



# Incorporation of near-real-time hospital occupancy data to improve hospitalization forecast accuracy during the COVID-19 pandemic



Alexander Preiss<sup>a,1,\*</sup>, Emily Hadley<sup>a,1</sup>, Kasey Jones<sup>a</sup>, Marie C.D. Stoner<sup>a</sup>, Caroline Kery<sup>a</sup>, Peter Baumgartner<sup>b</sup>, Georgiy Bobashev<sup>a</sup>, Jessica Tenenbaum<sup>c</sup>, Charles Carter<sup>c</sup>, Kimberly Clement<sup>c</sup>, Sarah Rhea<sup>d</sup>

<sup>a</sup> RTI International, Research Triangle Park, NC, USA

<sup>b</sup> Explosion, Berlin, Germany

<sup>c</sup> North Carolina Department of Health and Human Services, Raleigh, NC, USA

<sup>d</sup> North Carolina State University, Department of Population Health and Pathobiology, Raleigh, NC, USA

## ARTICLE INFO

### Article history:

Received 4 November 2021

Received in revised form 28 January 2022

Accepted 29 January 2022

Available online 4 February 2022

Handling editor: Dr HE DAIHAI HE

## ABSTRACT

Public health decision makers rely on hospitalization forecasts to inform COVID-19 pandemic planning and resource allocation. Hospitalization forecasts are most relevant when they are accurate, made available quickly, and updated frequently. We rapidly adapted an agent-based model (ABM) to provide weekly 30-day hospitalization forecasts (i.e., demand for intensive care unit [ICU] beds and non-ICU beds) by state and region in North Carolina for public health decision makers. The ABM was based on a synthetic population of North Carolina residents and included movement of agents (i.e., patients) among North Carolina hospitals, nursing homes, and the community. We assigned SARS-CoV-2 infection to agents using county-level compartmental models and determined agents' COVID-19 severity and probability of hospitalization using synthetic population characteristics (e.g., age, comorbidities). We generated weekly 30-day hospitalization forecasts during May–December 2020 and evaluated the impact of major model updates on statewide forecast accuracy under a SARS-CoV-2 effective reproduction number range of 1.0–1.2. Of the 21 forecasts included in the assessment, the average mean absolute percentage error (MAPE) was 7.8% for non-ICU beds and 23.6% for ICU beds. Among the major model updates, integration of near-real-time hospital occupancy data into the model had the largest impact on improving forecast accuracy, reducing the average MAPE for non-ICU beds from 6.6% to 3.9% and for ICU beds from 33.4% to 6.5%. Our results suggest that future pandemic hospitalization forecasting efforts should prioritize early inclusion of hospital occupancy data to maximize accuracy.

© 2022 The Authors. Publishing services 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. RTI International.

E-mail address: [apreiss@rti.org](mailto:apreiss@rti.org) (A. Preiss).

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

<sup>1</sup> These authors contributed equally to this work.

## 1. Introduction

Coronavirus disease 2019 (COVID-19) hospitalizations are a crucial metric used to monitor and respond to the COVID-19 pandemic. COVID-19, the disease caused by SARS-CoV-2, has been associated with over 2.2 million hospitalizations and over 600,000 deaths in the United States as of June 2021 (Centers for Disease Control and Prevention, 2021b). Throughout the COVID-19 pandemic, many U.S. health systems have struggled to meet the demand for patient care with limited resources of hospital beds, ventilators, and qualified health care workers (Maxmen, 2020; Schenck et al., 2020). When these resources become stretched, hospitals may be unable to deliver high-quality care to patients with and without COVID-19 (Daugherty Biddison et al., 2019; Rosenbaum, 2020; Ventura et al., 2020). Projections of hospital bed demand, including when capacity may be exceeded, can aid public health and health care stakeholders in planning for and responding to these scenarios (Weissman et al., 2020). These forecasts can model capacity under a range of possible scenarios by varying the effective reproduction number, or  $R_e$  (Inglesby, 2020). Higher  $R_e$  values correspond to faster spread of SARS-CoV-2, while lower  $R_e$  values correspond to slower spread (Petersen et al., 2020).

Researchers have developed a variety of models to forecast COVID-19 hospitalizations, including simple formulas (Cavallo et al., 2020; Holmes et al., 2020), statistical models (Castro et al., 2020; Srivastava et al., 2020), deep learning models (Rodríguez et al., 2020), compartmental models (Arik et al., 2020; Massachusetts General Hospital Institute for Technology Assessment, 2020; Weissman et al., 2020), and agent-based models (ABMs) (No author, 2020). Others have used ABMs to answer what-if questions related to COVID-19 interventions (Holmdahl et al., 2021; Naimark et al., 2021). Numerous publicly available, routinely updated national and state models have accurately predicted COVID-19 cases and deaths (University of Southern California, 2021). However, publicly available models offer limited flexibility to end users who want to apply their own substantive, on-the-ground knowledge and unique data to obtain more customized model output. Most publicly available forecasts do not consider changes in non-COVID-19 hospitalizations, locations of individual patients, or capacities of individual hospitals (Centers for Disease Control and Prevention, 2021a; IMHE, 2021). Incorporation of population, geospatial, and health system dynamics in forecasting models can enhance hospitalization forecasts.

To address this need, we adapted an existing ABM of the North Carolina health system (Rhea et al., 2019) by integrating a COVID-19 module. The ABM was originally developed to model healthcare-associated infections. It modeled patient movement among healthcare facilities and was calibrated to contemporary hospital discharge data (Jones et al., 2019). The existing ABM provided a framework from which to realistically model movement of patients with COVID-19 into hospitals for forecasting purposes. This framework allowed us to consider disease dynamics on a local, individual scale (i.e., hospital), as well as on a large, aggregate scale (i.e., state/region). In addition to model accuracy, this modeling effort had three overarching goals. First, forecasts were needed as soon as possible for rapid decision support. However, generating forecasts early in the pandemic was a challenge as ABMs are complex and require substantial development time. Second, forecasts needed to be generated as close to real time as possible, minimizing the time between receiving data and returning forecasts. Other modeling teams have reported the value of quick forecast turnaround by responding to the demand for SARS-CoV-2 forecasts in as little as 2 days (Weissman et al., 2020). Third, because of the dynamic nature of the pandemic, forecasts needed to be updated frequently.

