European Journal of Information Systems

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjis20

Chasing John Snow: data analytics in the COVID-19 era

Jesse Pietz , Scott McCoy & Joseph H. Wilck

To cite this article: Jesse Pietz , Scott McCoy & Joseph H. Wilck (2020): Chasing John Snow: data analytics in the COVID-19 era, European Journal of Information Systems, DOI: 10.1080/0960085X.2020.1793698

To link to this article: https://doi.org/10.1080/0960085X.2020.1793698

Published online: 20 Jul 2020.

Submit your article to this journal

Article views: 285

View related articles

View Crossmark data

Citing articles: 2 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tjis20

EUROPEAN JOURNAL OF INFORMATION SYSTEMS
https://doi.org/10.1080/0960085X.2020.1793698

EMPIRICAL RESEARCH

Chasing John Snow: data analytics in the COVID-19 era

Jesse Pietz[a], Scott McCoy[b] and Joseph H. Wilck[b]

[a]Department of Management, US Air Force Academy, Colorado Springs, CO, USA; [b]Raymond A. Mason School of Business, William and Mary, Williamsburg, VA, USA

ABSTRACT
During the first half of 2020, the lives of people around the world abruptly changed due to COVID-19. Data visualisations and models related to the spread of the disease became ubiquitous. In this paper, we survey 25 different data analytics dashboards, highlight the modelling approach taken by each, and develop a multi-attribute utility theory model to assess their effectiveness in communicating key features that explain the spread of infectious disease. We show that the dashboards that feature dimensions that span the categories associated with compartmental epidemiology models tend to be relatively robust data visualisations, and we highlight that information systems need to be improved to include data on actions to reduce the spread of the disease. We analyse the actions taken by countries around the world and show that when governments employ strict measures early, particularly those that enforce social distancing and include widespread testing and comprehensive contact tracing, they are more likely to experience better outcomes. Recommendations for how countries should respond in future pandemics are detailed.

ARTICLE HISTORY
Received 6 June 2020
Accepted 6 July 2020

SPECIAL ISSUE EDITORS
Pär Ågerfalk, Kieran Conboy and Michael Myers

KEYWORDS
COVID-19; data visualisation; multi-attribute utility; pandemic modelling; contact tracing; information systems

1. Introduction

In 1854 John Snow meticulously catalogued the outbreak of cholera in London to identify its mode of communication (Snow, 1855). Lacking modern information systems, Snow used public records and interviewed countless residents to identify the source of the outbreak in a sort of premodern contact tracing study. Snow's renowned dot map visualisation of cholera cases in and around the water pump near the intersection of Cambridge and Broad Streets in London's Soho district (Snow, 1855, p. 45) serves as a powerful example of how data analytics and data visualisation can be used to understand the spread of an infectious disease.

Cameron and Jones (1983, p. 393) argue that Snow's genius was in his ability to thoroughly study the "mechanisms and processes of every aspect he had chosen to study" and to present his findings internally to the medical community and externally for public consumption. In his studies, he used data in order to confirm his germ theory of disease. Snow's analysis was detailed, textured, and multidimensional. In order to understand how cholera communicated from person to person, he considered interactions between people and their environment within communities.

Data visualisations and analyses used to understand complex problems should be textured and multidimensional. The complement of Tufte's (2001) famous principle of visual integrity, which requires that the number of information-carrying dimensions in a data visualisation be no larger than the number of dimensions in the data, applies here. In order to understand a problem in all its complexity, a data visualisation should present as many of the problem's dimensions as possible. Indeed, Tufte (2001, p. 40) goes on to praise Charles Minard's classic visualisation of the fate of Napoleon's army in 1812 as perhaps "the best statistical graphic ever drawn" because of the skill with which Minard was able to display six dimensions of data on a two-dimensional visualisation.

In today's era of the COVID-19 pandemic, data visualisations and models that deal with the spread of infectious disease are ubiquitous. And this is a good thing. Lee and Jung (2019) include accurate information and information sharing as among the most important factors influencing the response to infectious disease. Central to this idea is that only accurate information be shared. Laato et al. (2020) and Zarocostas (2020) highlight the problems that misinformation and information overload pose to public health. The response to infectious disease improves as officials and the public understand the factors that influence its spread with greater clarity. In January 2020, just days after the first confirmed case of COVID-19 in the United States, Johns Hopkins University (2020) shared its COVID-19 Tracking Map in order to systematically track and analyse the outbreak as it happened. We believe this to be the first such product released in the United States. Shortly thereafter several other academic, research,

CONTACT Jesse Pietz jesse.pietz@afacademy.af.edu
© Operational Research Society 2019.

government, news, and independent organisations released COVID-19 data visualisation dashboards and models, each taking different approaches and with different features. COVID-19 data visualisation dashboards, hereafter referred to as *dashboards*, use data analytics to present a variety of information related to the spread of the disease and the policies that public health officials are taking to combat it. COVID-19 models, hereafter referred to as *models*, use a variety of data to predict how the disease is likely to spread under a variety of scenarios, including public health policy. These models are often supported by dashboards to help readers understand the model's data and predictions.

Unfortunately, the dimensions considered in many of these dashboards and models are insufficient to understand the problem of reducing the spread of infectious disease. The United States Centers for Disease Control and Prevention (2020a) references a collection of models and explains that their aim is to predict the number of deaths using different types of data, methods, and estimate the impacts of interventions. However, in order to understand how interventions may affect death rates, dashboards and models must consider myriad dimensions that relate these factors. This is a difficult task, and one that some studies do better than others.

In this paper, we survey 25 publicly available dashboards and models, chosen as a representative sampling of products across various analysis sectors (academia, research organisations, government, news outlets, and independents) and those listed on the aforementioned United States Centers for Disease Control and Prevention (US CDC) website. Assessing the value of these technologies in supporting the fight against COVID-19 is among the many open research questions regarding information systems in the age of pandemics (Ågerfalk et al., 2020). We discuss the modelling approaches taken in these studies and present a multi-attribute utility theory (MAUT) framework to assess their effectiveness in helping to understand the status of the COVID-19 pandemic and how interventions may affect death rates. Multi-attribute decision modelling has been studied widely as an approach to assessing decision support systems and information technology investments. Ahituv (1980) presents a MAUT approach to assess the value of a reporting system that uses three major categories of attributes that represent its capability. Forgionne (1999), Liebowitz (1986), Phillips-Wren et al. (2004), (2009), and Phillips-Wren et al. (2009) examine how an Analytical Hierarchy Process (AHP), a type of multi-attribute decision modelling that relies on pairwise comparisons, may be used to assess the value of decision support systems. Angelou and Economides (2008) also uses AHP, but the aim of their study is to prioritise information and communication technology investments. These studies all consider general-purpose information systems. We constrain our study to the subject of reporting on and analysing the spread of infectious disease. This focus allows us to tailor our MAUT framework to consider the factors that are known to be important for understanding how disease spreads.

