Iterative near-term forecasting of the transmission and management of SARS-CoV-2/COVID-19 using social interventions at the county-level in the United States

Kenneth Newcomb  
University of South Florida

Morgan E. Smith  
University of Notre Dame

Rose E. Donohue  
University of Notre Dame

Sebastian Wyngaard  
University of Notre Dame

Caleb Reinking  
University of Notre Dame

Christopher R. Sweet  
University of Notre Dame

Marissa J. Levine  
University of South Florida

Thomas R. Unnasch  
University of South Florida

Edwin Michael (emichael443@usf.edu)  
University of South Florida  https://orcid.org/0000-0002-9473-4245

Research Article

Keywords: SARS-CoV-2, data-driven mathematical modelling, Bayesian Melding-based sequential model calibration, near-term epidemic forecasting, non-pharmaceutical mitigation, geographically-targeted pandemic control

Posted Date: October 23rd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-96665/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

The control of the initial outbreak and spread of SARS-CoV-2/COVID-19 by the implementation of unprecedented population-wide non-pharmaceutical mitigation measures has led to remarkable success in dampening the pandemic globally. With many countries easing or beginning to lift these measures to restart activities presently, concern is growing regarding the impacts that such reopening of societies could have on the subsequent transmission of the virus. While mathematical models of COVID-19 transmission have played important roles in evaluating the general population-level impacts of these measures for curbing virus transmission, a key need is for models that are able to effectively capture the effects of the spatial and social heterogeneities that drive the epidemic dynamics observed at the local community level. Iterative near-term forecasting that uses new incoming epidemiological and social behavioural data to sequentially update locally-applicable transmission models can overcome this gap, potentially leading to better predictions and intervention actions. Here, we present the development of one such data-driven iterative modelling tool based on publically-available data and an extended SEIR model for forecasting SARS-CoV-2 at the county level in the United States, and demonstrate, using data from the state of Florida, how this tool can be used to explore the outcomes of the social measures proposed for containing the course of the pandemic as a result of easing the initially imposed lockdown in the state. We provide comprehensive results showing the use of the locally identified models for accessing the impacts and societal tradeoffs of using specific strategies involving movement restriction, social distancing and mass testing, and conclude that while it is absolutely vital to continue with these measures over the near-term and likely to the end of March 2021 in all counties for containing the ongoing pandemic before less socially-disruptive vaccination strategies come into play, it could be possible to lift the more disruptive movement restriction/social distancing measures by end of December 2020 if these are accompanied by widespread testing and contact tracing. Our findings further show that such intensified social interventions could potentially also bring about the control of the epidemic in low and some medium incidence counties first, supporting the development and deployment of a geographically-phased approach to reopening the economy of Florida. We have made our data-driven forecasting system publicly available for policymakers and health officials to use in their own locales, with the hope that a more efficient coordinated strategy for controlling SARS-CoV-2 state-wide, based on effective control of viral transmission at the county level, can be developed and successfully implemented.

Introduction

The implementation of population-wide non-pharmaceutical socially-based suppressive measures, focused on lockdowns of whole communities, social distancing, travel restrictions, and increasingly the deployment of testing and contact tracing, has led to remarkable success in dampening the initial waves of the ongoing severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2/COVID-19) pandemic globally (https://aatishb.com/coviddtrends/). This has resulted in many countries considering the lifting of these unprecedented behavioral measures to allow the re-opening of their economies while awaiting the arrival of less socially-disruptive technologies, such as viable vaccines. Health officials are, however, becoming increasingly concerned about the likely impacts that such re-openings could have on the subsequent transmission dynamics of the virus. One reason for this concern is that in many areas, easing of the above non-pharmaceutical interventions (NPIs) has taken or are taking place before the initial epidemics have reached their endpoints. Another is that herd immunity in communities that have been under these interventions has not developed to levels that would mitigate the possibility of infection resurgences. These possibilities, including the economic and political imperatives for easing these NPIs, have led to increased attention being paid to the identification and deployment of those social mitigation measures that will enable the containment of viral transmission to levels that would allow reopening of societies with minimal harmful health and social side-effects. These currently include the prolonging of social distancing measures, a phased approach to easing current movement restrictions, and ramping up of testing and contact tracing capacity for identifying and isolating the remaining or new infectious cases.

A distinctive feature of the policy response to the management of the COVID-19 pandemic worldwide has been the role played by epidemiological modelling for evaluating the use of behavioural interventions exclusively for controlling epidemic outbreaks in populations. These mathematical models, based primarily thus far on extensions to the standard SEIR compartmental model, have enabled predictions of the course of the epidemic to warn policy-makers of the gravity of potential impacts, as well as help them in making comparisons of the various social measures proposed for suppressing viral transmission in exposed communities. For example, these tools have played critical roles for evaluating the comparative effects of locking down communities versus allowing a portion of the population to be exposed and develop immunity as alternate strategies for containing both first and subsequent epidemic waves.

However, while these models have been important for supporting generalized, scenario-based, policy planning, they can be less useful for simulating dynamics after the disease breaks out in real communities. First, such models have little support for capturing the rapidly changing social contexts in which an epidemic evolves. This is an important need since for fast evolving phenomena, such as an epidemic, feedbacks from changing conditions need to be processed speedily to correct for the invariable modelling errors surrounding a dynamic forecast. Second, addressing forecasts of the near future as an epidemic evolves requires models that give accurate forecasts with the least variance at the lead times required by management. Third, general models, even if supported by the available historical data, do not address the need for data assimilation in real-time forecasting. Such data assimilation, in which information regarding the extant transmission processes that are embedded in observational data is used to iteratively update the underlying dynamical principles represented by the structure and parameters of a model, has been shown to provide near-term forecasts of the state of a dynamical system which are better than could be obtained with just data or the model alone. For forecasting epidemics over the near future, this data-model assimilation framework will allow shrinkage of forecast variance while also correcting for model bias and drift. Finally, few currently available mathematical models of SARS-CoV-2 transmission provide predictions at lower administrative levels, such as the county. Thus, they do
Developing, testing and refining data-driven models for enabling reliable near- and long-term forecasts for decision-making can be a challenge especially when the objective is to control an ongoing epidemic. One problem is the requirement that locally-relevant observations are transmitted sufficiently rapidly to be useful for updating models within the lead times connected with making an effective public response. Fortunately, advances in communication systems, powerful data processing, storage software and hardware, and an increasing focus on the provision of standardized, open-source data for diseases, are now becoming available that are providing solutions to this problem. In parallel, major progress has also been made in the development of iterative statistical data-model assimilation techniques, whereby data of diverse types and prior information regarding model structures and parameters can be used reliably to constrain model parameters or states in a setting, including supporting evaluations of forecast uncertainty over time. Lastly, developments in cyberinfrastructures to automate the dynamic integration of new data and information to facilitate regular assessment of forecasts and active updating of models mean that the practical implementation of iterative data-driven epidemic forecasting is now increasingly becoming possible.

As countries relax the social restrictions imposed to suppress the initial pandemic waves, it is becoming evident that it is important to consider and predict the subsequent pathway that SARS-CoV-2 might follow. This is because, given the intensity and size of the initial outbreaks in a locality, the epidemic could simply be eradicated locally by the implementation of intensified social measures. Alternatively, especially in high incidence areas, these interventions need to be calibrated both in terms of duration and intensity to ensure that resurgences can be managed to ensure that critical care capacities are not exceeded. Distinguishing between these outcomes at local scales will therefore be key for formulating the requisite public health response for managing SARS-CoV-2. Data-driven locally applicable COVID-19 models that are regularly updated using the latest epidemiological data will clearly constitute a powerful tool for making these spatially-explicit assessments.