We used our updated ABM to generate 30-day North Carolina hospitalization forecasts at weekly intervals during May–December 2020 to address these goals. By leveraging an existing ABM of the North Carolina health system, we focused development efforts on the addition of a COVID-19 module. This approach allowed us to provide beta-version forecasts within 2 days of the North Carolina state of emergency declaration on March 10, 2020, and validated weekly forecasts within 2 months of that date. Here, we describe our experiences developing and implementing the ABM with the COVID-19 module throughout the early pandemic period. We provide an evaluation of our model's accuracy and evaluate the impact of major model updates on forecast accuracy.

## 2. Methods

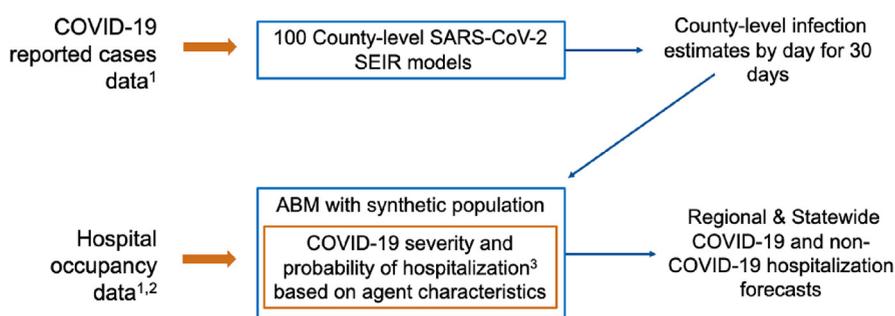
**ABM with COVID-19 Module.** The existing geospatially explicit ABM included the movement of agents (i.e., patients) among 104 acute care hospitals (“hospitals” throughout), 10 long-term acute care hospitals (LTACHs), 421 licensed nursing homes, and the community in North Carolina (Rhea et al., 2019). Agents were based on a synthetic population of North Carolina residents (i.e., >10.4 million agents), each with unique characteristics, including age, county of residence, and comorbidities (Jones et al., 2019). Within the ABM, hospital characteristics included location and number of licensed and staffed intensive care unit (ICU) beds and non-ICU (i.e., regular inpatient) beds. Statewide, the ABM included a total of approximately 17,900 non-ICU beds and approximately 3200 ICU beds. The exact bed counts varied as our methodology and data evolved over the course of forecasting. Relevant COVID-19 parameters are in Table 1, with additional parameters and model details available in the Overview, Design Concepts, and Details (ODD) protocol and eMethods (Jones et al., 2021).

We created a COVID-19 module consisting of county-level susceptible, exposed, infectious, recovered (SEIR) compartmental models for each of North Carolina's 100 counties. These SEIR models were used to produce 30-day SARS-CoV-2 infection projections (Fig. 1) under various  $R_e$  ranges (Table 1). For each forecast, we based the SEIR models on the latest available North Carolina county-level reported COVID-19 cases. We used official case data provided by the North Carolina Department of Health and Human Services (NCDHSS), in contrast to many modeling efforts which rely on publicly available

**Table 1**  
Relevant COVID-19 parameters.<sup>a</sup>

Parameter	Value	Source
R <sub>0</sub> minimum and maximum value bound	[1–1.2], [1.2–1.4], [1.4–1.6]	(COVID Act Now, 2021; Ferguson et al., 2020 March 16; Gu, 2020; Imperial College London, 2020; Institute of Global Health, 2021; Li et al., 2020; Rees et al., 2020; Rosenberg et al., 2020; Rt.live, 2020; Sun & Achenbach, 2020 June 25; Wake Forest Baptist Health, 2021; Zeek, 2020 May 13)
COVID-19 agent length of stay (days)	[median = 5], [mean = 3], [standard deviation = 5], [minimum = 0], [maximum = 50]; truncated normal distribution	Rees et al. (2020)
Proportion of population that remains susceptible when the simulation starts	0.9	Wake Forest Baptist Health (2021)
Infectious period (days)	6	(Ferguson et al., 2020 March 16)
Incubation period (days)	5	(Ferguson et al., 2020 March 16)
Length of infection (days) used for SEIR model and for calculating recovery days among COVID-19 agents not admitted to a hospital	14	
Initial case multiplier representing ratio of unreported infections to reported cases prior to the start of the model	10	(Li et al., 2020; Rosenberg et al., 2020; Sun & Achenbach, 2020 June 25; Zeek, 2020 May 13)
Proportion of hospitalized COVID-19 agents requiring an ICU bed	0.25	(Ferguson et al., 2020 March 16), Expert opinion
Proportion of infected agents that are tested	0.1	Inverse of initial case multiplier
Proportion of agents with asymptomatic, mild, or moderate symptoms that seek hospitalization	0	Expert opinion
Distribution of positive COVID-19 reported cases by age	[Age 0–49: 0.396], [50–64: 0.328], [65+: 0.276]	Bayesian calculation <sup>1</sup>
Probability of hospitalization by age given a positive, tested SARS-CoV-2 infection with comorbidities	[Age 0–49: 0.0], [Age 50–64: 0.4609], [Age 65+: 0.411]	Bayesian calculation <sup>1</sup>
Probability of hospitalization by age given a positive, tested SARS-CoV-2 infection without comorbidities	[Age 0–49: 0.0367], [Age 50–64: 0.035], [Age 65+: 0.1213]	Bayesian calculation <sup>1</sup>
Probability of hospitalization by age given a positive, untested SARS-CoV-2 infection with comorbidities	[Age 0–49: 0.0], [Age 50–64: 0.0651], [Age 65+: 0.058]	Bayesian calculation <sup>1</sup>
Probability of hospitalization by age given a positive, untested SARS-CoV-2 infection without comorbidities	[Age 0–49: 0.0052], [Age 50–64: 0.0049], [Age 65+: 0.0171]	Bayesian calculation <sup>1</sup>

<sup>a</sup> For a comprehensive list of parameters and details on Bayesian calculations (Jones et al., 2021).