There is surprisingly little analysis in the literature that relates the intervention measures that governments enact (e.g., stay home orders, school closures, contact tracing) to the problem of reducing the spread of or deaths caused by COVID-19. Oxford University has developed the Oxford COVID-19 Government Response Tracker (OxCGRT), which collects data from 73 countries based on a host of measures taken, such as school closings, workplace closings, shelter-in-place orders, income support, contact tracing, and a host of other measures (Our World in Data, 2020). The OxCGRT also includes composite assessment of these measures called stringency. Our work builds upon these data with a statistical analysis that relates the timing and magnitude of stringency measures to eventual COVID-19 cases and deaths.

In the next section, we highlight the different modelling approaches taken by the analyses surveyed. In the following section, we present our MAUT framework and assess how each dashboard presents relevant factors associated with the spread of COVID-19. We then discuss the global pandemic response and consider different strategies taken by countries around the world based on the stringency measures they have taken. In the penultimate section, we discuss varying responses taken by governments and highlight the importance of testing and contact tracing. We conclude with a summary of our findings and recommendations for governments.

2. COVID-19 modelling approaches

Our work starts by reviewing many of the different modelling approaches used to provide both descriptive analytics on the current cases and deaths, but also those used to predict the impact of the pandemic. The descriptive analytics models share current statistics on cases, deaths, recoveries, etc. The predictive models share current statistics and provide forecasts for deaths and cases in the future. The dashboards included in the list of data visualisations (Table A1) include models that use descriptive analytics and predictive modelling.

We observe that the authors of the dashboards considered regularly update their data and models. Our discussion and analyses are based on observations completed on May 29, 2020. A popular approach to modelling the spread of an infectious disease is to use a compartmental epidemiology model to understand the dynamic relationship between compartments that

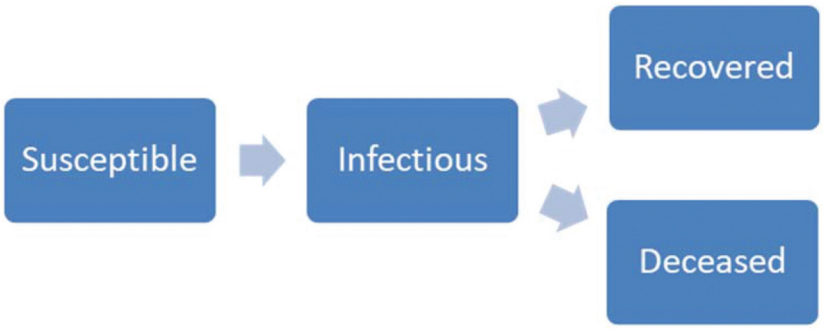

Figure 1. Flowchart of a basic SIRD Model.

describe an individual's disposition with respect to the disease (Hethcote, 2000; Ji & Jiang, 2014). The simplest version of this model includes susceptible (S), infected (I) and recovered (R) compartments. For diseases that cause fatalities, it is common to include the deceased (D) compartment (see Figure 1). The resulting SIRD model illuminates how certain key factors affect the spread of the virus. In addition to population data for each compartment (S, I, R & D), the SIRD model depends on the rate at which infected people interact with other people (α), the recovery rate (β), and the death rate attributed to the disease (δ) (Osemwinyen & Diakhaby, 2015).

The model is analysed by solving the following differential Equations (1–4).

$$\frac{dS}{dt} = -\alpha IS \tag{1}$$

$$\frac{dI}{dt} = \alpha IS - \beta I - \delta I \tag{2}$$

$$\frac{dR}{dt} = \beta I \tag{3}$$

$$\frac{dD}{dt} = \delta I \tag{4}$$

SIRD models can be solved using ordinary differential equations, curve-fitting measures using generalised linear models (GLMs), stochastic optimisation approaches, simulation, or some combination of those. This is the most common predictive modelling approach taken in the analyses we surveyed (e.g., CP, CU, ICL, IHME, JHU, MIT, MOBS, UCLA, and UG). Approaches vary in how rates (α, β, and δ) are estimated and by including other compartments (e.g., it is common to include another compartment to model the number of people that have been exposed to the virus).

Two alternative approaches to predictive modelling that are not based on compartmental epidemiology models include the LANL and UT predictive models. The LANL predictive model uses a growth model for infections, and then probabilistic forecasting for deaths. The UT predictive model uses curve-fitting approaches based on the underlying data. Another modelling approach is to use descriptive analytics and statistics to visualise different aspects of COVID-19. These models do not make predictions on the future, but instead, focus on visualising the current status.

All the dashboards use government data for confirmed deaths and infections (i.e., from the US CDC, state and local health departments, European Centre for Disease Prevention and Control, World Health Organization); most of these data are summarised and available on the JHU website. However, some of the predictive models include significant supplemental data, for example, the UT model includes mobile phone data from SafeGraph (2020), the MOBS predictive model includes data from the International Air Transport Association (2020), the MIT predictive model includes data from over 160 clinical trials, and the ICL model includes stages for hospitalisations (i.e., critical care versus general hospitalisation) and accounts for ages (i.e., using 5 year age bins).

The US CDC began reporting an ensemble model (i.e., an average of several predictive models) the last week of April 2020 and have included more predictive models in the ensemble with each update. In terms of model performance at predicting the total number of COVID-19 deaths in the USA, the CP predictive model has performed the best in May 2020. FiveThirtyEight (2020) also provides a visualisation of these various predictive models.

As of this writing, the European Centre for Disease Prevention and Control (ECDC) does not provide an ensemble model nor do they provide forecasts for deaths. However, several European models do provide forecasts, including the European-based Imperial College London (ICL) and University of Geneva (UG) models. Both predictive models use data from JHU and the ECDC and are based on SIRD methodology.

The COVID-19 dashboards, both descriptive analytics models and predictive models are data-driven and not process-driven. National and local policies are

changing from business-as-usual to stay-at-home, and back again. COVID-19 may also adapt and evolve with changes in weather and temperature. The dashboards are not designed to update disease transmission rates automatically based on policies and weather changes, and at best will include assumptions about changes to transmission rates and at worst will have a time lag as those changes are reflected in the data. The other major issue with these models is that they are focused primarily on reporting and projecting near-term deaths, hospital utilisation rates, infection rates, etc. However, they are not focused on identifying actions to reduce deaths, reduce infections, and eliminate the spread of the disease.

3. Assessment of COVID-19 dashboards

We begin by observing that dashboards may be created for any number of reasons. For the purposes of this assessment, we assume that all dashboards we consider are designed to help people to understand the status of COVID-19 pandemic and/or to take actions to reduce infections and deaths caused by the virus. We denote the former as Objective A and the latter as Objective B. Dashboards that were apparently produced for some other objective were removed from consideration and are not included in those assessed in this paper.

Truly understanding the status of COVID-19 at any point in time (Objective A) requires data on population and each of the SIRD model compartments. Knowing how to affect Objective B requires more nuanced data and analytics. Intervention policies (e.g., quarantines, travel restrictions, contact tracing) are designed to reduce the infection rate α, and these policies vary temporally and geographically. Medical capabilities (e.g., hospital beds, ventilators) can affect the recovery rate β and death rate δ. Demographic factors (e.g., age, ethnicity, pre-existing health conditions) can affect the infection rate α, the recovery rate β, and the death rate δ depending upon the context in which data are recorded.

