerned countries (e). Values in parentheses in (b) and (d) represent the number of predictions found for each case.

models than for statistical models where this ratio was large than one, thus suggesting relatively more accurate predictions with Bayesian models (Fig. 8b), although these differences are only suggestive due to the small size of the data. A greater number of predictions than we found in this study would be needed for robust significance test. Whether the models included asymptomatic or pre-symptomatics individuals does not seem to affect this ratio (Fig. 8c). Nevertheless, there was a significant negative correlation between this ratio and the number of days of the first infection in the target country, suggesting that the more data used to make the estimates cover a large period of time, the more accurate the estimates tend to be (Fig. 8d).

Among the above nine studies, seven reported the 95% CI or CrI for their predictions and they made ten predictions of the number cumulative deaths (Fig. 9a). For six of the ten predictions (60%), the ratio of the amplitude of the 95% CI or CrI over the predicted number of cumulative deaths was lower than 1 (Fig. 9a). Two predictions had values between 4 and 7 for this ratio (Fig. 9a). There was no evidence of difference for this ratio among categories of models, nor according to whether the models considered asymptomatic or pre-symptomatic individuals (Fig. 9 b, c). There was also no correlation of this ratio with the length of period of time, the more accurate the estimates tend to be (Fig. 9d).

Fig. 10 shows the graphical representation of the cross-tabulation of the number of predictions of the cumulative number of deaths that presented a 95% CI or CrI (10 in total) and whether or not the value actually observed belongs to the 95% CI or

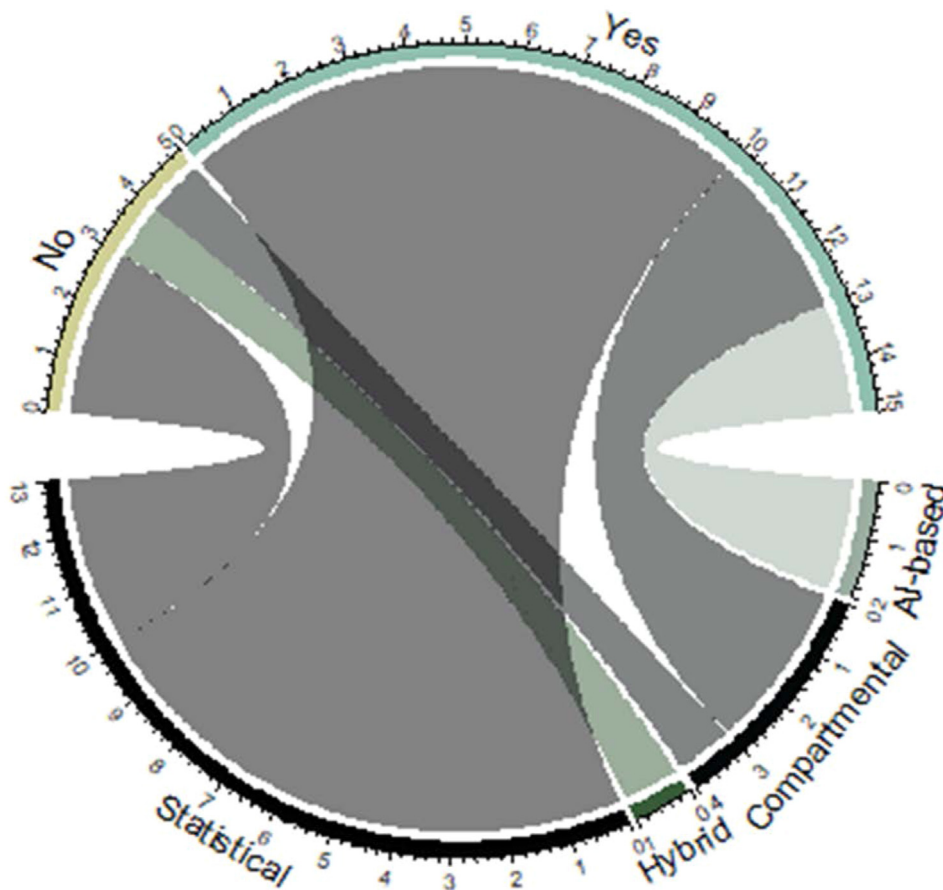


Fig. 7. Distributions of predictions (20) of the cumulative number of cases according to whether the values actually observed for the predictions fall within the 95%CI or 95%CrI of the prediction.