This paper describes the efforts of our team to develop and use such an iterative data-model assimilation-based forecasting system for SARS-CoV-2, wherein we use a SEIR-type model updated sequentially with publically available COVID-19 case and human movement data in order to provide predictions of the course of the pandemic under various social interventions at the county level in the United States. We use these data from counties in the state of Florida to demonstrate how the developed data-driven modelling system can allow forecasts and assessments of the local epidemic dynamics as well as enable evaluations of the relative impacts of the social interventions — viz. a phased easing of the lockdown, inclusion of different durations of social distancing measures, and tracing and isolation of infectious individuals — proposed for curbing viral transmission in the near-term until effective vaccination strategies can come into play in this state. We also examine how gaining a better understanding of the geographical variation in the propagation of the virus and our attempts to curtail it can allow the derivation of a regionally-varying or tailored response that is effective at minimizing contagion whilst offering at the same time the advantages of less restrictive rules for parts of a state's population.

Results

Model fits to data, validation, and assessment of geographic drivers of transmission

We modeled the observed SARS-CoV-2 outbreaks in the 67 counties of Florida (Table S1) based on reported case data, and information regarding local social behaviors, through to September 30th, 2020. This was carried out by using a Bayesian-melding data assimilation approach centered on the sequential calibration of our compartmental epidemic model to confirmed case and death time series data, and data on human movements, reported for each of these counties by the Johns Hopkins University Coronavirus Resource Center and the location data firm, Unacast. The model predictions for cumulative confirmed cases compared to data are shown in Figure 1. The results show that the 95% confidence interval bounds of the predictions from the ensemble of sequentially updated models for each county are able to envelop nearly all of the confirmed case data to date.

Because the dynamics of COVID-19 are significantly influenced by changes in social behaviours, it is critical to iteratively calibrate transmission models, such as the present SEIR model, to new data and actively update the resulting predictions accordingly. Figure S1 demonstrates the shift in model prediction performance as the model is updated sequentially with various lengths of incoming longitudinal data. Cross-validation analysis showed that carrying out such model calibrations to 2-week sequential blocks of incoming data can keep the relative mean-square error to consistently below 20% while also being computationally feasible. Implementing this 2-week sequential updating procedure also allowed us to incorporate information regarding changing transmission conditions in each county as rapidly as possible into the model.

Table S2 summarizes the values of the posteriors obtained by model fitting to the 14-day case day prior to September 30th for the social distancing parameter, \( d \), which modifies the transmission rate, \( \beta \), in each county, as well as for the fraction of the respective county population remaining under mobility restriction as estimated using the Unacast mobility data. The results show that while the values of each of these two key social parameters varied between the present counties and therefore are associated with the differences in the virus transmission dynamics observed for individual counties, the social distancing parameter, \( d \), appears to be comparatively less variable in its estimated between-county values compared to those estimated or calculated for the lockdown fractions (Table S2). This suggests that the between-county variations in SARS-CoV-2 outbreak dynamics, and the individual county-level response to the intervention scenarios, reported here (from October 1st 2020) may be a reflection more of the
Forecasting the epidemic and social interventions

We used the models updated using the cumulative infection/deaths data reported to September 30th in each county to simulate both the local epidemic dynamics and compare the dynamical impacts of six different social interventions as described in Methods and depicted in Figure S2. The model predictions under each intervention scenario through the end of the year are shown in Figures 2 and 3, stratified by initial incidence growth rate (Group a < 0.05, Group b < 0.15, Group c > 0.15; see Methods). Scenario 1 is the least aggressive option considered with no interventions put in place and a full release of strict stay-at-home orders after September 30th. The results for this scenario show that because the social restrictions in Florida were lifted before the first epidemics ended, resurgence of the epidemic will be inevitable in every county (Figure 2). Indeed, an average of 27% (with a range of 15-47%) of the overall population across all counties is projected to become infected at the peaks of the resulting 2nd wave local epidemics under this scenario (Figure 2, Table 1, Table S1). Note that these and subsequent model forecasts account for all infected cases, including those that are not yet infectious (exposed class), asymptomatic, presymptomatic, and/or not detected by health systems. The county level model predictions from this scenario also highlight the variation in the peak size of the 2nd waves that different counties could have faced with the lifting of all interventions, with these sizes ranging from 1,459 cases in Lafayette to as high as 742,898 cases in Miami-Dade county (Table 1). The time course of the epidemic is also variable with higher incidence counties expected to see generally later peaks compared to lower incidence counties (see legend to Figure 1 for peak dates). Figure 2d shows that the predicted size of the 2nd wave peaks are directly related to those of the first waves that occurred in each county, further underlining the impact that initial variation in local conditions of virus transmission can have on the size of subsequent county-level infection resurgences in each county. The simulations also reveal that while the size of the 2nd waves will be large with the full release of lockdowns, the epidemics in each county will nonetheless, as expected, eventually fade and end (Figure 2), with the possibility that this will occur earlier between December 2020 to January 2021 in the case of the low incidence counties compared to the case with high incidence counties where the corresponding 2nd waves are predicted to end as late as early March 2021.

The inclusion of social distancing measures through October and November in scenarios 2 and 3, respectively, is predicted to have two key effects. First, it is to be noted that these measures will not prevent the occurrence of sizable 2nd waves; however, they will shift the timing of the county-level 2nd epidemic peaks further into the future, with this shift more pronounced for scenario 3 (from the October/November peaks predicted for scenario 2 to the December/January peaks forecasted for scenario 3; see legend to Figure 3). These measures, however, will significantly reduce the size of the 2nd epidemic peaks, with the more intensive scenario 3 bringing about an average case reduction of 76% (with a range of 37-83%) compared to scenario 1 (Table 1). Again, these outcomes will vary by county group, with shifts to 2nd peaks, and resolution of the epidemics, occurring generally later among the high incidence countries (Group c), while the greatest reduction in peak cases will occur for the low incidence counties (Figure 3, Table 1).

Implementation of contact tracing and quarantine measures to prevent a fraction of the infectious population from spreading the disease (25% quarantine rate in scenario 4 and 50% quarantine rate in scenario 5 both through March 2021) along with sustained social distancing measures through December 2020, is predicted to have a uniformly high suppressive effect on the course of the 2nd epidemics in each county. Scenario 4 reduces the size of the averaged county-level epidemic peak to affect just 3% of the total population (with a range of 0-17%), while scenario 5 is predicted to reduce the average peak infection size to just 0.6% of the overall population in the state (with a range of 0-8.5%). This represents an average peak reduction compared to scenario 1 of 89% (with a range of 61-97%) for scenario 4 and 98% (with a range of 80-100%) for scenario 5. Furthermore, the results show that scenario 5 could even bring about breakage of epidemic transmission in some counties, particularly in the case of those that exhibited the lowest initial infection incidences (Table 1). Interestingly, these intervention scenarios also appear to generally lessen the between-county group variations in the timings and peaks of the predicted 2nd epidemics (Figure 3).

The above scenario differences are also apparent in the predictions of the required hospitalizations at the county level (Figure 4). However, the hospitalization forecasts also indicate that in the absence of more aggressive interventions, such as those modeled in scenarios 4 and 5, there will be a high risk that the predicted cases will exceed existing county-level hospital capacities (Table 1). This serious outcome will also vary significantly by geographic location.