We use each of these SIRD model terms to establish our MAUT framework. MAUT modelling has been used in the management sciences since the 1960s to aid decision-making when alternatives must be evaluated using more than one criterion. Huber (1974) surveys the various MAUT approaches. A data visualisation (and model) that presents data in the following SIRD model categories (C) is assessed to be strong: population, infections, recoveries, deaths, factors affecting infection rate, factors affecting recovery rate, and factors affecting death rate. For each category (c), we use an exponentially decaying utility function (u) to assess the strength of a data visualisation as a function of the number of data dimensions (x) considered (5).

$$u_c \equiv u_c(x) = 1 - e^{-xK} \tag{5}$$

We choose a risk avoidance policy and set shape parameter $K = 2$, taking the view that the most important consideration is for a dashboard to account for each category in some way. Moreover, we assume that the value of a dimension in each category diminishes as more are added. SIRD models are dynamic. All components must be represented in order to understand the spread of infectious disease. Depth in any one category (with many dimensions) has little value if an entire category is not represented. We considered excursions with varying values of K and present them as sensitivity analysis. Weights (w) are used to sum the resulting utility of each category. The utility of each dashboard j is determined by a weighted sum of utility scores (6).

$$U(j) = \sum\nolimits_{c \in C} w_c u_c \tag{6}$$

Each dashboard was assessed by counting dimensions (see Table A1) to determine values x and applying those data to Equations (5–6). Figure 2 depicts the resulting assessment assuming that each category c is equally important (i.e., weights w are all equal). Assessment score ranges 0–1. Dashboards with higher scores help the reader to understand more of the factors associated with the spread of COVID-19.

We observe that dashboards that are underpinned by predictive models tend to provide the reader with a richer understanding of the spread of COVID-19. This is because they generally account for all the SIRD model factors in order to make predictions. For example, JHU and MIT are assessed with high scores because the data visualisations presented include dimensions that span all seven SIRD model categories. Data are reported on a per capita basis to account for population. The number of known cases is combined with testing data to give the reader an understanding of infections. Recovery and death data are shown to account for those status groups. Case data are reported over time and geography to give the reader a sense for the infection rate. Hospital bed and ventilator availability data speak to recovery rate. Death rate estimates are provided. And MIT presents an infection and death risk calculator based on demographic factors.

While many dashboards that are underpinned by predictive models generated high assessment scores, some did not. The UT dashboard only displays predictive model output: deaths over time and geography. The reader is not presented with any of the other SIRD model categories. The CU dashboard presents projected cases and hospital bed usage over time and

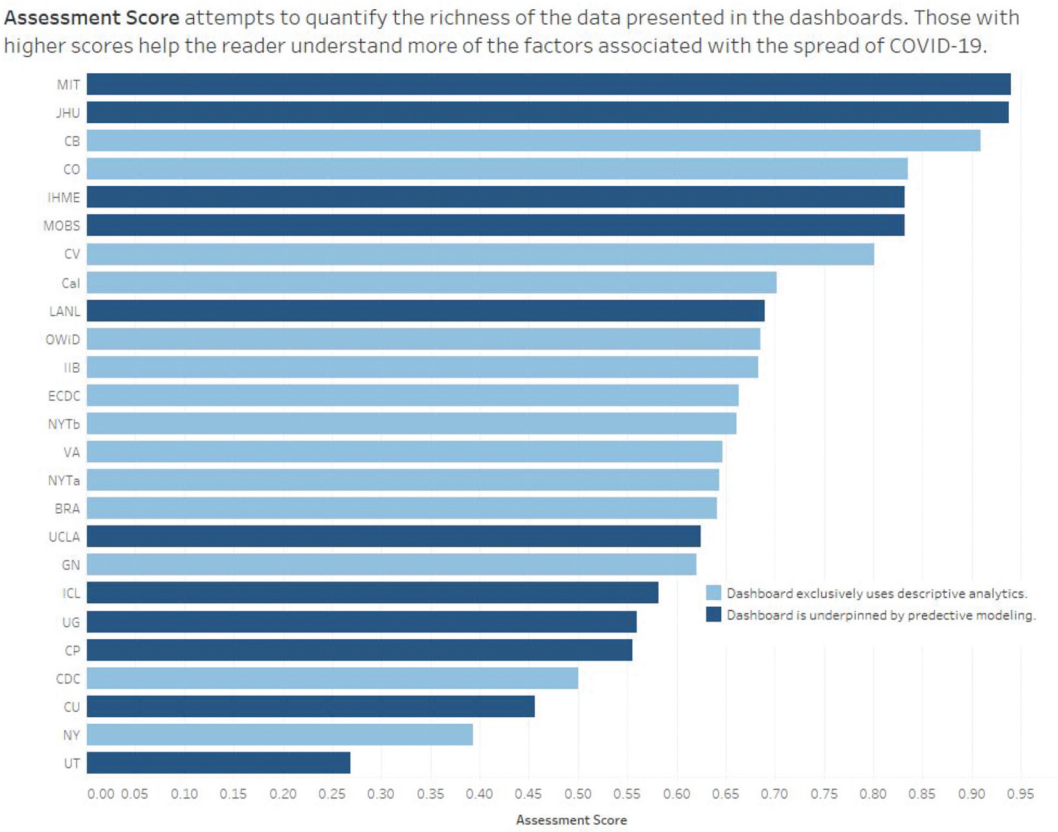

Figure 2. Assessment score assuming K = 2 and equally weighted SIRD model categories.

geography. Population, death, and recovery data are not shown.

Assessment scores for dashboards that exclusively use descriptive analytics with no predictive modelling varied substantially. Those with high scores provided data about most of the seven SIRD model categories. For example, CO includes population, cases, deaths, and testing data along with several demographic dimensions and data that speaks to hospital capabilities. The only category missing is recovery data. The CB and CV dashboards simply present active case, death, and recovery data over time and geography. CB also includes population data and CV does not. Those with low assessment scores are lacking data in multiple SIRD model categories.

Some dashboards perform better for Objective A than Objective B (and vice versa). We can see this when analysing the sensitivity of assessment score as we change the weighting w in (6). The first four SIRD model categories (population, infections, recoveries, and deaths) are associated with Objective A, while the last three categories (factors affecting infection rate, factors affecting recovery rate and factors affecting death rate) are associated with Objective B. Applying weight sensitivity parameter q in [0,1] to (6) we can see how assessment score changes.

$$U(j) = \sum_{c=1}^{4} q w_c u_c + \sum_{c=5}^{7} (1-q) w_c u_c \quad (7)$$

In Figure 3 we show how the assessment score for each dashboard varies as we adjust the weighting. Interestingly, the dashboards with the highest nominal assessment, MIT and JHU, vary the least (left column in Figure 3). They are always among the highest assessed dashboards (because the data presented spans all seven categories as previously discussed). As the weighing shifts to favour Objective A, dashboards that show data on SIRD model compartments see improved assessment scores. For example, we see UCLA and GN go from below average overall assessments to being assessed as among the best dashboards.