CrI. This figure shows that 60% (6 out of 10) of the values actually observed were within the 95% CI or CrI provided for the predictions. Among this, four were from statistical models (out of five predictions with statistical models), one from compartmental models (out one prediction with compartmental model), and one from Bayesian models (out of three predictions with Bayesian models). This might indicate a more reliable prediction with statistical model, but a greater number of predictions per group of models would also be needed for a more robust significance test.

4. Discussion

4.1. Modelling techniques applied to Covid-19 dynamics

The novel coronavirus pandemic (Covid-19) is causing devastating public health and socio-economic burden in affected areas. Understanding current patterns of the pandemic spread and forecasting its long-term trajectory is essential in guiding policies aimed at curtailing the pandemic (Taboe et al., 2020; Tang, Wang, et al., 2020). This situation induces a demand to the mathematical epidemiologist community (Rhodes et al., 2020) for revealing models of outbreak dynamics, which have not only explanatory but also a predictive potential while the outbreak is in an active phase. These models are aimed at the fast estimations of the future Covid-19 impact on the population, measures required from the public health system and effectiveness of different control measures (Postnikov, 2020). A wide-range of models were used; some were extremely simple models while others are more sophisticated. Most studies focused on compartmental models, SIR and SEIR models, to estimate the transmission dynamics and make predictions about the future growth of the pandemic. We found that the SIR model performs better than the SEIR model in representing the information contained in the confirmed-case data though Jewell et al. (2020) suggested that simpler models may provide less valid forecasts because they cannot capture complex and unobserved human mixing patterns and other time varying characteristics of infectious disease spread. Our finding is in line with that of Roda et al. (2020) who reported that predictions using more complex models may not be more reliable compared to using a simpler model. The models used included also statistical models, Bayesian models, Artificial Intelligence based

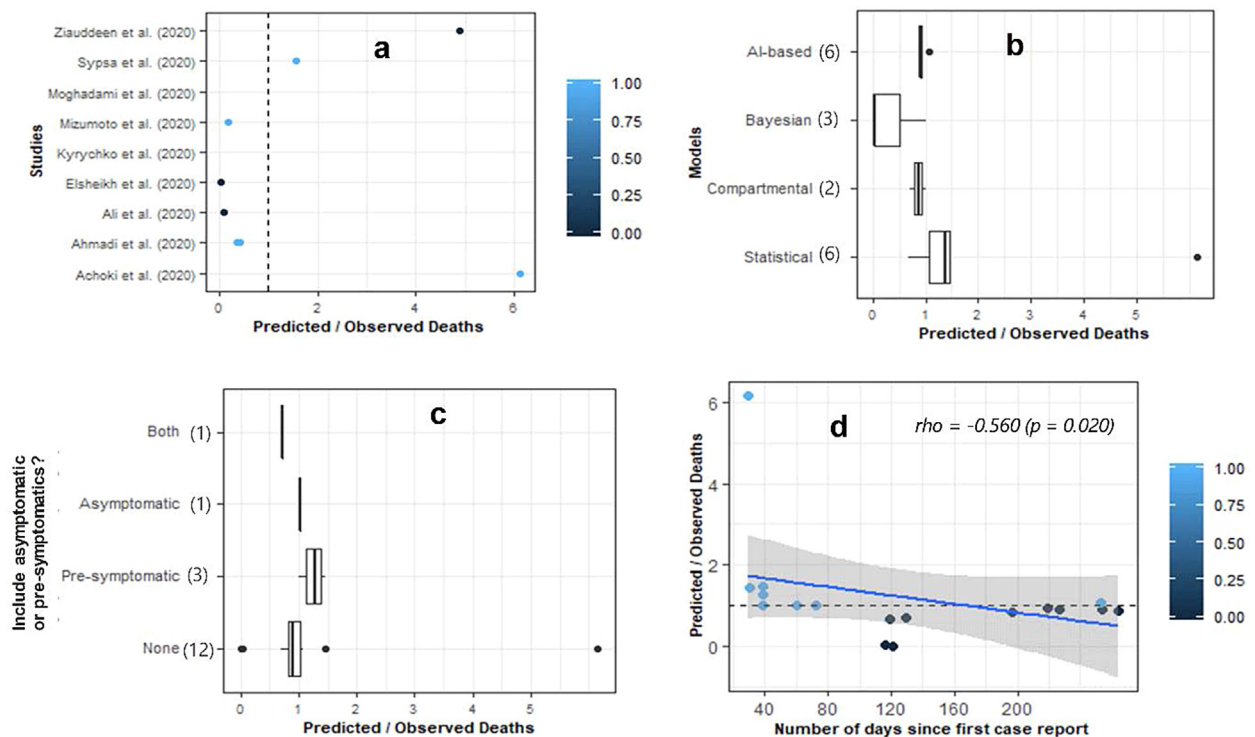


Fig. 8. Accuracy of the models' predictions: ratio of the number of cumulative deaths predicted over the actual number of cumulative deaths observed in nine studies (a), across types of models (b), according to whether models included asymptomatic or pre-symptomatics (c), and in relationships to the number of days since the first case was reported in concerned countries (d). Values in parentheses in (b) and (c) represent the number of predictions found for each case.