Figure 5 shows the predictions arising from the simulation of the most intensive of the present intervention scenarios (scenario 6), viz. maintaining current social distancing, lockdown, and a 25% quarantine rate from October 1st through the end of our simulation period, March 2021. The results show that this scenario is the only one among the six scenarios investigated that will prevent the occurrence of a 2nd wave of COVID–19 in all the currently modeled counties. It would also hasten the ending of the epidemic locally with all low incidence counties predicted to achieve fade-outs of their epidemics as early as between November 1st 2020 to January 26th 2021, whereas high incidence counties will see their corresponding epidemics fading out between December 20th 2020 to February 25th 2021.

Figure 6 depicts the proportions of the populations recovering from infection and thus developing immunity to infection through time in each county for scenarios 1 (top panel) and 6 (bottom panel). These results show that these scenarios may bring about extinction of the epidemic in each group of counties via different mechanisms. In the case of scenario 1, the epidemics are ended through the development of high levels of herd immunity (between 88 to 97%) in the community, with as expected lower population-level immunity required to bring out epidemic extinction in the lower incidence...
counties. By contrast, the results for scenario 6 indicate that extinction can also be brought about by instituting strong long-duration social mitigation measures that can reduce transmission to sufficiently low levels to bring about epidemic fade outs. However, it is important to note that this impact comes with the cost of generating very low levels of herd immunity in each community by the end of the epidemics, raising the possibility of the inevitable resurgence of transmission should new infecteds bring the virus back into these communities.

Although scenario 1 can produce high levels of herd immunity (Figure 6), it is clear, however, that this will come with higher death tolls than in the case of scenario 6 (Figure 7). Cumulative predicted deaths through the entire period of these simulations (i.e. October 1st 2020 to end of March 2021) ranged from 70 to 4,223 in the low incidence counties to as high as 28,130 in the high incidence counties in the case of scenario 1. These were significantly lower in the case of scenario 6 with the corresponding cumulative deaths ranging from 6 to 300 in the low incidence counties to between 185 to 3,330 in the high incidence counties (Figure 7). These findings underscore the health-economy trade-offs involved in using an approach focused on the evolution of herd immunity as opposed to one based on the use of more socially-disruptive mitigation measures for containing the present pandemic.

**Discussion**

Our goals in this work were essentially two-fold. The first was to assess if it is possible to use publically available longitudinal infection case and human movement data to derive reasonable mathematical models of SARS-CoV-2 transmission to allow the simulation and evaluation of the course of the ongoing pandemic at the local county level in the United States. The second goal was to evaluate how the contagion dynamics operating at the local county level might interact with the currently implemented or proposed social options for controlling COVID-19, such that the results may be used to identify those measures that will enable the containment of the virus to levels allowing the safe reopening of communities whilst we await the arrival of viable vaccines. We have also attempted to determine the implications of variance in virus transmission risk at smaller spatial units within a region, such as counties across a state, for the design of intervention strategies to curb the present contagion. We show in this regard that compared to a common control strategy applied uniformly across all counties, minimizing both human health impacts as well as disruptions to economic and other social activities across a state could be better achieved through the deployment of a geographically-targeted approach that explicitly takes into account the impacts of locally-variable risks of virus transmission. These results also question the use of homogenous aggregate level models for reliably predicting and managing the epidemic over a broader spatial region, such as a state, across which the processes that govern the transmission of the virus are likely to vary.

Developing and using reliable data-driven models for forecasting live local epidemics is challenging given the need for both the locality-specific temporal data required for updating models, and the necessity that predictions have to be made at frequencies or within the lead times requisite for making effective public responses. Here, we have addressed this problem for modelling SARS-CoV-2 transmission to support policy making via the implementation of an iterative data-model assimilation-based forecasting system that acquires and processes the latest data, updates our SEIR COVID-19 model, and generates new sequential forecasts over time. Key features of the system include procedures that leverage the availability of open source API-enabled case and mortality surveillance data that are reported daily by health departments at the county level, the incorporation of independently-quantified county-wide non-essential movement data to serve as an estimator of the level of population mixing over time, and the Bayesian calibration of our model on a sequential basis. An additional recent feature of this system is also the use of a continuous analysis framework to automate the computational pipeline to handle the various stages of converting the raw data into new forecasts, including: data assembly, modelling and forecasting, and presentation of the forecasts relevant to policy makers. The cross-validation exercise we carried out for assessing the variance of future forecasts using our iterative or sequential model updating procedure (Figure S1) show how the longer our forecasts remain out of date, the less accurate they become. Even with our choice of using 2-weekly sequential blocks of data to update our model here, we found that forecasts can result in an error rate of 20%. This outcome highlights the value of regular assessment and active updating of dynamic epidemic models for providing the forecasts required by policy-makers for decision-making.

We have used our modelling system to examine a range of likely SARs-CoV-2 transmission scenarios through March 2021, starting with simulations of the impact of the state-wide phased easing of the lockdown that was implemented in Florida between April 3 2020 to May 4 2020 (scenario 1). Figure 2 depicts the major outcome of this policy response, viz. the inevitability of the emergence of significant 2nd waves in all counties if all other social measures, such as social distancing (mask wearing and observation of physical distancing), are also discontinued fully after September 30th 2020. While this prediction can be considered to be as expected and so unremarkable, a striking and possibly less commented upon feature, however, is that the predicted sizes of the 2nd wave peaks or intensity of the 2nd waves will vary between counties as a positive function of variations in the size of their 1st wave peaks. This indicates that the subsequent intensity of virus transmission following the full release of all social protective measures in a community will depend fundamentally on the initial incidence at the time of epidemic establishment in a locality. It will additionally also depend on the number of infecteds remaining in each county following the ending of the state-wide lockdown that took place before the 1st waves had been fully controlled. These findings clearly highlight the dangers of imposing a one size fits all policy (here pertaining to the decision to ease lockdowns on a common date across Florida) for managing a spatially variable contagion.

The results pertaining to the numerical size of the 2nd peaks are shown in Table 1, and indicate that the health burden of the pandemic will also vary markedly between counties from as low as 1,459 cases in Lafayette to as high as 742,898 cases for the most populous Miami-Dade county, further underscoring the fact that apart from variable resurgences in infection, a wide spatial heterogeneity in infection burdens can also be expected if social measures are fully released in all counties. There will also be considerable variations in the course of the 2nd waves even within each incidence group.
Perhaps the first need is to address what impact the impending arrival of viable vaccines would have on the need for continuing with the social measures. We end by highlighting the future work that will be required to overcome the current limitations and applicability of our modelling system.

Our simulations of the impacts of social control measures of varying strength and nature demonstrate overall the vital importance of continuing with these measures following phased lockdown release for containing the epidemic while waiting for more effective and less-socially disruptive pharmaceutical measures. According to our findings, while the two social measures investigated here, viz. maintaining current social distancing over a shorter (to October 14th) and longer (to November 30th) periods, and testing, contact tracing and quarantining at a moderate (25%) or higher (50%) level, will not prevent the emergence of 2nd waves in the majority of cases, they will substantially delay infection peaks and reduce the numbers of patients requiring admissions to hospital, and in the case of the stronger versions of these measures also dramatically suppress the size of the 2nd wave peak cases/hospitalizations. Indeed, the inclusion of quarantine measures to March 2021, while ending social distancing by mid-October or end November 2020, for example, would decrease the 2nd wave peak numbers of infection or hospitalizations required to over 90% of the respective 1st peak numbers in each county. Inclusion of strong contracting tracing and quarantine (scenario 5) can lead to very low levels or even interruptions of epidemic transmission and zero hospitalization cases in those counties exhibiting the lowest incidence rates among the present counties. This is an important outcome as it suggests that if testing and contact tracing can be ramped up, then counties will be able to lift their social distancing measures and hence reopen their economies by end of November 2021 without fear of such an option would lead to a catastrophic overwhelming of their existing hospital capacities.