We submit that it is better for a dashboard to account for each SIRD model category and extra depth in any category is not as important. In other words, the shape parameter K in (1) should justifiably be greater than 1. But how much greater? If it is too high, say $K = 5$, then the utility of each category approaches a binary response and a score of approximately 1 is awarded for having at least one data dimension in the category. If it is too low, say $K = 1/5$, then the utility of each category increases with each dimension added. In Figure 4 we see how assessment score changes as we vary the shape parameter K between 1/5 and 5.

Overall, dashboards with data dimensions that span more of the SIRD model categories (e.g., MIT, JHU, CB, CO) provide their readers with a better understanding of the status of COVID-19 pandemic and

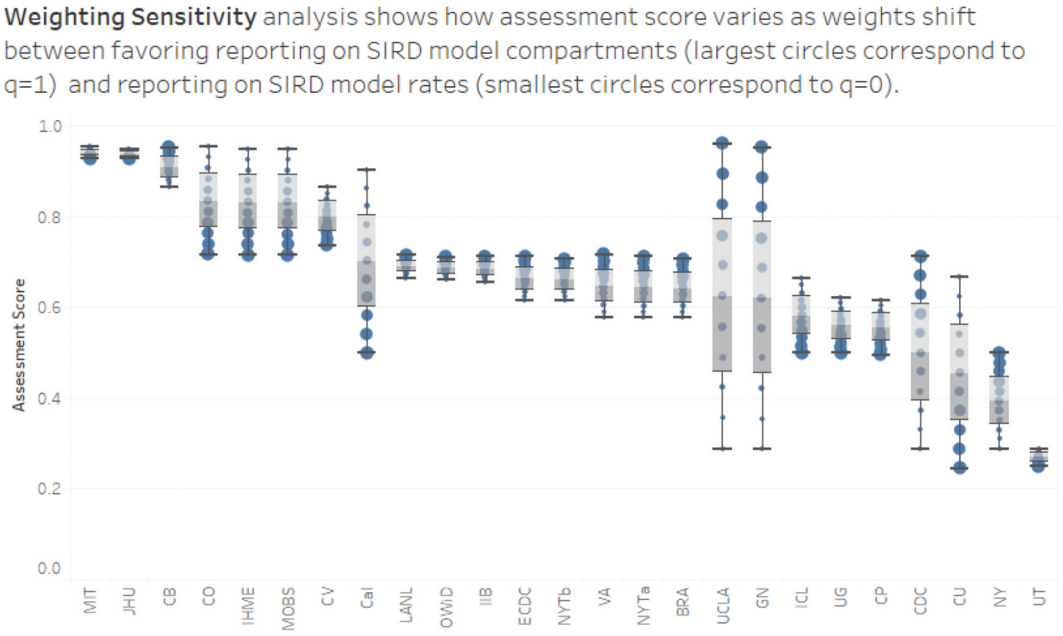

Figure 3. Assessment score as weighting is varied between favouring Objective A and favouring Objective B.

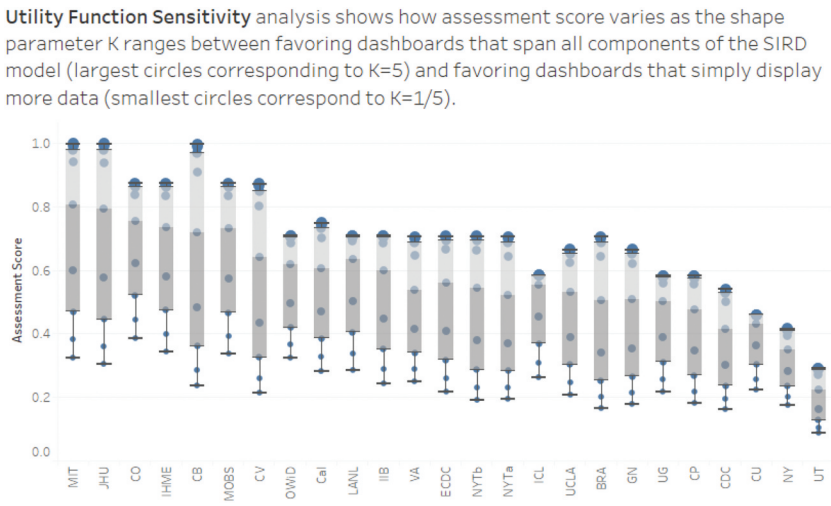

Figure 4. Sensitivity analysis depicting how assessment score changes as we vary the shape parameter *K* between 1/5 and 5.

how it is spreading. However, the dimensions presented by the dashboards are varied, offer little consistency and subject society to information overload, at a time when people need timely and accurate information. This is particularly important for 18 of the dashboards surveyed, which are produced or endorsed by government agencies. Moreover, none of the dashboards surveyed are assessed with perfect scores. Only 6 of the 25 dashboards surveyed present testing data, which is crucial to understanding the number of infections as reported cases are known to underestimate the actual number of infections. Only five dashboards present recovery data. Last and perhaps most importantly, while several of them present data that speaks to intervention policies (e.g., stay home orders, school closures, contact tracing), they all lack clarity on *how* each intervention has affected or is expected to affect infection rates. In the next section, we present an alternative to improve how data analytics dashboards convey the relationship between intervention policies and the spread of the disease in the context of the global response to COVID-19.

4. Global response

The COVID-19 pandemic started with a cluster of pneumonia cases in Wuhan, Hubei Province, China first reported by the Wuhan Municipal Health Commission on December 31, 2019 (WHO, 2020). The World Health Organization (WHO) issued infection and control guidance on January 10, 2020 to protect health workers. On January 13, 2020, the first confirmed case outside of China was reported in Thailand. COVID-19 has now been reported in at

least 188 countries (Johns Hopkins University, 2020) with only a handful of isolated countries reporting no cases.

Some countries have had very good success at preventing large outbreaks of the COVID-19 virus, while others have been hit especially hard. As discussed in the previous section, all the dashboards that we surveyed lack clarity as to how intervention measures affect the spread of the virus. In order to assess the impact measures taken and their impact on COVID-19 deaths, we combined JHU data, which included longitudinal data on cases and deaths, and a recently developed model of stringency built by Oxford University (OWiD). Oxford University has developed the Oxford COVID-19 Government Response Tracker which collects data from 73 countries based on 17 different types of response measures, such as school closings, workplace closings, shelter-in-place orders, income support, contact tracing, and a host of other measures. For a complete review of the factors considered and the scores of each country, please see Our World in Data (2020). Ideally, the higher the stringency measures taken, the lower the death toll of the virus.