model and Hybrid models (Taboe et al., 2020; Ziauddeen et al., 2020; Linka et al., 2020; Stübinger & Schneider, 2020; Tian et al., 2020).

4.2. Reliability of predictions on Covid-19 dynamics

The studies included in this review focused more on the prediction of the cumulative number of cases than the deaths caused by Covid-19. The models used for predicting the dynamics of the Covid-19 were mainly compartmental (SIR and SEIR models) and statistical (Linear regression model, time series model, growth models). We did not find evidence for difference across models while some predicted values that far exceeded true values. However, our findings should be considered with caution as for the reviewed studies, the number of estimations were not fairly distributed across models. The simplest SIR model is used to predict diseases in which individuals can obtain permanent immunity after infection and is only applicable when there is a non-drug prevention intervention (Bai, Gong, et al., 2020). This model has shown better predictive performance relatively to the SEIR model for number of cases forecasting. Our findings are contrary to that of Bai, Gong, et al. (2020) which argued that the estimated numbers of infected people far exceed reported cases in the available literature which used the SIR model and suggest more the use of complex models. Bayesian models were also used by Ziauddeen et al. (2020) and Mizumoto et al. (2020). Though AI based models were used in few studies of our review, it is shown that the ongoing development in AI has significantly improved prediction, and forecasting for the Covid-19 pandemic (Lalmuanawma et al., 2020). However, some factors could explain the departure of predictions from actually observed data. Estimates that emerge from modelling studies are only as good as the validity of the epidemiological or statistical models used, the extent and accuracy of the assumptions made and, perhaps most importantly, the quality of the data to which models are calibrated. Early in an epidemic, the quality of data on infections, deaths, tests, and other factors often are limited by under-detection or inconsistent detection of cases, reporting delays, and poor documentation, all of which affect the quality of any model output (Jewell et al., 2020). In the particular case of Covid-19, Achoki et al. (2020) argued that the fact that some people only experience mild symptoms and that even the best health system can only detect and treat those presenting to facilities also means that the available data on 'confirmed' cases represents only a fraction of the true picture of the pandemic. Additionally, the key to establish a reliable model is to track the epidemic dynamics and release the clinical information and epidemiological data in a timely manner. However, official data are often uncertain because medical resources are limited. The available