We show that the most intensive social intervention modelled in this study, viz. phased lockdown release, maintenance of current social distancing measures and a 25% quarantine rate all maintained from October 1st to end of March 2021, is the only social option that would not only prevent the occurrence of a 2nd wave, but also hasten the sustained ending of the epidemic in all counties, with fade-outs possible as early as by November 1st in some low-medium incidence counties (Figure 5). This finding underscores the need for combining and implementing social measures as early as possible and maintaining them throughout the next few months (to end of March 2021 at least) for achieving the eradication of the pandemic in Florida. However, like scenarios 4 and 5 but unlike the full release of social measures modelled in scenario 1, this scenario will also be marked by the low level of herd immunity that would develop in each county population by the end of the local epidemics, leaving such communities vulnerable to the real threat of future epidemic resurgences should the virus be re-introduced after the lifting of interventions (Figure 6). This finding indicates that either maintaining continued vigilance and control by testing and contact tracing measures will be required to counter this prospect of epidemic resurgence in these communities over the foreseeable future, or that ultimately, control of the epidemic will only be achieved through effective vaccination of county populations. Our results show that in the latter case, vaccination rates (with a highly effective vaccine) will need to be above 85% and even above 90% in the medium-high incidence counties (Figure 6) to accomplish the resolution of the pandemic, although note that if significant population heterogeneity underlines virus transmission within a county, much lower rates (to as low as 50%) might be sufficient to arrest the local epidemic.

Our findings, finally, are insightful with regard to the need to address the spatial heterogeneity of pandemic risk for effectively managing the SARS-CoV-2 both locally and globally. In particular, our results suggest that changing from using uniform policy approaches to the making of tailored responses might be more effective at minimizing harmful effects in communities, not just in relation to health outcomes, but also in terms of minimizing the disruption to the local and global economic and other social systems. Thus, we show that while combining social distancing measures to end of December 2020 with high intensity contact tracing and quarantine to March 2021 (scenario 5) will depress hospitalization cases drastically and within manageable levels across virtually all county incidence groups, it might be possible to contain the pandemics in some low and medium incidence counties with a version of this scenario (scenario 4) that implements only low intensity quarantine (Figure 4, Table 1). This will allow reopening of the economies of these counties earlier than for high incidence counties, lessening the economic and other social disruptions faced by the populations of these counties. Similarly, our predictions of the impact of scenario 6, in which all interventions are implemented from October 1st onwards, indicate that resolutions of the epidemic will occur significantly earlier in low incidence counties than in the case of medium and high incidence counties, suggesting that a safe reopening of the state of Florida and indeed other US states can be effectively accomplished in a geographically phased manner that takes into account county-level variations in epidemic risk explicitly. Indeed, our web-based SEIRcast COVID-19 simulation tool (https://seircast.org/) that implements our iterative data-driven continuous integration modelling framework, is designed to provide policymakers with the means to devise precisely such spatially-explicit management plans. We believe that including this spatial dimension into both models and in mitigation plans will provide not only better predictions of the pandemic dynamics across a spatial domain but will also offer significantly better overall social outcomes for populations.

We end by highlighting the future work that will be required to overcome the current limitations and applicability of our modelling system. Perhaps a first need is to address what impact the impending arrival of viable vaccines would have on the need for continuing with the social measures.
investigated in this paper to allow the safe reopening of parts of the populations in Florida as early as possible. The key question here is whether our findings would still apply over the near-future (until at least end of March 2021) if these vaccines become available for use as expected by early 2021. We believe that given the present uncertainties regarding when production and widespread delivery can be achieved even if first vaccines are released at that time, the likely moderate efficacies expected for these vaccines, and the likely intensification of general viral circulation in the coming autumn and winter seasons, the present NPI strategies will need to continue and indeed have to remain the mainstay of our attempts to contain the contagion over the next few months to at least the end of spring 2021. Indeed, if the coming vaccines are not perfect but instead reduces susceptibility by a fraction, and if the immunity induced is not long-term, then there is a need to investigate how best to adapt the social measures studied here along with vaccination to bring about the containment or resolution of the pandemic effectively 47. We are currently extending our model to include these various vaccination scenarios to address this fast emerging policy question.

It is also clear going forward that we need to consider the effects that between-county movement might have on the current model predictions. While personal movement was curtailed drastically by lockdown, and the phased ending of the lockdown has led to increased movement within counties – both of which we have been able to incorporate into our model via parameterization of the within-county movement data provided by Unacast - details of inter-county movement and its reliable incorporation into our model will be required if we are to better capture the impacts of state-wide policies that are beginning to focus on lifting of all restrictions fully 48. Recently, Unacast 43 has begun to publish population migration data in the US using cell-phone signals, which will provide a means to address this topic.

Our current model also does not represent the age-structure and health status of the county-level populations. Partly this is an outcome of our goal to develop a modelling system that would support the generation of forecasts for the contagion in all counties of the United States based on the data presently publically available for facilitating model configurations - and these currently lack information on these variables 38. Extending our SEIR model to include these features, however, would allow better treatments of the exposure, risk, and transmission conditions that are likely to underlie the spatial heterogeneity in epidemic dynamics observed at the county level 15,24,46. The addition of population structure and health composition into our current SEIR model will require deriving and adding more compartments and the applicable contact matrices 7,49,50, but also, as noted, the configuration data for parameterizing these additions appropriately. We are currently in the process of adapting the data from the POLYMOD study 44,50 to begin the construction of the relevant social contact matrices and parameterizations for accomplishing these major extensions to the model.

Our sequential data-assimilation framework, while allowing the incorporation of longitudinal changes in transmission conditions into the model, has the outcome, as for all dynamical models, that prediction error will increase the further out of sample forecasts are made 16-19,21,22. While we have attempted to reduce forecast errors by model fitting to two sets of variables (cumulative infectious cases and deaths), obtaining new data on other currently latent states (eg. the fraction of asymptomatic infecteds) would offer better constraining of parameters and hence forecast variance. However, this has to be balanced by appropriately addressing the effects of parameter degeneracy and sample impoverishment which would impact the ability of the model to fit novel data as transmission conditions change drastically over the near future 28,51-53. We have used a resampling approach whereby at each sequential updating point, we have blended in 25% random samples from initial priors to the posteriors obtained during the uptake made a time step (every 2-weeks) previously to keep forecast error below 20% to address this problem in the simulations reported here. However, future work might need to consider the development of appropriate adaptive approaches developed in the field of particle filtering 34 to resolve this problem more effectively. Regardless, we note that while our forecasts beyond 2 weeks ahead could attain variances as high as 40%, and so can affect the peak sizes and extinction dates reported here, this will have lesser impacts on the conclusions reached regarding the comparative outcomes of the interventions investigated in this study.