The tracker records government responses worldwide and aggregates the scores into a "Stringency Index" which allows one to understand the strict measures governments have taken and see the impact on rates of infection. These scores range from 0 to 100 with a score of 100 having the highest stringent policies. We have used this model to see the corresponding impact on COVID-19-related deaths. Although the model predicted similar results with infections, we felt deaths were a more appropriate outcome measure given that an accurate measure of infections has been elusive to most countries because of a lack of testing. In order to control for differences in the timing when each country took stringent measures, we shifted the timeline for each country to coincide with the date its 10th case was reported. All else being equal we would expect cases and deaths for each country to follow a similar exponential growth pattern from that point. We assume that differences in cases and deaths from one country to the next can be attributed to differences in both the timing and nature of stringent policies enacted. Figure 5 shows the scatterplot of early stringency, defined as cumulative stringency up until the day of the 10th reported case, and the deaths per 100 K people (on a log scale) by country as of May 24 (the last date that we collected data). We chose to use the 10th case because this was a constraint chosen by some of the models (e.g., IHME, CP, and ICL) to determine when to include a locality for sample size and data reliability purposes. The trendline is negatively sloped with statistically significant p-value and explains approximately 18% of the variance in COVID-19 deaths. Similar is true when we consider cases per 100 K people (significant p-value with 23% variance explained), but we chose not to present that chart for brevity.

The individual response measures that comprise stringency vary in importance depending on timing and geographic scope. We observe that cancelling public events, restrictions on gatherings and international travel controls are the most important of the 17 measures when it comes to early stringency. The numbers of cases and deaths (per capita) experienced by countries that enacted these three measures early (before their 10th reported case) were approximately 80% lower than those of the countries that did not (see Table A2). Additionally, the importance of cancelling public events varies based on geographic scope. Event cancellations nationwide are more impactful in reducing deaths and infections than when events are only cancelled in targeted regions. The other 14 response measures can be thought of as important pieces of

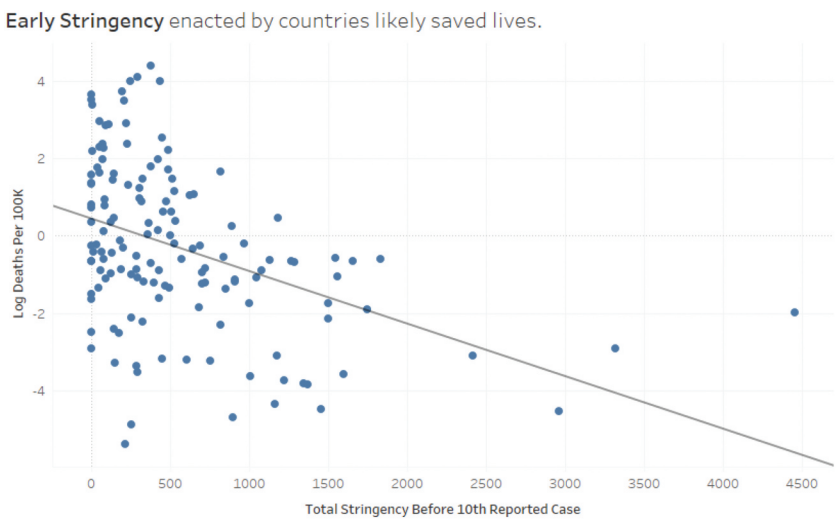

Figure 5. Scatterplot of total stringency before the 10th reported case of COVID-19 and the deaths per 100 K people (on a log scale) by country.

aggregate early stringency response, but not necessarily important as early measures individually. When we consider the period after a country reported its 10th COVID-19 case, we see changes in the importance of individual response measures. While preventative social distancing measures like cancelling public events, restrictions on gatherings and international travel controls continue to be important, other response measures like testing and contact tracing are also important in combating the growing infection. The numbers of deaths (per capita) experienced by countries that had widespread testing and comprehensive contact tracing in place for most of the time through the 60-day period following the date of their 10th reported case were approximately 50% and 70% lower (respectively) than those of the countries that did not (see Table A3).

To better understand the measures taken and the corresponding impact on COVID-19 cases and deaths, we have chosen to compare a relatively negative outcome country with a relatively positive outcome country in the same geographic region. We discuss when the first stringency measures were taken and when the stringency score exceeded 50 (on a 0 to 100 scale), along with details on what measures were taken.

In Europe, we consider Czech Republic and Sweden. Czech Republic acted faster and more stringently than Sweden and experienced fewer cases and deaths as a result[1] (see Figure 6). Sweden had their first case on February 1 but made no attempt to mitigate the cases. The government resisted telling its citizens how to behave, their schools and restaurants remained open, and the initial results seemed similar to neighbouring countries. While no restrictions were in place when the country had its 10th case, on March 9 with 203 cases the country took their first stringency measures (11.11) by having a public information campaign, a symptomatic testing policy, and comprehensive contact tracing. Their stringency measures never exceeded 50 (40.74) and only included recommended school and workplace closings and restrictions on large gatherings. On May 24, their case count was 326.17 per 100 K and their death rate was 39.23 per 100 K with a stringency score of 40.74. Czech Republic, on the other hand, took relatively quick action on January 24 (42 days before their 10th case) when they took their first stringency measures (5.56) by urging caution about the virus. By the time they had their 10th case on March 6, their Early Stringency[2] totalled 669.52 and included an international travel ban, a coordinated public information campaign, and comprehensive contact tracing. On March 12, their stringency measures exceeded 50 (50) with a case count of 94 and extended the previous measures to include requiring all schools to close, cancelling public events, restricted gatherings larger than 10, and provided symptomatic testing. On May 24, their case count was 83.63 per 100 K and their death rate was 2.95 per 100 K with a stringency score of 54.63.

In Africa, we focus on Egypt and Zimbabwe. While both countries eventually enacted a similar degree of stringent measures, Zimbabwe acted much faster than Egypt in taking them and experienced fewer cases and deaths as a result (see Figure 7). By the time Egypt reached its 10th case, no action had been taken to curb the spread of the virus. Egypt took its first measures (11.11) on March 15 when its COVID-19 case count was 93. These actions included only closing schools. Their stringency measures exceeded a score of 50 (51.85) on March 21 when their case count was 256. These actions extended school closings to also include some workplace closings, cancelling of public events, and border closures. On May 24, their case count was 16.88 per 100 K and their death rate was 0.75 per 100 K with a stringency score of 84.26. In contrast,