<sup>1</sup> Typically from the day before each model run  
<sup>2</sup> Including intensive care unit (ICU) beds and non-ICU beds  
<sup>3</sup> For COVID-19 agents and for non-COVID-19 agents

**Fig. 1.** Integration of susceptible-exposed-infectious-recovered (SEIR) models with an agent-based model (ABM) for hospitalization forecasts by state and region in North Carolina.

data (Patel et al., 2021). To account for under-reporting, reported case counts were multiplied by a fixed multiplier of 10 (Table 1).

SARS-CoV-2 infection status was assigned to agents in the ABM's community node using the infection projections (Fig. 1). Subsequently, COVID-19 severity and probability of hospitalization for agents in the ABM with SARS-CoV-2 were determined

by agent-specific characteristics (e.g., age, comorbidities) (Jones et al., 2021). Other agent hospitalizations in the ABM were according to previous calibrations (Jones et al., 2021; Rhea et al., 2020). The code base for the ABM is open source and available on GitHub (RTI International, 2021).

**Model Output.** Every model run generated estimates of multiple variables for each day of the 30-day forecast, including new and cumulative SARS-CoV-2 infections, new agents seeking a hospital bed, census of hospitalized agents, and hospital demand (the number of agents in need of a hospital bed). Hospitalization variables were disaggregated by COVID-19 status and ICU status (Jones et al., 2021). Hospital demand could be larger than the number of hospitalized agents when the model predicted that hospital bed capacity was reached.

For each weekly forecast, we conducted 100 simulations per  $R_e$  range. We summarized output variables from the 100 simulations by mean and by 25th and 75th percentiles. Forecasts were produced at various geographic levels. We performed a series of validation checks on model output for each run (**eMethods**).

**Major Model Updates.** During May–December 2020, we produced forecasts approximately once weekly. Over this period, we made three major model updates, as described below.

**Adjustment of county-level  $R_e$  values for SEIR models.** Prior to August 2020, we adjusted county  $R_e$  values based on each county's case counts (**eMethods**). Although the assumption of homogeneous mixing might not be applicable to small and sparse populations, it has been shown that  $R_e$  adjustments can lead to a reasonable approximation of a detailed model with complex contact matrices (Rahmandad & Sterman, 2008). However, this approach led to some smaller counties seeing unrealistic  $R_e$  values. To avoid spurious estimates caused by an occasional local outbreak, beginning in August 2020, we adjusted each county-level  $R_e$  using the following equation designed to dampen extreme  $R_e$  values:

$$\text{County corrected } R_e = \sqrt[3]{\frac{\text{County } R_e}{\text{State } R_e}} * \text{Model } R_e \quad (1)$$

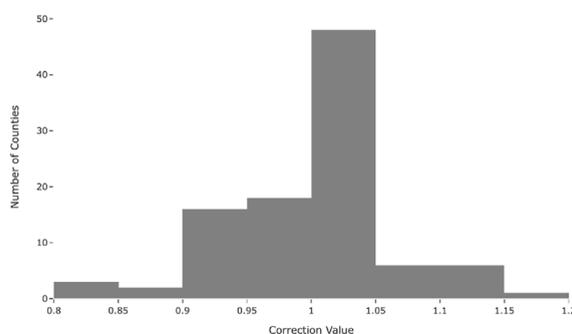
The correction value generally ranged between 0.8 and 1.2 (Fig. 2). Without this correction, some calculated county-level  $R_e$  values were too large or too small to be realistic.

**Better initialization of non-COVID-19 hospitalizations.** Prior to September 2020, we applied an estimated mean hospital occupancy of 56% across all modeled hospitals. Beginning in September 2020, we used the latest available, near-real-time North Carolina hospital occupancy data to initialize hospitalized agents for each hospital in the ABM.

**Implementation of updated location transition probabilities.** Location transition probabilities in the original ABM were determined using pre-pandemic hospital discharge data (Rhea et al., 2020). However, during the pandemic, non-COVID-19 related hospitalizations decreased (Stradling, 2021). We used the near-real-time hospital occupancy data for North Carolina to recalculate transition probabilities each week. These updated transition probabilities ensured that a steady state of non-COVID-19 agent hospitalizations occurred during a 30-day model run (i.e., non-COVID-19 agent hospital occupancy remained at a consistent level). These updates were fully integrated by October 2020.

**Forecast Accuracy.** We assessed accuracy by retrospectively comparing our state-level  $R_e = 1.0$ –1.2 forecasts to North Carolina hospitalization data, as reported daily by 104 North Carolina hospitals to the NCDHHS. In this assessment, we included all state-level  $R_e = 1.0$ –1.2 forecasts produced during May–December 2020 in which at least 80% (i.e., 24 days) of the forecast values could be compared with NCDHHS hospitalization data (**eMethods**). The  $R_e = 1.0$ –1.2 range was chosen for this assessment because statewide  $R_e$  estimates during the forecast production period fell within this range (Rt.live, 2020).

We compared the means and interquartile ranges (IQRs) (across the 100 simulations) of forecasted demand for ICU beds and non-ICU beds to reported hospitalizations by date. We also compared the means and IQRs of forecasted change in bed demand by COVID-19 agents to the change in reported hospitalizations by COVID-19 patients by date. This change in demand



**Fig. 2.** Example of  $R_e$  correction multipliers used during the project. These values were multiplied by the modeled  $R_e$  value of the corresponding county to generate county corrected  $R_e$  values.

for a model run was measured by the difference in beds needed between Day 1 and Day 30 of the forecast. We quantified forecast accuracy using mean absolute error (MAE) and mean absolute percentage error (MAPE). MAE measures the absolute difference between the predicted value and the observed value. MAPE measures the percentage difference between the predicted value and the observed value. Smaller MAE and MAPE values indicate higher forecast accuracy.