### Methods

### Epidemic model

We simulated the ongoing SARS-CoV-2 outbreaks at the county level using a variation of a SEIR model. The model compartments and transitions are shown in Figure S3. Full equations are also given in Supplementary Material. We assume each county is a closed population and ignore demographic changes such that the total population size remains constant. The population is divided into compartments representing various infection stages: susceptible (S), susceptible but removed from the transmission process via lockdown policies (R₁), exposed (E), infectious asymptomatic (Iₐ), infectious pre-symptomatic (Iₚ), infectious with mild symptoms (Iₘ), infectious with severe symptoms requiring hospitalization (Iₕ), infectious with severe symptoms requiring intensive care including ventilation (Iₗ), recovered and immune (R₂), and deceased (D). The model equations describing the transitions in and out of each class are given in the appendix along with a description of all the model parameters and their prior values (Table S3). Note the model considers the fraction of the population classified in each compartment (all compartments sum to 1), which is then scaled to the appropriate county population size to get counts for each compartment.

### Data

To calibrate the model to the local county setting, we fitted the SEIR model sequentially (see below) to cumulative confirmed case and deaths data assembled from the start of the epidemic at the county level and published for public access by the Johns Hopkins University Coronavirus Resource
Estimation of initial epidemic growth rate

The initial incidence growth rate, $\tau$, was estimated by fitting a log-linear model to the daily new cases reported during the early exponential phase (the first 4 weeks generally) of the epidemic curve observed in each county. The values estimated for $\tau$ in each county were used to stratify the counties in Florida into each of initially low (<0.05), medium (0.05 to <0.15) and high (>0.15) incidence or epidemic groups.

Bayesian Melding data assimilation

We used a Monte-Carlo-based Bayesian Melding framework to undertake the sequential updating of the model to the cumulative case and death data. We began by first defining uniform prior distributions for each of the model parameters based on current understanding of SARS-CoV-2 transmission and disease characteristics. The initial parameter priors and relevant references are given in Table S3. Note that the number of initial infected cases at the start of the simulation period is sampled as the parameter $E_0$ (Figure S4), the number of exposed cases when the first cases began to be confirmed. We consider the start of the epidemic in a given county to be when there are at least 10 cases reported. At this point, we sampled $N = 50,000$ parameter vectors from the initial priors and simulate the outbreak for 14 days forward. The resulting 50,000 model predictions of the epidemic are then compared to the confirmed case and death data observed during the 14-day forecast period using a normalized root-mean-square error distance metric that normalizes a traditional RMSE by the standard deviations of these data. This facilitated the combination of prediction errors with respect to case and death data together despite their different orders of magnitude:

$$NRMSE = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} \left( \frac{\hat{y}_i - y_i}{\text{std}(y)} \right)^2 + \left( \frac{\hat{x}_i - x_i}{\text{std}(x)} \right)^2}$$

where $n$ is the number of time points over which to compare the model predictions to data, $\hat{y}_i$ is the model-predicted confirmed case data on a given date $i$, and $y_i$ is the observed confirmed case count for the same date, $\hat{x}_i$ is the model-predicted death data on a given date $i$, and $x_i$ is the observed death count for the same date. Based on this performance metric, the best-fitting 500 parameter vectors are retained as the most likely parameter sets to describe the local outbreak during the chosen 14-day window. For simulating the epidemic for the next 14 days, another 50,000 parameters are sampled of which 75% are randomly sampled from the posterior distribution of the most recent 14 day window, while another 25% are sampled from the initial parameter priors to avoid sample depletion. These set of blended parameter vectors are used to sequentially select the best-fitting models over time, and are used to forecast the impacts of the interventions described above. Different fitting windows were tried, and a 14-day window was found to be long enough to be computationally feasible for the entire dataset for all counties, while being short enough to capture the changing epidemic behavior and keep forecast error consistently low (below 20%).

Simulating interventions

We used the latest sequentially fitted model in each county to simulate the impacts of different social intervention scenarios on the course of the outbreak in the future (beyond October 1st). We simulated six different scenarios, which are outlined graphically in Figure S2. Scenario 1 represents the least aggressive option where lockdown and social distancing measures are fully lifted after September 30th. Scenario 2 considers this lockdown release in addition to keeping social distancing measures in place for 2 weeks from October 1st to October 14th. We consider scenarios 1 and 2 to mimic the State of Florida’s current state reopening plan (https://floridahealthcovid19.gov/plan-for-floridas-recovery/). Scenario 3 extends the social distancing intervention by maintaining it over a longer 8 week period to November 30th. Scenarios 4 and 5 represent maintaining current social distancing and movement restrictions through the end of the year (December 2020) in addition to implementing contact tracing and quarantine efforts at either low ($q = 0.25$) or high ($q = 0.50$) intensity, respectively, from October 1st to end of March 2021. Finally, Scenario 6 represents the most intense intervention scenario, viz. maintaining social distancing, lockdown, and low quarantine starting from October 1st through to the end of March 2021. Outbreaks were simulated under these conditions and the predicted number of cases are compared between each scenario and county group. The numbers of hospitalized ($I_{H} + H_{C}$) are also forecasted to evaluate the potential resource needs under each scenario.

The effect of statewide lockdown measures in these simulations is implemented by adding a distinct susceptible class which is assumed to not contribute to disease transmission ($R_0$). The proportion of this class that complies with strict stay-at-home orders is controlled through the ratio of parameters $a$ and $\lambda$. This ratio is informed by the fraction of non-essential trips made by the population in each county estimated by Unacast based on analyses of GPS mobility data. We interpret the reduction in such on-essential trips from prior to the lockdown as a proxy for the proportion of the susceptible population remaining unexposed in a county during any time of a simulation – both during the lockdown and after lockdown measures were lifted. Social distancing measures are modeled as a reduction in transmissibility of the pathogen through the parameter $d$ and is primarily based on the
effectiveness of masks against transmission of similar diseases. Quarantine of infectious cases through contact tracing and/or testing is modeled simply as a proportion, $q$, of $I_A$, $I_P$, and $I_M$ as not contributing to transmission as a result of being detected and made to isolate themselves at home.

Availability of materials and data

All data analysed during this study are included in this published article and its supplementary information files. Code for sequentially calibrating and running the SEIR model is available at https://github.com/EdwinMichaelLab/COVID-SEIR-Paper