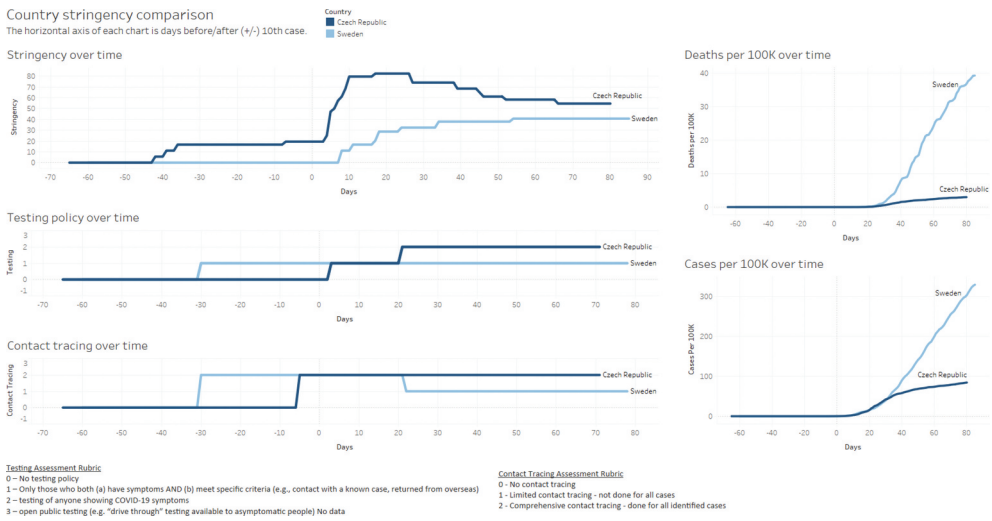

Figure 6. Stringency, cases and deaths comparison for Czech Republic and Sweden. Area under the stringency, testing and contact tracing curves to the left of zero indicate early action.

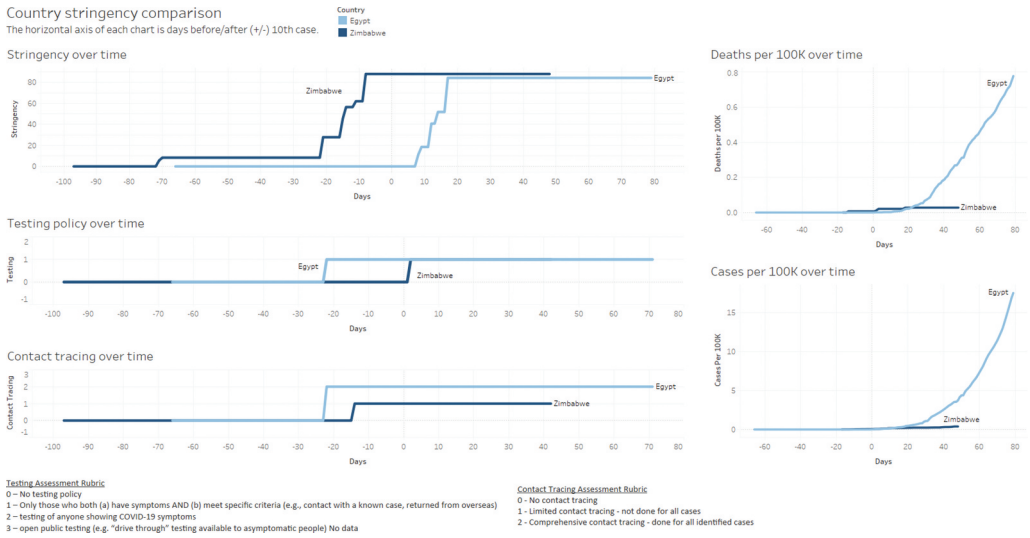

Figure 7. Stringency, cases and deaths comparison for Egypt and Zimbabwe. Area under the stringency, testing and contact tracing curves to the left of zero indicate early action.

Zimbabwe took their first measures (5.56) on January 27 (72 days before their 10th case) when their case count was zero by urging the public to be cautious about the virus. Their Early Stringency was among the highest totals in the world at 1,685.02. On March 24 with only 2 cases, their stringency score had exceeded 50 (56.48) and included closing all schools, cancelling public events, restricted gatherings of large groups (more than 10), recommending people stay at home, recommending people restrict their internal country movements, closing their border, and enacting limited contact tracing. On May 24, their case count was 0.39 per 100 K and their death rate was 0.03 per 100 K with a stringency score of 87.96. In comparing these two countries in Africa, the stark differences are the timing and level of stringent measures taken. Egypt acted after their first cases were reported and initially only closed schools. Six days later, when their cases had tripled, they took more action. Zimbabwe, on the other hand, took more stringent measures initially and when their case count was zero and then exceeded 50 when they had only two cases. While both countries have relatively low death rates compared to the many countries on May 24, it is important to note that Zimbabwe has been able to limit their case count by taking action very early, while Egypt's case count is 43 times higher and likely due to a slow start in their mitigation strategies.

In Asia, we focus on Iran and Vietnam. Vietnam acted faster and more stringently than Iran and experienced fewer cases and deaths as a result (see Figure 8). Iran has been one of the hardest hit countries in Asia. Their first reported case and death happened on the same day, February 19. The very next day (only 2 days before their 10th case) Iran took its first stringent measure (2.78) by recommending cancelling public

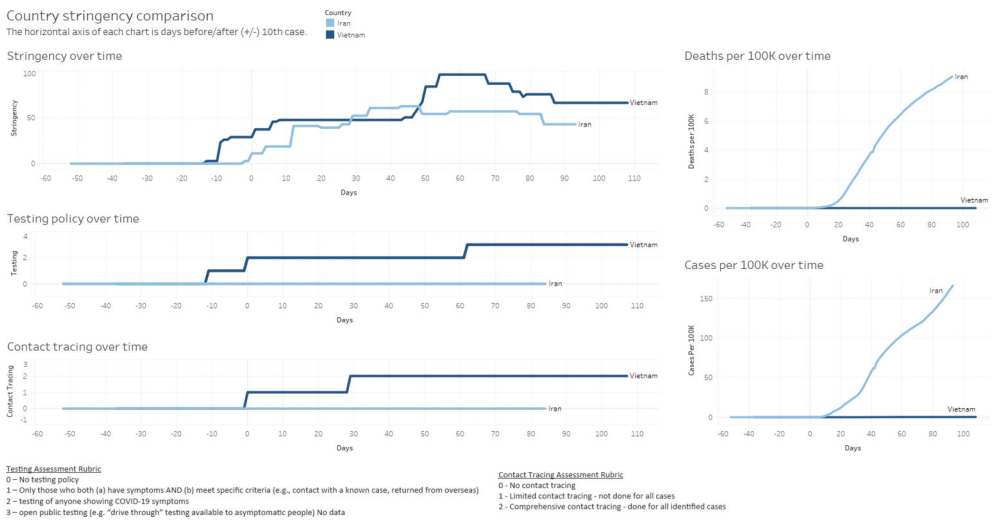

Figure 8. Stringency, cases and deaths comparison for Iran and Vietnam. Area under the stringency, testing and contact tracing curves to the left of zero indicate early action.

events. On February 22, Iran had its 10th case and an Early Stringency total of 5.56, which involved adding the outright cancelling of public events (previously it had just been recommended). Exactly 1 month later, on March 22 Iran's stringent measure surpassed 50 (51.85) for the first time when cases reached 133,521. Their measures extended to closing all schools, closing all businesses that were non-essential, recommending people stay home, and restricting internal movements. However, they have had no testing policy and no contact tracing throughout the pandemic. On May 24, their case count was 163.23 per 100 K and their death rate was 8.99 per 100 K with a stringency score of 42.59. In contrast, Vietnam was able to control the spread of the virus by controlling its population. Vietnam has dealt with pandemics before and knew their health system could be overrun if the virus was to spread. They took their first action (2.78) on January 25 (11 days before their 10th case) by screening international travellers when their case count was 2. Their Early Stringency totalled 258.33 and on March 22 their stringency measures reached 50 when their case count was 95. By this time, they had a coordinated public information campaign, a symptomatic testing policy in place, comprehensive contact tracing, borders were closed, international movements were restricted, all public events had been cancelled, and all schools were closed. The real impact of their measures came from their early actions which included sending everyone entering Vietnam to quarantine centres where everyone was tested, whether they showed symptoms or not. While the entire country was not on lockdown, thousands of people in communities where a positive case was detected were sealed off until 2 weeks after no confirmed cases were found. The result is 0.34 cases per 100 K and not a single death has been reported (Johns Hopkins University, 2020).