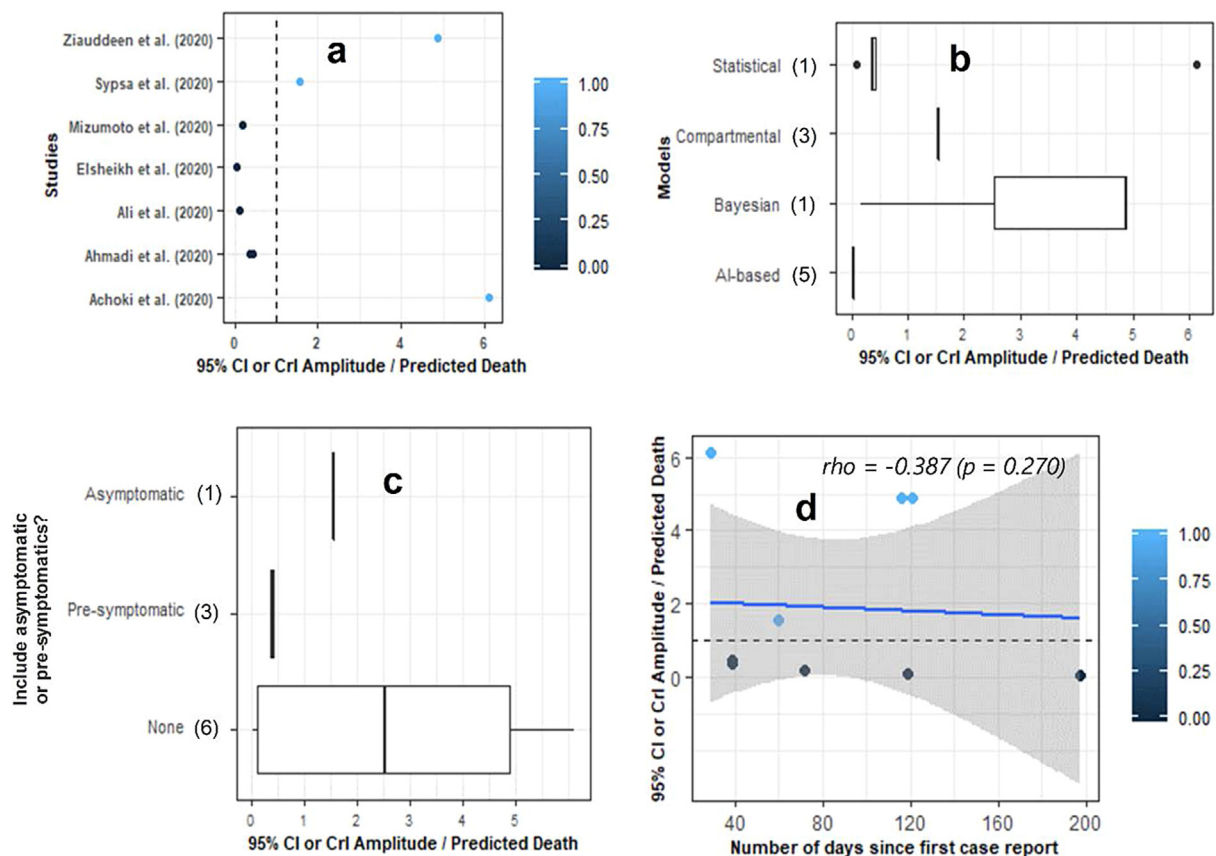


Fig. 9. Precision of the models' predictions: ratio of the amplitude of the 95%CI or 95%CrI of the predicted cumulative number of deaths over the predicted cumulative number of deaths in nine studies (a), across types of models (b), according to whether models included asymptomatic or pre-symptomatics (c), and in relationships to the number of days since the first case was reported in concerned countries (d). Values in parentheses in (b) and (c) represent the number of predictions found for each case.

data only reports confirmed cases in hospitals and ignores infected people who do not have access to medical services. This makes it difficult to accurately predict the course of the epidemic (Bai, Gong, et al., 2020; Roda et al., 2020). Moreover, regarding the problem of data quality, especially in developing countries, though we assume that data are collected in real time and adequately in developed countries, this is undoubtedly not the case in underdeveloped countries that do not have the means for doing so. Achoki et al. (2020), when forecasting cumulative cases, new infections, and mortality due to Covid-19 in Africa found their work further complicated because in Africa, data on key covariates are either lacking or when they exist, they tend to be biased or derived from other global covariate-based modelling exercises. For example, for a growth models at a growth phase of the epidemic, the fact that for 3–4 consecutive days no case was reported (often in week-end and 1–2 days after) is not consistent and will certainly introduce bias in predictions.

Another factor that might contribute to the estimations biases is that models are often built on strong assumptions (Bai, Gong, et al., 2020; Tang, Bragazzi, et al., 2020) that may not hold. Models may capture aspects of epidemics effectively while neglecting to account for other factors, such as the accuracy of diagnostic tests; whether immunity will wane quickly; if reinfection could occur; or population characteristics, such as age distribution, percentage of older adults with co-morbidities, and risk factors (e.g., smoking, exposure to air pollution) (Jewell et al., 2020). Additionally, predictions were made in some studies considering scenarios about control measures that do not always match with the reality on the ground. Then the predictions are prone to biases. In most of the studies, parameters estimated from data collected in the first affected countries such as China were used to derive estimates of parameters in other countries (Zareie, Roshani, Mansournia, Rasouli, & Moradi, 2020) even though it is unlikely that epidemics follow identical paths in all regions of the world (Jewell et al., 2020). An additional point that might also explain the departure of predictions from values actually observed is linked to the fact that predictions are among others intended to guide public health policies for controlling spread of epidemics. As such, based on the predictions, different control measures might have been taken which might have allowed to considerably reduce the number of cases, and hence the observed values, resulting in a “false” overestimation appreciation.