References

1. Kupferschmidt, K. The lockdowns worked—but what comes next? Science 368, 218-219, doi:10.1126/science.368.6488.218 (2020).
2. Byambasuren, O. et al. Estimating the seroprevalence of SARS-CoV-2 infections: systematic review. medRxiv, doi:10.1101/2020.07.13.20153163 (2020).
3. Fontanet, A. & Cauchemez, S. COVID-19 herd immunity: where are we? Nat Rev Immunol 20, 583-584, doi:10.1038/s41577-020-00451-5 (2020).
4. Chowdhury, R. et al. Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries. Eur J Epidemiol 35, 389-399, doi:10.1007/s10654-020-00649-w (2020).
5. Giordano, G. et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nat Med 26, 855-860, doi:10.1038/s41591-020-00683-4 (2020).
6. Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 368, 860, doi:10.1126/science.abb5793 (2020).
7. Prem, K. et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Publ Hlth 5, e261-e270, doi:10.1016/S2468-2667(20)30073-6 (2020).
8. Leung, K., Wu, J. T., Liu, D. & Leung, G. M. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. The Lancet 395, 1382-1393, doi:10.1016/S0140-6736(20)30746-7 (2020).
9. Peng, L., Yang, W., Zhang, D., Zhuge, C. & Hong, L. Epidemic analysis of COVID-19 in China by dynamical modeling. medRxiv, doi:10.1101/2020.02.16.20023465 (2020).
10. Read, J. M., Bridgen, J. R. E., Cummings, D. A. T., Ho, A. & Jewell, C. P. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. medRxiv, doi:10.1101/2020.01.23.20018549 (2020).
11. Roda, W. C., Varughese, M. B., Han, D. & Li, M. Y. Why is it difficult to accurately predict the COVID-19 epidemic? Infect Dis Model 5, 271-281, doi:10.1016/j.idm.2020.03.001 (2020).
12. Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet 395, 689-697, doi:10.1016/S0140-6736(20)30260-9 (2020).
13. Brett, T. S. & Rohani, P. Transmission dynamics reveal the impracticality of COVID-19 herd immunity strategies. Proc Natl Acad Sci U S A, 117, 25897-25903, doi:10.1073/pnas.200807117 (2020).
14. Britton, T., Ball, F. & Trapman, P. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. Science 369, 846-849, doi:10.1126/science.abc6810 (2020).
15. Venkatramanan, S. et al. Using data-driven agent-based models for forecasting emerging infectious diseases. Epidemics 22, 43-49, doi:10.1016/j.epidem.2017.02.010 (2018).
16. Beven, K. Environmental modelling: An uncertain future? (CRC press, 2010).
17. Dietze, M. C. Prediction in ecology: a first-principles framework. Ecol Appl 27, 2048-2060, doi:10.1002/eap.1589 (2017).
18. Dietze, M. C. et al. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proc Natl Acad Sci U S A, 115, 1424, doi:10.1073/pnas.1710231115 (2018).
Keenan, T. F., Carbone, M. S., Reichstein, M. & Richardson, A. D. The model-data fusion pitfall: assuming certainty in an uncertain world. *Oecologia* **167**, 587-597, doi:10.1007/s00442-011-2106-x (2011).

Niu, S. *et al.* The role of data assimilation in predictive ecology. *Ecosphere* **5**, 65, doi:10.1890/es13-00273.1 (2014).

White, E. P. *et al.* Developing an automated iterative near-term forecasting system for an ecological study. *Methods Ecol Evol* **10**, 332-344, doi:10.1111/2041-210x.13104 (2019).

White, B. G. *et al.* Short-Term Forecast Validation of Six Models. *Weather Forecast* **14**, 84-108, doi:10.1175/1520-0434(1999)014<0084:STFVOS>2.0.CO;2 (1999).

Calvetti, D., Hoover, A. P., Rose, J. & Somersalo, E. Metapopulation Network Models for Understanding, Predicting, and Managing the Coronavirus Disease COVID-19. *Front Phys* **8**, 261, doi: 10.3389/fphy.2020.00261 (2020).

O’Sullivan, D., Gahegan, M., Exeter, D. J. & Adams, B. Spatially explicit models for exploring COVID-19 lockdown strategies. *T GIS* **24**, 967-1000, doi:10.1111/tgis.12660 (2020).

Bisset, K. R. *et al.* INDEMICS: An Interactive High-Performance Computing Framework for Data-Intensive Epidemic Modeling. *Acm T Model Comput S* **24**, 10.1145/2501602, doi:10.1145/2501602 (2014).

Chao, D. L., Halloran, M. E., Obenchain, V. J. & Longini, I. M., Jr. FlUTE, a Publicly Available Stochastic Inuenza Epidemic Simulation Model. *PLOS Comput Biol* **6**, e1000656, doi:10.1371/journal.pcbi.1000656 (2010).

Marathe, M. V. & Ramakrishnan, N. Recent Advances in Computational Epidemiology. *IEEE Intel Sys* **28**, 96-101, doi:10.1109/MIS.2013.114 (2013).

Dowd, M. A sequential Monte Carlo approach for marine ecological prediction. *Environmetrics* **17**, 435-455, doi:10.1002/env.780 (2006).

Gu, F. On-demand data assimilation of large-scale spatial temporal systems using sequential Monte Carlo methods. *Simul Model Pract Theory* **85**, 1-14, doi:https://doi.org/10.1016/j.simpat.2018.03.007 (2018).

Michael, E. *et al.* Continental-scale, data-driven predictive assessment of eliminating the vector-borne disease, lymphatic filariasis, in sub-Saharan Africa by 2020. *BMC Med* **15**, 176, doi:10.1186/s12916-017-0933-2 (2017).

Poole, D. & Raftery, A. E. Inference for Deterministic Simulation Models: The Bayesian Melding Approach. *J Am Stat Assoc* **95**, 1244-1255 doi:10.1080/01621459.2000.10474324 (2000).

Singh, B. K. & Michael, E. Bayesian calibration of simulation models for supporting management of the elimination of the macroparasitic disease, Lymphatic Filariasis. *Parasites Vectors* **8**, 522, doi:10.1186/s13071-015-1132-7 (2015).

Sisson, S. A., Fan, Y. & Tanaka, M. M. Sequential Monte Carlo without likelihoods. *Proc Natl Acad Sci US A* **104**, 1760, doi:10.1073/pnas.0607208104 (2007).

Spear, R. C., Hubbard, A., Liang, S. & Seto, E. Disease transmission models for public health decision making: toward an approach for designing intervention strategies for Schistosomiasis japonica. *Environ Health Perspect* **110**, 907-915, doi:10.1289/ehp.02110907 (2002).

Taylor, S. D. & White, E. P. Automated data-intensive forecasting of plant phenology throughout the United States. *Ecol Appl* **30**, e02025, doi:10.1002/eap.2025 (2020).

Beaulieu-Jones, B. K. & Greene, C. S. Reproducibility of computational workflows is automated using continuous analysis. *Nat Biotechnol* **35**, 342-346, doi:10.1038/nbt.3780 (2017).

Delgoshaei, P., Austin, M. A. & Pertzborn, A. J. A Semantic Framework for Modeling and Simulation of Cyber-Physical Systems. *Int J Adv Sys Measure* **7**, 223-238, (2014).

Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* **20**, 533-534, doi:10.1016/S1473-3099(20)30120-1 (2020).

Henkel, R., Wolkenhauer, O. & Wullemath, D. Combining computational models, semantic annotations and simulation experiments in a graph database. *Database* **2015** (2015).

Merkel, D. Docker: lightweight linux containers for consistent development and deployment. *Linux J* **2014**, 2 (2014).

Nakamura, K., Higuchi, T. & Hirose, N. Sequential Data Assimilation: Information Fusion of a Numerical Simulation and Large Scale Observation Data. *J UCS* **12**, 608-626 (2006).
42 Stodden, V. & Miguez, S. Best Practices for Computational Science: Software Infrastructure and Environments for Reproducible and Extensible Research. *J Open Res Soft* 2, e21, doi: http://doi.org/10.5334/jors.ay (2014).

43 Unacast. Social distancing Scoreboard. https://www.unacast.com/covid19/social-distancing-scoreboard (2020).

44 Willem, L. et al. SOCRATES: an online tool leveraging a social contact data sharing initiative to assess mitigation strategies for COVID-19. *BMC Res Notes* 13, 293, doi:10.1186/s13104-020-05136-9 (2020).

45 Luo, Y. et al. Ecological forecasting and data assimilation in a data-rich era. *Ecol Appl* 21, 1429-1442, doi:10.1890/09-1275.1 (2011).

46 Hunter, E., Mac Namee, B. & Kelleher, J. An open-data-driven agent-based model to simulate infectious disease outbreaks. *Plos One* 13, e0208775, doi: 10.1371/journal.pone.0208775 (2018).

47 Iboi, E. A., Ngonghala, C. N. & Gumel, A. B. Will an imperfect vaccine curtail the COVID-19 pandemic in the U.S.? *Infect Dis Model* 5, 510-524, doi:10.1016/j.idm.2020.07.006 (2020).