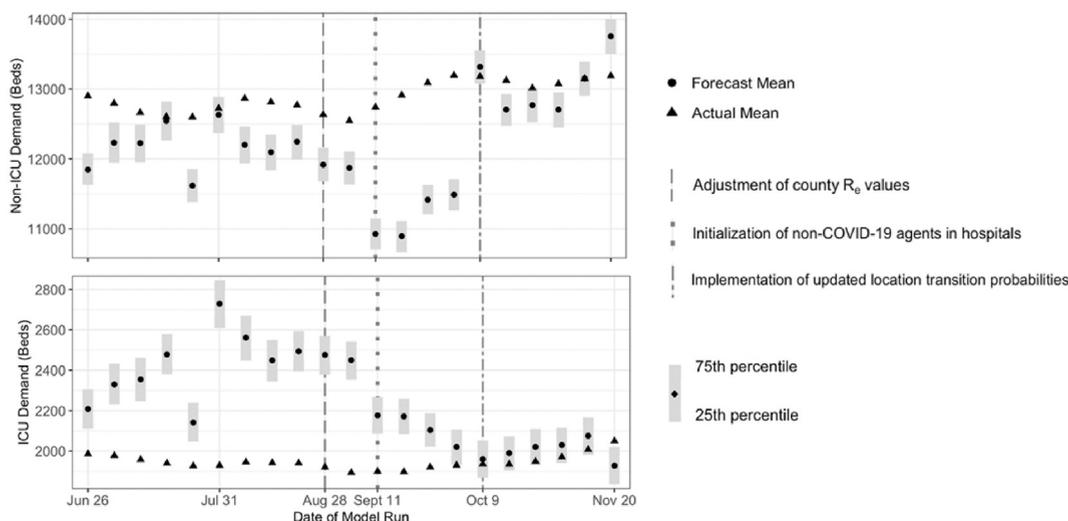
### 3. Results

During May–December 2020, we produced 31 weekly forecasts. Of these, 21 state-level  $R_e = 1.0$ – $1.2$  weekly forecasts, from June 26 to November 20, 2020, had comparison data for at least 80% of the forecast values and were included in the accuracy assessment. Across all included forecasts, the average MAPE was 7.8% for non-ICU beds and 23.6% for ICU beds. Following full implementation (by October 2020) of near-real-time hospital occupancy data in the model (i.e., initialization of non-COVID-19 agents in hospitals and updates to location transition probabilities), the average MAPE was reduced to 3.9% for non-ICU beds and 6.5% for ICU beds.

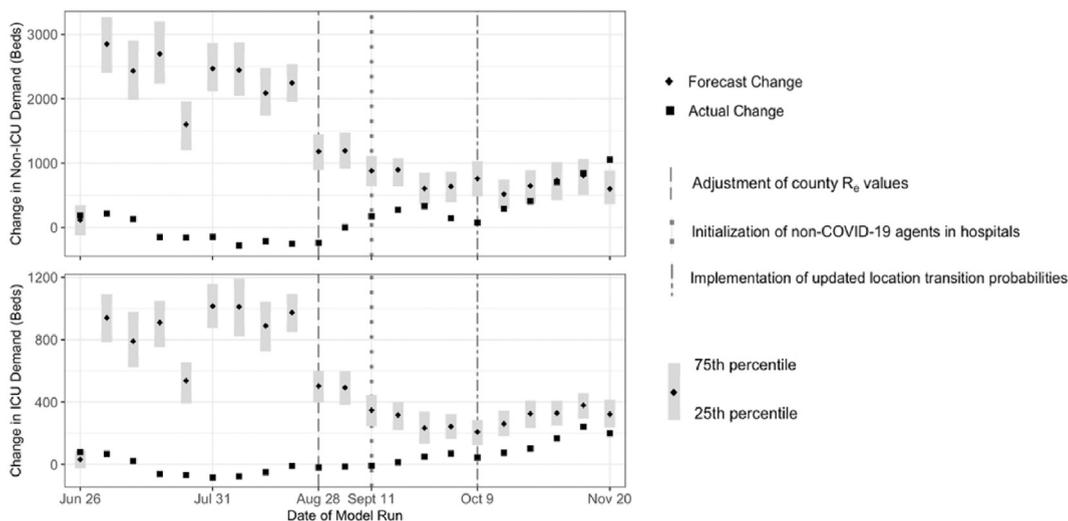
The mean non-ICU bed demand forecasts were generally lower than reported non-ICU hospitalizations, particularly during August–September 2020 (Fig. 3). In 4 of the 21 forecasts, the IQR of non-ICU bed demand included the reported non-ICU hospitalizations. The mean ICU bed demand forecasts were generally higher than reported ICU hospitalizations until early October 2020, when the forecasts became better aligned with reported values. In 5 of the 21 forecasts, the IQR of ICU bed demand included the reported ICU hospitalizations. Among the three major model updates, initialization of non-COVID-19 agents in hospitals (September 11, 2020) and implementation of updated location transition probabilities (October 9, 2020) had the largest impact on improving bed demand forecast accuracy.

For the forecasts produced during June–September 2020, the forecasted changes in non-ICU bed demand and ICU bed demand by COVID-19 agents were higher than the changes in reported hospitalizations for patients with COVID-19 (Fig. 4). The July–August 2020 timeframe is particularly notable because the forecasted change in demand increased while the reported change in demand decreased. Among the three major model updates, adjusting the county-level  $R_e$  correction values to avoid unrealistically extreme  $R_e$  values (August 28, 2020) had the largest impact on improving the accuracy of forecasted changes. This update resulted in fewer projected SARS-Cov-2 infections and, subsequently, a smaller change in demand for hospital beds by COVID-19 agents for the duration of the forecast period.