Deterministic models have a long history of application in the study of infectious disease epidemiology. Yet, many infectious disease systems are fundamentally individual-based stochastic processes, and are more naturally described by

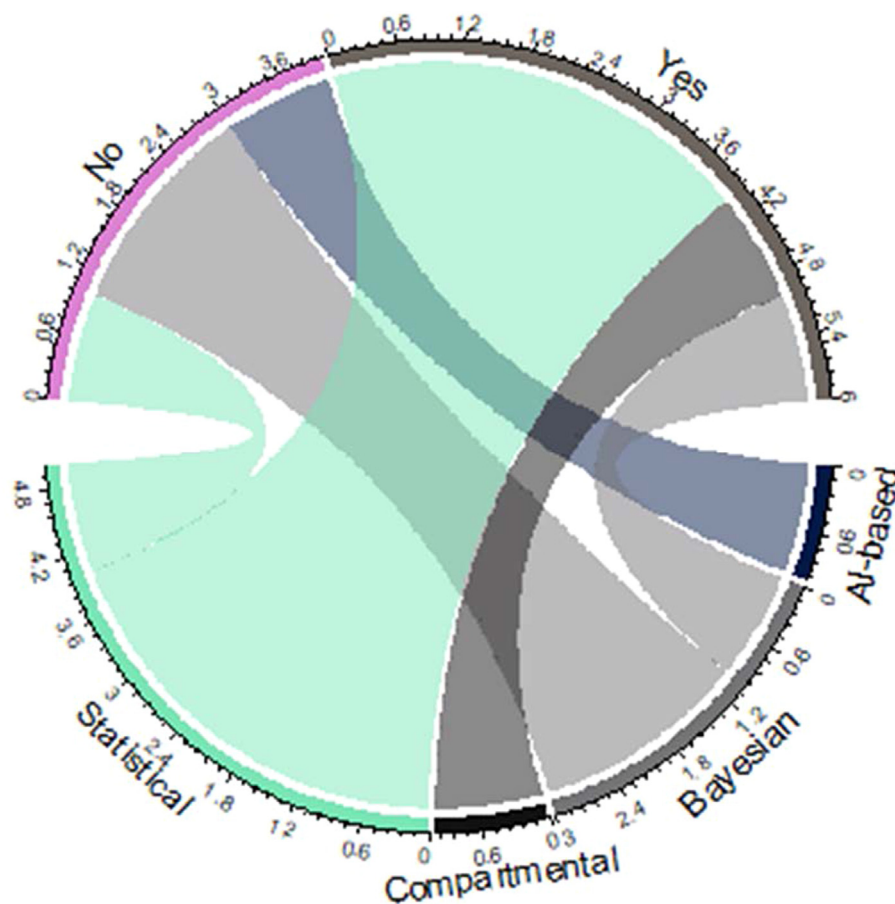


Fig. 10. Distributions of predictions (10) of the cumulative number of deaths according to whether the values actually observed for the predictions fall within the 95%CI or 95%CrI of the prediction.

stochastic models (Roberts et al., 2015). Therefore, using deterministic models for a stochastic process could also be a source of bias in the estimations. Deterministic model typically describes the average behavior of a system (e.g., populations or sub-populations) without taking into account stochastic processes or chance events in single entities (e.g., individuals). Hence, such models are typically applied to situations with a large number of individuals where stochastic variation becomes less important and heterogeneity can be accounted for using various sub-populations (Porgo et al., 2019). Stochastic models are models where the parameters, variables, and/or the change in variables can be described by probability distributions. This type of model can account for process variability by taking into account the random nature of variable interactions, or can accommodate parameter uncertainty, and so may predict a distribution of possible health outcomes. Considering process variability can be particularly important when populations are small or certain events are very rare (Porgo et al., 2019). Stochastic models make it possible to take into account several factors and lead to a more realistic research. Zhang, Zeb, et al. (2020) studied the effects of the environment on the spread of Covid-19 using stochastic mathematical model. Modelling studies have contributed vital insights into the Covid-19 pandemic, and will undoubtedly continue to do so (Jewell et al., 2020) although modelling and predicting the epidemiology and trajectory of a disease such as Covid-19 is a challenging exercise (Achoki et al., 2020). However, mathematicians and statisticians should be rigorous in their methodology in order to provide robust and reliable results based on which appropriate and optimal management strategies to contain the disease efficiently would be made.

5. Conclusion

Using modelling techniques to predict the course of Covid-19 is important to effectively guide public health policy-making. Nevertheless, ensuring that predictions are accurate and precise is necessary to optimize the allocation of the limited resources available, especially in resource-challenged communities. Based on the sample of 242 papers analysed, we showed that compartmental and statistical growth models are so far the most used modelling techniques to predict Covid-19 dynamics. AI-based models, Agent-based models, and Bayesian models, were also used, but to a lower extent. For about 1/3 of

the predictions of the cumulative number of cases, predicted values were larger than observed values and 1/3 of the predictions departed from the values actually observed (ratio of the predicted value to the observed value less than 0.8 or greater than 1.2). We also showed that predictions based on larger dataset (i.e. longer period) were more accurate than predictions based on smaller dataset for the cumulative number of deaths and the Bayesian models seem to provide more accurate and precise predictions. Ensuring identifiability in model calibrations, data quality, and larger period should provide better confidence in predictions. We also remarked that most of studies did not report the CI or CrI of their predictions, thus limiting ability for deeper comparative analyses.