48 Badr, H. S. et al. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect Dis*, doi:10.1016/S1473-3099(20)30553-3 (2020).

49 Contreras, S., Villavicencio, H. A., Medina-Ortiz, D., Biron-Lattes, J. P. & Olivera-Nappa, A. A multi-group SEIRA model for the spread of COVID-19 among heterogeneous populations. *Chaos Sol Fract* 136, 109925, doi:10.1016/j.chaos.2020.109925 (2020).

50 Mossong, J. et al. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLOS Med* 5, e74, doi:10.1371/journal.pmed.0050074 (2008).

51 Chen, R. Sequential Monte Carlo methods and their applications. in *Markov Chain Monte Carlo* Vol. Volume 7 Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore (eds. Kendell, W. S., Liang, F. & Wang, J-S.) 147-182 (World Scientific, 2005).

52 Doucet, A., Godsill, S. & Andrieu, C. On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat Comput* 10, 197-208, doi:10.1023/A:1008935410038 (2000).

53 Fearnhead, P. & Kunsch, H. R. Particle Filters and Data Assimilation. *Annu Rev Stat Appl* 5, 421-449, doi:10.1146/annurev-statistics-031017-100232 (2018).

54 Gu, F., Butt, M., Ai, C., Shen, X. & Xiao, J. Adaptive particle filtering in data assimilation of wildfire spread simulation. in *Proceedings of the Conference on Summer Computer Simulation* 1–10 (Society for Computer Simulation International, Chicago, Illinois, 2015).

55 Polonsky, J. A. et al. Outbreak analytics: a developing data science for informing the response to emerging pathogens. *Philos T R Soc B* 374, 20180276, doi:10.1098/rstb.2018.0276 (2019).

56 Gambhir, M. et al. Geographic and ecologic heterogeneity in elimination thresholds for the major vector-borne helminthic disease, lymphatic filariasis. *BMC biology* 8, 22 (2010).

57 Spear, R. C. & Hubbard, A. Parameter estimation and site-specific calibration of disease transmission models. in *Modelling Parasite Transmission and Control* (eds. Michael, E. & Spear, R.C.) 99-111 (Springer, 2010).

58 Brienen, N. C., Timen, A., Wallinga, J., Van Steenbergen, J. E. & Teunis, P. F. The effect of mask use on the spread of influenza during a pandemic. *Risk Anal* 30, 1210-1218 (2010).

---

**Declarations**

**Acknowledgements**
This work was made possible by an internal grant from the University of South Florida. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. A portion of the model runs was carried out using the MATLAB Parallel Computing Toolbox made available on Compute Clusters of the University of Notre Dame’s Center for Research Computation.

Authors’ contributions

Conceived the paper: EM, KN, MES, CRS; Conducted the analysis: KN, MES, SW, CR, CRS, EM; Wrote the first draft of the manuscript: EM, MES, KN; Data compilation: MES, RED; Contributed to the writing of the manuscript: All authors; Agree with the manuscript’s results and conclusions: All authors.

Competing interests

The authors declare that they have no competing interests.

Tables

Table 1. Total median infected and hospitalized cases predicted at the epidemic peak in each county. Shaded cells denote situations when hospitalization cases are predicted to be below the corresponding county-level care (bed/ICU) capacity.
| Group A | County | Peak Total Infected Cases | Peak Hospitalized Cases | Capacity |
|---------|--------|--------------------------|------------------------|----------|
|         |        | $S_1$                    | $S_2$                  | $S_3$    |
| Alachua | 82696  | 67899                    | 26572                  | 7161     |
| Baker   | 6856   | 5121                     | 1769                   | 696      |
| Bay     | 51066  | 39280                    | 14864                  | 399       |
| Bradford | 9479   | 9000                     | 9191                   | 325      |
| Calhoun | 2567   | 1830                     | 496                    | 89       |
| Clay    | 78051  | 71980                    | 43312                  | 18053    |
| Columbia | 20045  | 17378                    | 9488                   | 3406     |
| DeSoto  | 6013   | 4831                     | 1592                   | 348      |
| Flagler | 23992  | 19667                    | 7942                   | 1571     |
| Franklin | 2872   | 2234                     | 1755                   | 211      |
| Gadsden | 11163  | 9126                     | 3676                   | 1268     |
| Gilchrist | 3386  | 2630                     | 871                    | 197      |
| Gulf    | 2057   | 1483                     | 375                    | 61       |
| Hamilton | 4298   | 3550                     | 1479                   | 577      |
| Hardee | 6867   | 6877                     | 3033                   | 1274     |
| Highlands | 30539  | 24480                    | 8743                   | 2160     |
| Holmes | 5527   | 4437                     | 1649                   | 551      |
| Jackson | 14501  | 13119                    | 7631                   | 2737     |
| Jefferson | 3279  | 2609                     | 968                    | 454      |
| Lafayette | 1459  | 1014                     | 275                    | 50       |
| Lake    | 9037   | 76121                    | 30544                  | 7937     |
| Levy    | 9600   | 8061                     | 3621                   | 1002     |
| Liberty | 2782   | 2502                     | 1522                   | 809      |
| Madison | 4990   | 4206                     | 1878                   | 540      |
| Marion | 383145 | 63118                   | 19313                  | 3779     |
| Monroe | 14350  | 10199                    | 2954                   | 633      |
| Manatee | 20632  | 16795                    | 6453                   | 1412     |
| Okaloosa | 58620  | 50121                    | 26223                  | 8021     |
| Putnam | 24005  | 21207                    | 10829                  | 3847     |
| Santa Rosa | 34202 | 25531                    | 7368                   | 1225     |
| Sumter | 34849  | 28042                    | 10592                  | 2447     |
| Suwannee | 9988   | 7334                     | 2051                   | 430      |
| Taylor | 4521   | 3477                     | 1170                   | 398      |
| Union | 4701   | 3926                     | 1886                   | 561      |
| Wakulla | 9333   | 7997                     | 3061                   | 975      |
| Walton | 24418  | 21068                    | 10072                  | 3227     |
| Washington | 4756  | 3660                     | 1135                   | 207      |
|         |        | $S_4$                    | $S_5$                  |          |
| Brevard | 216546 | 194227                   | 101280                 | 41631    |
| Charlotte | 50932  | 44307                    | 15841                  | 4175     |
| Citrus | 30776  | 24863                    | 7968                   | 1431     |
| Collier | 81605  | 61625                    | 19574                  | 3953     |
| Dixie | 3875   | 3071                     | 1189                   | 235      |
| Escambia | 76781  | 59268                    | 19828                  | 4584     |
| Glades | 5967   | 56121                    | 20544                  | 2348     |
| Hendry | 6591   | 5053                     | 1469                   | 268      |
| Hernando | 48155  | 38208                    | 13378                  | 3046     |
| Hillsborough | 397091 | 296947                   | 89333                  | 21603    |
| Indian River | 45001 | 37314                    | 12981                  | 3739     |
| Leon | 96968   | 85935                    | 43604                  | 13896    |
| Martin | 39199   | 31207                    | 10958                  | 2816     |
| Okaloosa | 17530  | 15780                    | 7888                   | 4710     |
| Pasco | 211838 | 195573                   | 110066                 | 45306    |
| Pinellas | 249782 | 184207                   | 54312                  | 11694    |
| St. Johns | 90187  | 79325                    | 42884                  | 15740    |
| St. Lucie | 77214  | 61019                    | 18514                  | 3588     |
| Sarasota | 142268 | 115675                   | 44632                  | 13217    |
| Broward | 587916 | 488812                   | 204061                 | 53764    |
| Dade | 263303  | 208691                   | 71687                  | 22480    |
| Lee | 160900  | 113444                   | 30393                  | 4979     |
| Manatee | 86346   | 60866                    | 17509                  | 2743     |
| Miami-Dade | 742598 | 577962                   | 182730                 | 47067    |
| Orange | 825542  | 557470                   | 190520                 | 163852   |
| Osceola | 87801   | 64189                    | 18324                  | 3667     |
| Palm Beach | 460637 | 379619                   | 141320                 | 39597    |
| Polk | 232456  | 203981                   | 99111                  | 31266    |
| Seminole | 112562 | 85203                    | 25402                  | 5711     |
| Volusia | 140718  | 97188                    | 23409                  | 4716     |
|         |        | $S_6$                    | $S_7$                  |          |
|         |        | $S_8$                    | $S_9$                  |          |