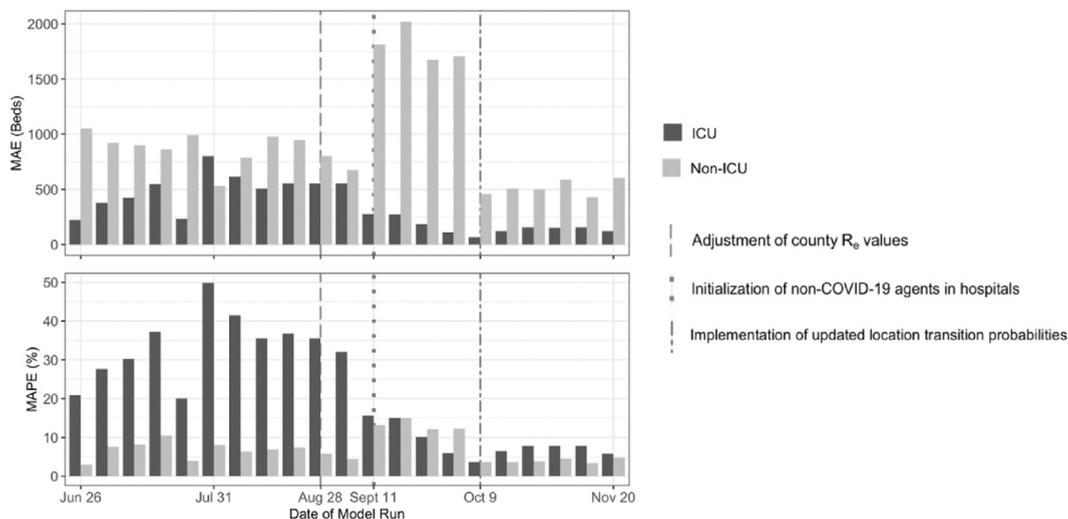
MAE was higher for non-ICU bed demand forecasts than for ICU bed demand forecasts, reflecting the larger number of non-ICU beds (Fig. 5). MAPE was generally higher for ICU bed demand forecasts. During the forecasting period, the lowest MAPE and MAE for non-ICU and ICU bed demand forecasts occurred after the implementation of updated location transition probabilities in October 2020. After this improvement, the average MAPE value for non-ICU bed demand forecasts was less than two-thirds of the average initial MAPE value. The average MAPE value for ICU bed demand forecasts was nearly one-fifth of the initial average (Table 2).



**Fig. 3.** Mean forecasted demand<sup>a</sup> for non-intensive care unit (ICU) beds and ICU beds by date of model run and retrospectively compared to reported demand,<sup>b</sup> North Carolina, June 26, 2020–November 20, 2020



**Fig. 4.** Change in forecasted demand<sup>a</sup> by COVID-19 agents for non-intensive care unit (ICU) beds and ICU beds by date of model run and retrospectively compared to reported demand,<sup>b</sup> North Carolina, June 26, 2020–November 20, 2020



**Fig. 5.** Mean absolute error (MAE) and mean absolute percentage error (MAPE) for non-intensive care unit (ICU) beds and ICU beds under an effective reproduction number range of 1.0–1.2 and by date of model run, North Carolina, June 26, 2020–November 20, 2020.<sup>a</sup>

**Table 2**

Average mean absolute error (MAE) and mean absolute percentage error (MAPE) by hospital bed type (non-intensive care unit [ICU] bed and ICU bed) before, during, and after integration of the near-real-time<sup>a</sup> hospital occupancy data.<sup>b</sup>

Hospital Occupancy Data	Time Period	Bed Type	Average	
			MAE	MAPE
Not available	June 26, 2020–September 4, 2020	Non-ICU	859	6.6%
		ICU	490	33.4%
Used to initialize non-COVID-19 agents in hospitals in the ABM	September 11, 2020–October 2, 2020	Non-ICU	1804	13.1%
		ICU	212	11.7%
Used to determine updated location transition probabilities in the ABM	October 9, 2020–November 20, 2020	Non-ICU	514	3.9%
		ICU	128	6.5%

<sup>a</sup> Typically from the day before the model run.

<sup>b</sup> Hospital occupancy data as reported by hospitals to the North Carolina Department of Health and Human Services and provided for model input.

In May 2020, a full deployment of the ABM and COVID-19 module took 20 h on a local server with 40 cores. After implementing model efficiencies (i.e., cloud computing, GNU parallel processing (Tange, 2020), code refactoring), we reduced model runtime from 20 h to approximately 15 min and typically turned around forecasts within 1 business day.

#### 4. Discussion

Public health and health care stakeholders can use hospital bed demand forecasts for COVID-19 pandemic planning and response. We adapted an existing ABM to address this need by adding a COVID-19 module of county-level SEIR models. We generated the forecasts using near-real-time data and updated the forecasts weekly during May–December 2020. During this time, we made numerous model updates, including three major updates: (a) adjustment of county-level  $R_e$  values for SEIR models; (b) better initialization of non-COVID-19 hospitalizations; and (c) implementation of updated location transition probabilities. The addition of near-real-time hospitalization data, enabling updates b and c, had the largest impact on our model's forecast accuracy.

After the hospitalization data were fully integrated, MAPE and MAE values decreased and both non-ICU and ICU demand forecasts were more accurate. After incorporating the hospitalization data, the average MAPE was 3.9% for non-ICU beds and 6.5% for ICU beds. These relatively small (though not negligible) MAPE values provide useful context for stakeholders when interpreting the potential accuracy of future forecasts. This was the most substantial change for the project and highlights that obtaining relevant data through close collaboration with stakeholders can improve model output considerably.

Forecasts shared with public health decision makers included information about the assumptions, parameters, and limitations of the model. Making this information transparent supported open dialog about parameter updates. In addition, the transparency and open dialog helped support model development and enabled adding hospital occupancy data when they became available. This approach improved our forecasts and helped us collect end-user feedback, which supported trust in the model and the output.

Employing a variety of technical best practices, including cloud computing and automation, greatly improved the model's speed. This allowed us to provide results regularly and rapidly to stakeholders, while still permitting integration of new features and data sources that were often requested by stakeholders. This development framework was crucial in a public health crisis, where stakeholders required immediate information that could adjust to changing circumstances.

Public health stakeholders in North Carolina used the model's results in a variety of ways. First, stakeholders used results to reinforce trends seen from other data sources. For example, when hospitalization rates were increasing but hospitals were not reporting concern that they were likely to hit capacity in the next several weeks, it was reassuring to see that the models showed the same thing. Second, stakeholders used regional forecasts to identify which regions were most likely to face hospital capacity challenges and prepare accordingly. This helped inform discussions regarding which regions could take transfers from other regions if necessary. Finally, stakeholders used state and regional forecasts to predict upcoming need for personal protective equipment (PPE) and inform PPE purchasing and allocation.