Credit author statement

Janyce Gnanvi: Conceptualization, Methodology, Writing, review and editing; Kolawolé Valère Salako: Conceptualization, Methodology, Writing, review and editing; Gaëtan Brezesky Kotanmi: Formal analysis and Writing; Romain Glèlè Kakaï: Conceptualization, Methodology, Supervision, Validation, Writing, review and editing.

Declaration of competing interest

The author declares no conflict of interest.

Acknowledgments

KVS acknowledges the support of the Wallonie-Bruxelles International Post-doctoral Fellowship for Excellence, Belgium (Fellowship N°SUB/2019/443681). RGK acknowledges the support from the African German Network of Excellence in Science (AGNES) and the Alexander von Humboldt Foundation (AvH). Authors acknowledge the assistance of Sacla Aide Edmond in mobilising additional articles when revising previous version of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.idm.2020.12.008>.

References

- Achoki, T., Alam, U., Were, L., Gebremedhin, T., Senkubuge, F., Lesego, A., Liu, S., Wamai, R., & Kinfu, Y. (2020). *Covid-19 pandemic in the African continent: Forecasts of cumulative cases, new infections, and mortality*. medRxiv. <https://doi.org/10.1101/2020.04.09.20059154>
- Ahmadi, A., Fadaei, Y., Shirani, M., & Rahmani, F. (2020). Modeling and forecasting trend of covid-19 epidemic in Iran until May 13, 2020. *Medical Journal of the Islamic Republic of Iran*, 34(1), 183–195.
- Bai, Z., Gong, Y., Tian, X., Cao, Y., Liu, W., & Li, J. (2020b). The rapid assessment and early warning models for covid-19. *Virologica Sinica*, 35(3), 272–279.
- Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.-Y., Chen, L., & Wang, M. (2020a). Presumed asymptomatic carrier transmission of covid-19. *Jama*, 323(14), 1406–1407.
- Bartolomeo, N., Trerotoli, P., & Serio, G. (2020). Estimating the size of the covid-19 outbreak in Italy: Application of an exponential decay model to the weighted and cumulative average daily growth rate. medRxiv. <https://doi.org/10.1101/2020.05.20.20108241>
- de Camino-Beck, T. (2020). A modified SEIR model with confinement and lockdown of covid-19 for Costa Rica. medRxiv. <https://doi.org/10.1101/2020.05.19.20106492>
- Cao, J., Jiang, X., Zhao, B., et al. (2020). Mathematical modeling and epidemic prediction of covid-19 and its significance to epidemic prevention and control measures. *Journal of Biomedical Research & Innovation*, 1(1), 1–19.
- Ceylan, Z. (2020). Estimation of covid-19 prevalence in Italy, Spain, and France. *The Science of the Total Environment*, 729, 138817.
- Chauhan, P., Kumar, A., & Jamdagni, P. (2020). Regression analysis of covid-19 spread in India and its different states. medRxiv. <https://doi.org/10.1101/2020.05.29.20117069>
- Elmousalimi, H. H., & Hassanien, A. E. (2020). Day level forecasting for coronavirus disease (covid-19) spread: Analysis, modeling and recommendations. *arXiv preprint*. arXiv:2003.07778.
- Hamidouche, M. (2020). *Covid-19 outbreak in Algeria: A mathematical model to predict the incidence*. medRxiv. <https://doi.org/10.1101/2020.03.20.20039891>
- He, J., Guo, Y., Mao, R., & Zhang, J. (2020). Proportion of asymptomatic coronavirus disease 2019: A systematic review and meta-analysis. *Journal of Medical Virology*, 93, 820–830. <https://doi.org/10.1002/jmv.26326>
- Hermanowicz, S. W. (2020). Forecasting the Wuhan coronavirus (2019-ncov) epidemics using a simple (simplistic) model. medRxiv. <https://doi.org/10.1101/2020.02.04.20020461>
- Jewell, N. P., Lewnard, J. A., & Jewell, B. L. (2020). Predictive mathematical models of the covid-19 pandemic: Underlying principles and value of projections. *Jama*, 323(19), 1893–1894.
- Kim, S., Kim, Y.-J., Peck, K. R., & Jung, E. (2020). School opening delay effect on transmission dynamics of coronavirus disease 2019 in Korea: Based on mathematical modeling and simulation study. *Journal of Korean Medical Science*, 35(13), e143.
- Lalmuanawma, S., Hussain, J., & Chhakchhuak, L. (2020). Applications of machine learning and artificial intelligence for covid-19 (sars-cov-2) pandemic: A review. *Chaos, Solitons & Fractals*, 139, 110059.
- Linka, K., Peirlinck, M., Sahli Costabal, F., & Kuhl, E. (2020). Outbreak dynamics of covid-19 in europe and the effect of travel restrictions. *Computer Methods in Biomechanics and Biomedical Engineering*, 23(11), 710–717.
- Liu, Z., Magal, P., Seydi, O., & Webb, G. (2020). A model to predict covid-19 epidemics with applications to South Korea, Italy, and Spain. *SIAM News* (to appear).
- Mizumoto, K., Kagaya, K., Zarebski, A., & Chowell, G. (2020). Estimating the asymptomatic proportion of coronavirus disease 2019 (covid-19) cases on board the diamond princess cruise ship, Yokohama, Japan, 2020. *Euro Surveillance*, 25(10), 2000180.
- Ngonghala, C. N., Iboi, E., Eikenberry, S., Scotch, M., MacIntyre, C. R., Bonds, M. H., & Gumel, A. B. (2020). Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel coronavirus. *Mathematical Biosciences*, 325, 108364.
- Nyabadza, F., Chirove, F., Chukwu, W. C., & Visaya, M. V. (2020). Modelling the potential impact of social distancing on the covid-19 epidemic in South Africa. medRxiv. <https://doi.org/10.1101/2020.04.21.20074492>
- Pasayat, A. K., Pati, S. N., & Maharana, A. (2020). Predicting the covid-19 positive cases in India with concern to lockdown by using mathematical and machine learning based models. medRxiv. <https://doi.org/10.1101/2020.05.16.20104133>