| Group B | County | Peak Total Infected Cases | Peak Hospitalized Cases | Capacity |
|---------|--------|--------------------------|------------------------|----------|
|         |        | $S_1$                    | $S_2$                  | $S_3$    |
|         |        | $S_4$                    | $S_5$                  |          |

| Group C | County | Peak Total Infected Cases | Peak Hospitalized Cases | Capacity |
|---------|--------|--------------------------|------------------------|----------|
|         |        | $S_1$                    | $S_2$                  | $S_3$    |
|         |        | $S_4$                    | $S_5$                  |          |

Total: 6200885 503473 2008993 662339 174513 186269 146465 78075 26527 4398 28122
Figure 1

Model fits compared to confirmed cumulative case data in 67 Florida counties. Gray curves represent county-specific model predictions and red points represent confirmed case data obtained from Johns Hopkins University. Fitting was started after at least 10 confirmed cases were reported.
Figure 2

Epidemic forecasts for individual counties in the scenario where all interventions are lifted following the initial lockdown (S1). Each curve represents the median prediction for a given county. The results are stratified by initial incidence growth rate in each county (Group a < 0.05, 0.05 < Group b < 0.15, Group c > 0.15). The intervention scenario (also see Figure S2) represents a full release of lockdown and social distancing from October 1st. The y-axis is shown in log-scale to better visualize the difference between the size of the first and second epidemic waves. The range of epidemic peaks and epidemic ending dates for each group is shown by red and blue dotted vertical lines, respectively. The 2nd wave peaks of infected cases occurs between September 27th – November 2nd for Group a, October 3rd – November 7th for group b, and October 3rd - November 2nd for Group c. The epidemic ending dates occur between December 2nd – January 21st, 2021 for Group a, December 2nd – January 25th, 2021 for group b, and January 16th, 2021 - March 2nd, 2021 for Group c, respectively.
Figure 3

Epidemic forecasts for individual counties under four different social intervention scenarios. Note the differences in y-axis values when comparing scenarios. Each curve represents the median prediction for a given county. The results are stratified by initial incidence growth rates observed in each county (Group b < 0.05, 0.05 < Group b < 0.15, Group c > 0.15). The intervention scenarios are as follows (also see Figure S2): S2: maintain movement estimate and social distancing measures over 2 weeks from October 1st to October 14th; S3: maintain movement estimate and social distancing measures over 8 weeks from October 1st to November 30th; S4: maintain movement estimate and social distancing measures through end December 2020 + begin low intensity contact tracing and quarantine efforts on October 1st through March 2021; S5: maintain movement estimate and social distancing measures through end December 2020 + begin high intensity contact tracing and quarantine efforts on October 1st through March 2021. The red shading in each plot indicates the duration of movement and social distancing in each scenario, while the blue shading shows quarantining measures without movement restrictions or social distancing. The range of 2nd wave infection peaks for each county group is shown by dotted vertical lines. For scenario 2, the peak of infected cases occurs between October 29th - November 16th for Group a, October 29th – November 14th for group b, and November 2nd - November 13th for Group c. For scenario 3, the peak of infected cases occurs between December 17th - January 7th 2021 for Group a, December 16th – January 5th 2021 for group b, and December 17th - January 3rd 2021 for Group c. Epidemic ending dates for scenario 2 occur between January 2nd – February 17th 2021 for Group a, January 6th – February 21th 2021 for group b, and February 11th - February 27th, 2021 for Group c, respectively. For scenario 3, the corresponding end dates were between February 20th – March 2nd 2021 for Group a, and between February 22th to beyond March 2nd for Groups b and c.
Forecasts of corresponding hospitalized cases for individual counties under the four different social intervention scenarios. Hospitalized cases include both hospital and ICU model compartments. Note the differences in y-axis values when comparing scenarios. Each curve represents the median prediction for a given county. The results are stratified by initial county-level incidence growth rate (Group a < 0.05, 0.05 < Group b < 0.15, Group c > 0.15). The intervention scenarios are as follows (also see Figure S2): S2: maintain movement estimate and social distancing measures over 2 weeks from October 1st to October 14th; S3: maintain movement estimate and social distancing measures over 8 weeks from October 1st to November 30th; S4: maintain movement estimate and social distancing measures through end December 2020 + begin low intensity contact tracing and quarantine efforts on October 1st through March 2021; S5: maintain movement estimate and social distancing measures through end December 2020 + begin high intensity contact tracing and quarantine efforts on October 1st through March 2021. The red shading in each plot indicates the duration of movement and social distancing in each scenario, while the blue shading shows quarantining measures without movement or social distancing. The range of peak hospitalized cases for each group is shown by dotted vertical lines. For scenario 2, the peak of hospitalized cases occurs between November 9th - November 24th for Group a, November 7th – November 23rd for group b, and November 12th - November 22nd for Group c. For scenario 3, the peak of hospitalized cases occurs between December 25th - January 15th 2021 for Group a, December 27nd – January 13th 2021 for group b, and January 2nd, 2021 - January 18th 2021 for Group c.
Figure 5

Epidemic forecasts for individual counties in the scenario where all interventions (movement restriction, social distancing, and contact tracing) are held through to March 2021 starting from October 1st 2020. Each curve represents the median prediction for a given county. The results are stratified by initial incidence growth rate (Group a < 0.05, 0.05 < Group b < 0.15, Group c > 0.15). The y-axis is shown in log-scale to better visualize the difference between the size of the first and second epidemic waves. Red vertical lines indicate the ranges of peak infected cases, while blue vertical lines indicate the typical ending times of the epidemic. The ending dates for the epidemic are predicted to occur between November 1st - January 26th 2021 for Group a, November 14th – February 25th 2021 for Group b, and December 20th – February 25th 2021 for Group c.
Figure 6

Proportion recovereds predicted over time for Scenario 1 (full release of interventions on October 1st) and Scenario 6 (lockdown, social distancing, and quarantining through to March 2021). Each curve represents the median prediction for a given county. The results are stratified by initial incidence growth rate in each county (Group a < 0.05, 0.05 < Group b < 0.15, Group c > 0.15).
Figure 7
Cumulative number of deaths predicted over time for Scenario 1 (full release of interventions on October 1st) and Scenario 6 (lockdown, social distancing, and quarantining through March 2021). Each curve represents the median prediction for a given county. The results are stratified by initial incidence growth rate in each county (Group a < 0.05, 0.05 < Group b < 0.15, Group c > 0.15).

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- CovidPaperSI10.19.2020.docx