In a period of extreme uncertainty, with numerous tactical and strategic decisions to be made each week in the face of uncertainty and ever-changing conditions, this model complemented other sources in enabling data- and science-driven policy. If a similar situation is encountered in the future, researchers should prioritize acquiring and integrating relevant and up-to-date hospitalization data into the model, deploying cloud computing resources, automating time-consuming processes, and ensuring transparent communication with stakeholders.

#### 5. Limitations

Our approach to modeling SARS-CoV-2 transmission and COVID-19 illness was subject to several limitations. Foremost, agents were infected with SARS-CoV-2 using estimates from the SEIR models rather than by explicitly modeling SARS-CoV-2 transmission within the ABM. Additionally, SARS-CoV-2 infections were only assigned to agents in the ABM's community node who had not previously been infected. Finally, once infected with SARS-CoV-2, agents did not move among COVID-19 severity states (e.g., mild, moderate, severe). Death associated with COVID-19 was not explicitly modeled because hospitalization was the public health priority of this modeling effort. Furthermore, since SARS-CoV-2 transmission was modeled with SEIR models, mortality in the ABM would not affect transmission. Therefore, we did not expect COVID-19 mortality to have a practically significant effect on hospitalization outcomes.

Missing data also affected our model. Actual available ICU bed and non-ICU bed counts in North Carolina varied over the course of forecasting. Although the ABM was updated with the latest values as they became available, outdated bed counts could have affected transition probabilities and hospital demand estimate accuracy. Although we did not assess changes to individual model parameters for each weekly forecast, we performed weekly comparisons of model output to reported COVID-19 cases from the previous week (**eMethods**).

Bias in underlying data could also affect model results. In particular, the SEIR models relied on reported COVID-19 cases to produce projected COVID-19 cases. Although we accounted for under-reporting with a 10x multiplier, the true rate of under-reporting was unknown and likely varied over time. We ran sensitivity analysis on this multiplier at two points during this project to determine how sensitive projected cases were to under-reporting. Larger under-reporting multipliers decreased the number of "susceptible" individuals in the SEIR models over time, and thus led to slightly smaller case estimates. Both analyses showed that the multiplier had a small impact on our forecasts for COVID-19 reported case counts, especially at the

beginning of the pandemic. Bias in underlying data compounded over time, because knowledge of the number of cumulative cases, and thus susceptible individuals, was critical to estimate future cases. We accounted for this uncertainty by varying  $R_e$  ranges to model a range of scenarios, including worst-case scenarios. By running a variety of  $R_e$  values we also captured the uncertainty in the under-reporting multiplier.

Future model updates may include the ability to directly implement policy changes in the model and to include indicators for warning signs at certain thresholds (e.g., 85% capacity reached). Ensemble forecasting may be beneficial for reducing the variability and uncertainty inherent in a single-model approach (Ray et al., 2020).

## 6. Conclusion

Forecasts of hospital demand during the COVID-19 pandemic can be powerful tools for intervention planning and resource allocation. We adapted an ABM of the North Carolina health care system to create state and regional hospital demand estimates for public health stakeholders. By varying  $R_e$  input parameters, we provided multiple hospital capacity estimates that corresponded to theoretical changes in COVID-19 spread. The technical best practices we used facilitated rapid development and deployment.

We draw two conclusions from this modeling effort. First, the accuracy of our forecasts improved over time, particularly with the inclusion of near-real-time hospitalization data. As researchers have shown in myriad contexts, improving the underlying data is often the most effective way to improve a model. Rapid response modeling projects should make every effort to secure crucial data as early as possible. Second, this project demonstrates how an ABM can be built rapidly and continually improved to support public health decision makers in the COVID-19 pandemic. An iterative modeling process and collaborative relationship between modelers and public health decision makers was crucial to continually improve this model. This close relationship also enabled modelers to communicate thorough and realistic information about the limitations and assumptions of the forecasts to public health decision makers.

## Funding sources

North Carolina Department of Health and Human Services Contract #2020-DHHS-COVPS 010; Centers for Disease Control and Prevention Contract #75D30120P07909.i

## Disclaimers

The findings and conclusions in this publication are those of the authors and do not necessarily represent the views of the North Carolina Department of Health and Human Services, Division of Public Health, or the Centers for Disease Control and Prevention.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We are grateful for the support and input of Susan Eversole, Stacy Endres-Dighe, and Alex Giarrocco of RTI International; and of Brad Wheeler and Katy Donovan of the North Carolina Department of Health and Human Services. This activity was based on a model originally developed through support from Centers for Disease Control and Prevention's Modeling Infectious Disease in Healthcare Network.

## References

- Arik, S. O., Chun-Liang, L., Yoon, J., Sinha, R., Epsteyn, A., Le, L. T., Menon, V., Singh, S., Zhang, L., Yoder, N., Nokiltchev, M., Sonthalia, Y., Nakhost, H., Kanal, E., & Pfister, T. (n.d.). Interpretable sequence learning for COVID-19 forecasting. White paper. Google Cloud AI. Retrieved from <https://storage.googleapis.com/covid-external/COVID-19ForecastWhitePaper.pdf>.
- Castro, L., Fairchild, G., Michaud, I., & Osthus, D. (2020). Coffee: COVID-19 forecasts using fast evaluations and estimation. *Los Alamos National Laboratory*. <https://covid-19.bsvgateway.org/static/COFFEE-methodology.pdf>.
- Cavallo, J. J., Donoho, D. A., & Forman, H. P. (2020). Hospital capacity and operations in the coronavirus disease 2019 (COVID-19) pandemic—planning for the Nth patient. *JAMA Health Forum*, 1(3). <https://doi.org/10.1001/jamahealthforum.2020.0345>. e200345–e200345.
- Centers for Disease Control and Prevention. (2021a). Interpretation of forecasts of new hospitalizations. Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/hospitalizations-forecasts.html>.
- Centers for Disease Control and Prevention. (2021b). Testing...testing...1, 2, 3. COVID data tracker weekly review. *Interpretive summary*. Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html>.
- Covid Act Now. (2021, April 1). North Carolina. Retrieved from [https://covidactnow.org/us/north\\_carolina-nc/?s=1723645](https://covidactnow.org/us/north_carolina-nc/?s=1723645).
- Daugherty Biddison, E. L., Faden, R., Gwon, H. S., Mareiniss, D. P., Regenberg, A. C., Schoch-Spana, M., Schwartz, J., & Toner, E. S. (2019). Too many patients... a framework to guide statewide allocation of scarce mechanical ventilation during disasters. *Chest*, 155(4), 848–854. <https://doi.org/10.1016/j.chest.2018.09.025>

- Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba, Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L. C., van Elsland, S., Thompson, H., Verity, R., et al. (2020 March 16). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team. Retrieved from <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID-19-NPI-modelling-16-03-2020.pdf>.
- Gu, Y. (2020, October 6). COVID-19 projections. Retrieved from <https://covid19-projections.com/us-nc>.
- Holmdahl, I., Kahn, R., Hay, J. A., Buckee, C. O., & Mina, M. J. (2021). Estimation of transmission of COVID-19 in simulated nursing homes with frequent testing and immunity-based staffing. *JAMA Network Open*, 4(5), e2110071. <https://doi.org/10.1001/jamanetworkopen.2021.10071>
- Holmes, M., Campbell, H., & McKethan, A. (2020). Available hospital capacity and "runway" in North Carolina. COVID-19 in North Carolina. Brief 7. *The Cecil G. Shep Center for Health Services Research*. Retrieved from [https://www.shepscenter.unc.edu/wp-content/uploads/2020/04/Brief-7\\_12-8-2020.pdf](https://www.shepscenter.unc.edu/wp-content/uploads/2020/04/Brief-7_12-8-2020.pdf).
- Imhe. (2021). COVID-19 projections. *Hospital resource use*. Retrieved from [https://covid19.healthdata.org/global?view=resource-use&tab=trend&resource=all\\_resources](https://covid19.healthdata.org/global?view=resource-use&tab=trend&resource=all_resources).
- Imperial College London. (2020). COVID-19 model. Retrieved from <https://mrc-ide.github.io/covid19usa/#/details/NC>.
- Inglesby, T. V. (2020). Public health measures and the reproduction number of SARS-CoV-2. *JAMA*, 323(21), 2186–2187. <https://doi.org/10.1001/jama.2020.7878>
- Institute of Global Health. (2021). COVID-19 daily epidemic forecasting. Retrieved from [https://renkulab.shinyapps.io/COVID-19-Epidemic-Forecasting/\\_w\\_d371d176/\\_w\\_878ebf61/?tab=jhu\\_us\\_pred&country=North%20Carolina](https://renkulab.shinyapps.io/COVID-19-Epidemic-Forecasting/_w_d371d176/_w_878ebf61/?tab=jhu_us_pred&country=North%20Carolina).
- Jones, K., Hadley, E., Preiss, S., Kery, C., Baumgartner, P., Stoner, M., & Rhea, S. (2021). North Carolina COVID-19 agent-based model framework for hospitalization forecasting overview, design concepts, and details protocol. Retrieved from <https://arxiv.org/abs/2106.04461>.
- Jones, K., Munoz, B., Rineer, J. I., Bobashev, G., Hilscher, R., & Rhea, S. (2019). On calibrating a microsimulation of patient movement through a healthcare network. Retrieved from <https://www.informs-sim.org/wsc19papers/437.pdf>.
- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., & Shaman, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490), 489–493. <https://doi.org/10.1126/science.abb3221>
- Massachusetts General Hospital Institute for Technology Assessment. (2020). Policy simulator methodology. *Harvard Medical School*. Retrieved from [https://www.covid19sim.org/images/docs/Policy\\_simulator\\_methodology\\_20201001.pdf](https://www.covid19sim.org/images/docs/Policy_simulator_methodology_20201001.pdf).
- Maxmen, A. (2020). The race to unravel the biggest coronavirus outbreak in the United States. *Nature*. Retrieved from <https://www.nature.com/articles/d41586-020-00676-3>.
- Naimark, D., Mishra, S., Barrett, K., Khan, Y. A., Mac, S., Ximenes, R., & Sander, B. (2021). Simulation-based estimation of SARS-CoV-2 infections associated with school closures and community-based nonpharmaceutical interventions in Ontario, Canada. *JAMA Network Open*, 4(3), e213793. <https://doi.org/10.1001/jamanetworkopen.2021.3793>
- No author. (2020). Modeling of COVID-19 epidemic in the United States. Retrieved from [https://uploads-ssl.webflow.com/58e6558acc00ee8e4536c1f5/5e8bb44f5baae4c12a75d2\\_GLEAM\\_web.pdf](https://uploads-ssl.webflow.com/58e6558acc00ee8e4536c1f5/5e8bb44f5baae4c12a75d2_GLEAM_web.pdf).
- Patel, M. D., Rosenstrom, E., Ivy, J. S., Mayorga, M. E., Keskinocak, P., Boyce, R. M., Hassmiller Lich, K., Smith, R. L., 3rd, Johnson, K. T., Delamater, P. L., & Swann, J. L. (2021). Association of simulated COVID-19 vaccination and nonpharmaceutical interventions with infections, hospitalizations, and mortality. *JAMA Network Open*, 4(6), e2110782. <https://doi.org/10.1001/jamanetworkopen.2021.10782>
- Petersen, E., Koopmans, M., Go, U., Hamer, D. H., Petrosillo, N., Castelli, F., Storgaard, M., Al Khalili, S., & Simonsen, L. (2020). Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet Infectious Diseases*, 20(9), e238–e244. [https://doi.org/10.1016/S1473-3099\(20\)30484-9](https://doi.org/10.1016/S1473-3099(20)30484-9)
- Rahmandad, H., & Sterman, J. (2008). Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science*, 54(5), 998–1014. <https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1070.0787>