In North America, our comparison is with the USA and Cuba, two countries with many stark differences. Cuba acted faster and more stringently than the United States and experienced fewer cases and deaths as a result (see Figure 9). The United States took their first stringent measures (5.56) on February 2 (only 1 day before their 10th case) with a limited travel ban from China (excluding US Citizens and permanent residents), including a quarantine for those who were in the Hubei province the previous 14 days and limited testing and contact tracing. Their Early Stringency totalled 11.12. Approximately 6 weeks later, on March 16, the USA's stringent measures exceeded 50 (52.31) when cases totalled approximately 6,000. These additional measures included closing schools, cancelling public events, restricting gatherings over 100, stay at home requirements with exceptions, recommended restricting internal movements, ban on international travel, a coordinated public information campaign, generally available testing, and limited contact tracing. On May 24, their case count was 496.70 per 100 K and their death rate was 29.72 per 100 K with a stringency score of 72.69. In contrast, Cuba took its first steps (11.11) on January 28 (51 days before their 10th case) when they had no reported cases. These steps included international travel screening, a coordinated public information campaign, and testing of anyone symptomatic. By March 24, their Early Stringency totalled 655.56. Just five days later, on March 24, with 40 cases, Cuba's stringency measures exceeded 50 (66.67) and extended the previous measures to include closing all schools, cancelling public events, closing public transportation, restricting internal movement, banning international travel, and limited contact tracing. On May 24, their case count was 17.03 per 100 K and their death rate was 0.72 per 100 K with a stringency score of 100, which includes open public testing and comprehensive contact tracing.

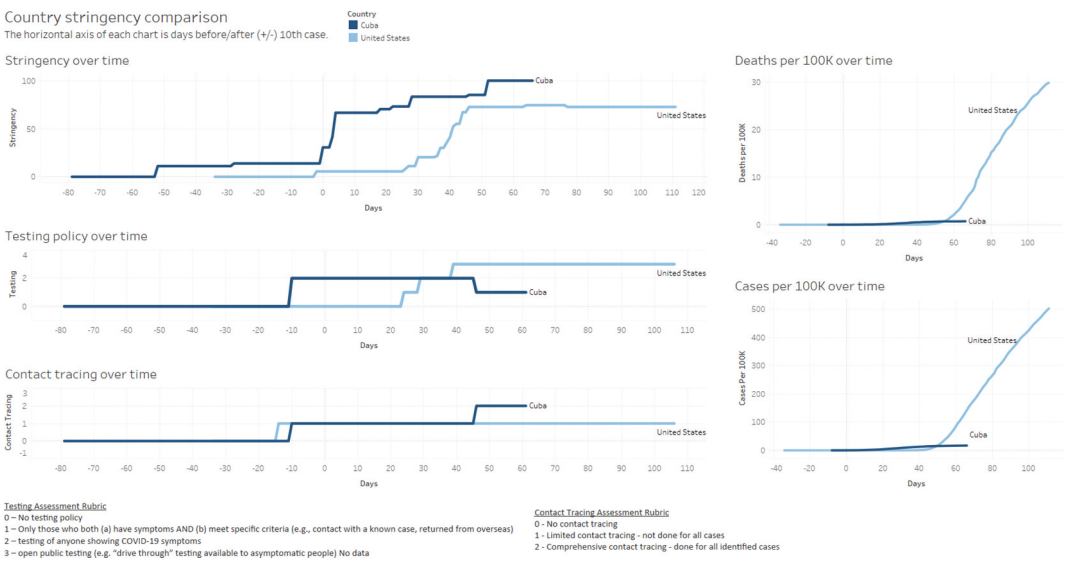

Figure 9. Stringency, cases and deaths comparison for Cuba and United States. Area under the stringency, testing and contact tracing curves to the left of zero indicate early action.

In South America, we compare Ecuador and Paraguay. While both countries eventually enacted a similar degree of stringent measures, Paraguay acted faster than Ecuador in taking them and experienced fewer cases and deaths as a result (see Figure 10). Ecuador's first stringent measures (5.56) started on January 26 (39 days before their 10th reported case) when, with zero cases, public officials urged caution. Their Early Stringency totalled 225.15, which is a high figure for a country that experienced a relatively large number of cases and deaths. The first time their stringency measures exceeded 50 (52.78) was on March 15 when they had 28 cases. The measures were extended to include closing all schools, cancelling public events, closing public transportation, and banning international travel. It is important to note that their stringency measure jumped to 93.52 just two days later but has never included more than limited contact tracing. On May 24, their case count was 212.23 per 100 K and their death rate was 18.12 per 100 K with a stringency score of 86.11. Paraguay announced their first stringent measures (5.56) on January 23 (55 days before their 10th reported case) with public officials urging caution about the virus. On March 13, their stringency measures (50.93) exceeded 50 for the first time and extended measures to include closing all schools, closing all but essential businesses, cancelling public events, restricting gatherings over 10, recommending people stay home, symptomatic testing, and comprehensive contact tracing (which came into effect March 7). At the time, Paraguay had just six cases. By the time they had their 10th case, their Early Stringency totalled 752.05 with measures extending to also include restricting internal movements in the country and an international travel ban. On May 24, their case count was 12.22 per 100 K and their death rate was 0.16 per 100 K with a stringency score of 94.44. It is important to note that on March 24, these two countries had the same stringency scores of 93.52 but drastically different numbers of cases (Ecuador had 981 cases while Paraguay had 27). Also, while Paraguay had a comprehensive contact tracing system in place before their first case, Ecuador has only had limited contact tracing capability.

5. Discussion of government responses

We now consider the varied government responses presented in the previous section and highlight the importance of early stringency, testing, and contact tracing. Table A4 provides a summary of government responses, including the initial stringent measure and case count, Early Stringency, the stringent measure when first exceeding 50 (scale is 0 to 100) and corresponding case count, whether the country implemented widespread testing and contact tracing, and cases and deaths as of May 24.