- Porgo, T. V., Norris, S. L., Salanti, G., Johnson, L. F., Simpson, J. A., Low, N., Egger, M., & Althaus, C. L. (2019). The use of mathematical modeling studies for evidence synthesis and guideline development: A glossary. *Research Synthesis Methods*, 10(1), 125–133.
- Postnikov, E. B. (2020). Estimation of covid-19 dynamics “on a back-of-envelope”: Does the simplest sir model provide quantitative parameters and predictions? *Chaos, Solitons & Fractals*, 135, 109841.
- Rhodes, T., Lancaster, K., & Rosengarten, M. (2020). A model society: Maths, models and expertise in viral outbreaks. URL <https://doi.org/10.1080/09581596.2020.1748310>.
- Roberts, M., Andreasen, V., Lloyd, A., & Pellis, L. (2015). Nine challenges for deterministic epidemic models. *Epidemics*, 10, 49–53.
- Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the covid-19 epidemic? *Infectious Disease Modelling*, 5, 271–281.
- Shaikh, A. S., Shaikh, I. N., & Nisar, K. S. (2020). A mathematical model of covid-19 using fractional derivative: Outbreak in India with dynamics of transmission and control. *Advances in Difference Equations*, 2020, 373. <https://doi.org/10.1186/s13662-020-02834-3>
- Shim, E., Tariq, A., Choi, W., Lee, Y., & Chowell, G. (2020). Transmission potential and severity of covid-19 in South Korea. *International Journal of Infectious Diseases*, 93. <https://doi.org/10.1016/j.ijid.2020.03.031>
- Shmueli, G., et al. (2010). To explain or to predict? *Statistical Science*, 25(3), 289–310.
- Silva, T. C., Anghinoni, L., & Zhao, L. (2020). Quantitative analysis of the effectiveness of public health measures on covid-19 transmission. medRxiv. <https://doi.org/10.1101/2020.05.15.20102988>
- Stallings, W. M., & Gillmore, G. M. (1971). A note on “accuracy” and “precision”. *Journal of Educational Measurement*, 8(2), 127–129.
- Stübinger, J., & Schneider, L. (2020). Epidemiology of coronavirus covid-19: Forecasting the future incidence in different countries. *Healthcare*, 8, 99.
- Taboe, B. H., Salako, V. K., Tison, J., Ngonghala, C. N., & Glèlè Kakai, R. (2020). Predicting covid-19 spread in the face of control measures in West-Africa. *Mathematical Biosciences*, 328, 108431.
- Tang, B., Bragazzi, N. L., Li, Q., Tang, S., Xiao, Y., & Wu, J. (2020a). An updated estimation of the risk of transmission of the novel coronavirus (2019-nCoV). *Infectious disease modelling*, 5, 248–255.
- Tang, B., Wang, X., Li, Q., Bragazzi, N. L., Tang, S., Xiao, Y., & Wu, J. (2020b). Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *Journal of Clinical Medicine*, 9(2), 462.
- Tian, T., Jiang, Y., Zhang, Y., Li, Z., Wang, X., & Zhang, H. (2020). Covid-net: A deep learning based and interpretable predication model for the county-wise trajectories of covid-19 in the United States. medRxiv. <https://doi.org/10.1101/2020.05.26.20113787>
- Toda, A. A. (2020). Susceptible-infected-recovered (sir) dynamics of covid-19 and economic impact. arXiv preprint. arXiv:2003.11221.
- Velavan, T. P., & Meyer, C. G. (2020). The covid-19 epidemic. *Tropical Medicine & International Health*, 25(3), 278.