- Ray, E. L., Wattanachit, N., Niemi, J., Kanji, A. H., House, K., Cramer, E. Y., Bracher, J., Zheng, A., Yamana, T. K., Xiong, X., Woody, S., Wang, Y., Wang, L., Walraven, R. L., Tomar, V., Sherratt, K., Sheldon, D., Reiner, R. C., Prakash, B. A., Osthus, D., Li, M. L., et al. (2020). Ensemble forecasts of coronavirus disease 2019 (COVID-19) in the U.S. medRxiv, 2020.2008.2019.20177493. <https://doi.org/10.1101/2020.08.19.20177493>.
- Rees, E. M., Nightingale, E. S., Jafari, Y., Waterlow, N. R., Clifford, S. B., Pearson, C. A., Group, C. W., Jombart, T., Procter, S. R., & Knight, G. M. (2020). COVID-19 length of hospital stay: A systematic review and data synthesis. *BMC Med*, 18(1). <https://doi.org/10.1186/s12916-020-01726-3>, 270.
- Rhea, S., Hilscher, R., Rineer, J. I., Munoz, B., Jones, K., Endres-Dighe, S. M., DiBiase, L. M., Sickbert-Bennett, E. E., Weber, D. J., MacFarquhar, J. K., Dubendris, H., Bobashev, G., & CDC MInD-Healthcare Network. (2019). Creation of a geospatially explicit, agent-based model of a regional healthcare network with application to clostridioides difficile infection. *Health Security*, 17(4), 276–290.
- Rhea, S., Jones, K., Endres-Dighe, S., Munoz, B., Weber, D. J., Hilscher, R., MacFarquhar, J., Sickbert-Bennett, E., DiBiase, L., Marx, A., Rineer, J., Lewis, J., & Bobashev, G., for the, C. D. C. M.-H. N. (2020). Modeling inpatient and outpatient antibiotic stewardship interventions to reduce the burden of Clostridioides difficile infection in a regional healthcare network. *PLoS One*, 15(6). <https://doi.org/10.1371/journal.pone.0234031>, e0234031.
- Rodríguez, A., Tabassum, A., Cui, J., Xie, J., Ho, J., Agarwal, P., Adhikari, B., & Prakash, B. A. (2020). DeepCOVID3: An operational deep learning-driven framework for explainable real-time COVID-19 forecasting. medRxiv. <https://doi.org/10.1101/2020.09.28.20203109>, 2020.2009.2028.20203109.
- Rosenbaum, L. (2020). Facing Covid-19 in Italy — ethics, logistics, and therapeutics on the epidemic's front line. *New England Journal of Medicine*, 382(20), 1873–1875. <https://doi.org/10.1056/NEJMp2005492>
- Rosenberg, E. S., Tesoriero, J. M., Rosenthal, E. M., Chung, R., Barranco, M. A., Styer, L. M., Parker, M. M., Leung, S.-Y. J., Morne, J. E., Greene, D., Holtgrave, D. R., Hoefler, D., Kumar, J., Udo, T., Hutton, B., & Zucker, H. A. (2020). Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York. medRxiv. <https://doi.org/10.1101/2020.05.25.20113050>, 2020.2005.2025.20113050.
- Rti International. (2021). COVID-19. RTI international. Retrieved from <https://github.com/RTIInternational/nc-mind-covid-19>.
- Rtlive. (2020). Twitter & intragram collaborative. Rt covid-19. *Previous Rt Calculations*. Retrieved from <https://d14wlfuexugcm.cloudfront.net/covid/rt.csv>.
- Schenck, E. J., Turetz, M. L., & Niederman, M. S. (2020). Letter from the United States: A New York experience with COVID-19. *Respirology*, 25, 8, 900.
- Srivastava, A., Xu, T., & Prasanna, V. K. (2020). Fast and accurate forecasting of COVID-19 deaths using the SIK $\alpha$  model. *Cornell University*. Retrieved from <https://arxiv.org/abs/2007.05180>.
- Stradling, R. (2021, January 6). As COVID cases grow, NC hospitals are once again delaying non-emergency procedures. *The News & Observer*. Retrieved from <https://www.newsobserver.com/news/local/article248281650.html>.
- Sun, L. H., & Achenbach, J. (2020 June 25). CDC chief says coronavirus cases may be 10 times higher than reported. *Washington Post*. Retrieved from <https://www.washingtonpost.com/health/2020/06/25/coronavirus-cases-10-times-larger/>.
- Tange, O. (2020, June 22). GNU parallel 2020062 ('Floyd'). Zenodo. Retrieved from <https://zenodo.org/record/3903853#.YGY720hKhPZ>.
- University of Southern California. (2021). Evaluation of forecast hub for US state-level COVID-19 forecasts. Retrieved from <https://scc-usc.github.io/ReCOVER-COVID-19/#/leaderboard>.
- Ventura, C., Gibson, C., & Collier, G. D. (2020). Emergency medical services resource capacity and competency amid COVID-19 in the United States: Preliminary findings from a national survey. *Heliyon*, 6, 5, e03900.
- Wake Forest Baptist Health. (2021). COVID-19 community research partnership - study results and data. Retrieved from <https://www.wakehealth.edu/Coronavirus/COVID-19-Community-Research-Partnership/Updates-and-Data>.
- Weissman, G. E., Crane-Droesch, A., Chivers, C., Luong, T., Hanish, A., Levy, M. Z., Lubken, J., Becker, M., Draugelis, M. E., Anesi, G. L., Brennan, P. J., Christie, J. D., Hanson, C. W., 3rd, Mikkelsen, M. E., & Halpern, S. D. (2020). Locally informed simulation to predict hospital capacity needs during the COVID-19 pandemic. *Annals of Internal Medicine*, 173(1), 21–28. <https://doi.org/10.7326/M20-1260>
- Zeek, A. (2020 May 13). IU, ISDH release preliminary findings about impact of COVID-19 in Indiana. *News at IUPUI*. Retrieved from <https://news.iu.edu/stories/2020/05/iupui/releases/13-preliminary-findings-impact-covid-19-indiana-coronavirus.html>.