- Wang, C., Liu, L., Hao, X., Guo, H., Wang, Q., Huang, J., He, N., Yu, H., Lin, X., Pan, A., et al. (2020). Evolving epidemiology and impact of non-pharmaceutical interventions on the outbreak of coronavirus disease 2019 in Wuhan, China. MedRxiv. <https://doi.org/10.1101/2020.03.03.20030593>
- WHO. (2020). Pneumonia of unknown cause—China, emergencies preparedness, response. Disease outbreak news, World Health Organization (WHO).
- Willem, L., Verelst, F., Bilcke, J., Hens, N., & Beutels, P. (2017). Lessons from a decade of individual-based models for infectious disease transmission: A systematic review (2006–2015). *BMC Infectious Diseases*, 17(1), 612.
- Worldometer. (2020). Coronavirus update. Worldometer.
- Wu, K., Darcet, D., Wang, Q., & Sornette, D. (2020). Generalized logistic growth modeling of the Covid-19 outbreak in 29 provinces in China and in the rest of the world. arXiv preprint, 101, 1561–1581. <https://doi.org/10.1101/2020.03.11.20034363>. arXiv:2003.05681.
- Xiao, Y., Tang, B., Wu, J., Cheke, R. A., & Tang, S. (2020). Linking key intervention timing to rapid decline of the COVID-19 effective reproductive number to quantify lessons from mainland China. *International Journal of Infectious Diseases*, 97, 296–298.
- Yang, S., Cao, P., Du, P., Wu, Z., Zhuang, Z., Yang, L., Yu, X., Zhou, Q., Feng, X., Wang, X., et al. (2020). Early estimation of the case fatality rate of covid-19 in mainland China: A data-driven analysis. *Annals of Translational Medicine*, 8(4).
- Yin, Y., & Wunderink, R. G. (2018). MERS, SARS and other coronaviruses as causes of pneumonia. *Respirology*, 23(2), 130–137.
- Zareie, B., Roshani, A., Mansournia, M. A., Rasouli, M. A., & Moradi, G. (2020). A model for covid-19 prediction in Iran based on China parameters. medRxiv, 23(4), 244–248. <https://doi.org/10.1101/2020.03.19.20038950>
- Zhang, J., Litvinova, M., Wang, W., Wang, Y., Deng, X., Chen, X., Li, M., Zheng, W., Yi, L., Chen, X., et al. (2020a). Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside hubei province, China: A descriptive and modelling study. *The Lancet Infectious Diseases*, 20(7), 793–802.
- Zhang, Z., Zeb, A., Hussain, S., & Alzahrani, E. (2020b). Dynamics of covid-19 mathematical model with stochastic perturbation. *Advances in Difference Equations*, 2020(1), 1–12.
- Zhao, S., Gao, D., Zhuang, Z., Chong, M. K., Cai, Y., Ran, J., . . . Wang, W., et al. (2020a). Estimating the serial interval of the novel coronavirus disease (Covid-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020. *Frontiers in Physics*, 8, 347.
- Zhao, Z., Li, X., Liu, F., Zhu, G., Ma, C., & Wang, L. (2020b). Prediction of the covid-19 spread in African countries and implications for prevention and controls: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. *The Science of the Total Environment*, 729, 138959.
- Ziauddeen, H., Subramaniam, N., & Gurdasani, D. (2020). Modelling the impact of lockdown easing measures on cumulative covid-19 cases and deaths in England. medRxiv. <https://doi.org/10.1101/2020.06.21.20136853>