Based on the country comparisons presented in the previous section and the summary details in Table A4, it is clear that those countries that fared well had widespread testing and comprehensive contact tracing in addition to taking more aggressive and early stringency measures overall. In fact, the only country comparison where the positive outcome did not involve widespread testing and comprehensive contact tracing was Zimbabwe. However, by the time, this country had its 10th case, its stringency score was an astounding 87.96. Although not directly compared in the preceding section because this was the first country affected and some have questioned the data reported, according to Brennan (2020), China has reported positive results because they were the first country to implement a quarantine. In addition, they locked down cities and controlled the

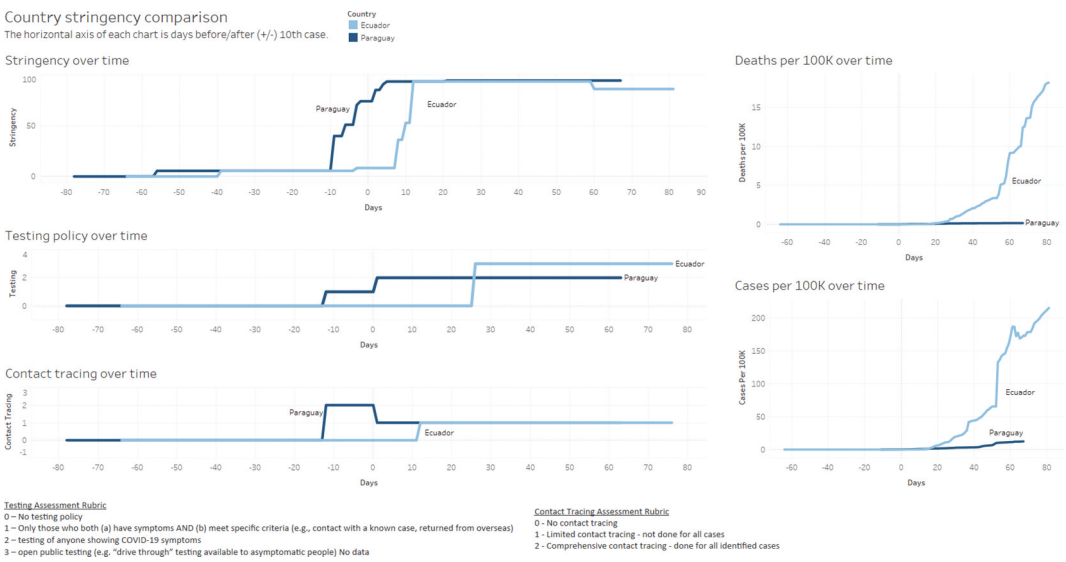

Figure 10. Stringency, cases and deaths comparison for Ecuador and Paraguay. Area under the stringency, testing and contact tracing curves to the left of zero indicate early action.

movement of millions of its citizens using surveillance systems. Anyone suspected of being infected was physically removed from their homes and taken to government quarantine locations. Countries like France deployed extra police to enforce their lockdown. Some states in India ordered to self-isolate were stamped on the left hand to help enforcement officers identify who should not be outside their home (Brennan, 2020).

It appears that when early and stringent measures were not both possible, aggressive, draconian measures by some governments resulted in similar outcomes. For example, as of May 24, China's case count was 6.04 per 100 K and their death rate was 0.33 per 100 K with a stringency score of 81.94. South Korea's technology-driven contact tracing was the real game changer for that country. Their contact tracing operation integrated GPS data, credit card data, surveillance footage, and information from almost 30 different data sources. This allowed the government to perform real-time analysis to identify people who had been in contact with a COVID-19 infected person (Kim et al., 2020). This is in stark contrast to the USA, where tests are not widely available and data used for contact tracing is based on voluntary disclosure, which is known to be problematic (Shanahan, 2020), not real-time objective data collection and analysis. As of May 24, South Korea's case count was 21.68 per 100 K and their death rate was 0.52 per 100 K with a stringency score of 39.81, whereas the USA had a case count of 496.70 per 100 K and their death rate was 29.71 deaths per 100 K with a stringency score of 72.69.

6. Conclusions and recommendations

What have we learned? First, disease spread information and modelling are imperfect, we must improve data collection and sharing, and we must develop a clearer understanding of how intervention (stringent) policies affect disease transmission rates. Second, governments that took swift action experienced the best results. Third, widespread testing and comprehensive contact tracing are important factors in keeping infection and death rates low. The following discussion summarises each of our findings in turn.

This paper studies 25 different data analytics dashboards and the models that underpin them. We build upon previous work in the area of using multi-attribute utility theory to assess general-purpose decision support systems (Forgionne, 1999; Liebowitz, 1986; Phillips-Wren et al., 2004; Phillips-Wren et al., 2009), and develop an infectious disease-focused assessment framework that highlights that as dashboards include more compartmental epidemiology model dimensions, they are more effective in providing society with a clear understanding of the spread of COVID-19. This clear understanding is known to be an important factor in the response to infectious disease (Lee & Jung, 2019). Observing that existing dashboards are generally lacking in their ability to illustrate how intervention policies affect the spread of COVID-19, we built a more comprehensive model to analyse the global response to the disease and demonstrate that early action results in lower infection and death rates. These findings together suggest a clear pathway forward for governments as they refine their information systems, data analytics dashboards and disease spread models. Governments should develop and refine information systems to systematically track, analyse, and disseminate data about the status and spread of infectious disease. At a minimum, these data should include all compartmental epidemiology model dimensions: population, infections (and testing), deaths, recoveries, infection rates, death rates, and recovery rates. As existing information systems are refined and new ones are developed, governments need to incorporate more action-oriented data that communicate the timing and stringency of response measures taken in order to give citizens a clear understanding of why compliance to governmental intervention policies is needed. The combination of disease status and response measure data needs to be presented in a single dashboard (or similar data report) in a consistent manner in order to improve public understanding. Looking forward, our proposed information system development pathway will be particularly important for countries like the USA where intervention policies are presented only as guidelines, and each of the 50 states and all the many cities and communities within those states are mostly allowed to respond how they see fit.

Using Oxford University's Stringency Index to assess the intervention policies enacted by governments around the world, we show that swift action brought about the best results. Our research demonstrates this by empirically analysing COVID-19 and government response data, whereas current thinking about early response is largely based on disease spread estimates arising from compartmental epidemiology modelling (Hethcote, 2000; Ji & Jiang, 2014; Osemwinyen & Diakhaby, 2015). Though our finding has implications for information system development (e.g., tracking and assessing governmental intervention policies), it extends beyond that to the broader conversation of how governments should respond to curtail the spread of infectious disease. Swift action is necessary. And all actions are not created equal. Our analysis suggests that cancelling public events (nationwide, not only in targeted regions), restrictions on gatherings and international travel controls are the most important early measures to take. This was the case of Zimbabwe, which took all these early measures, had the highest Early Stringency score of the countries we examined, and experienced among the fewest infections and deaths (per capita) in the world despite not having widespread testing and comprehensive contact tracing.

Finally, our analysis shows the importance of widespread testing and comprehensive contact tracing. Countries that had widespread testing and comprehensive contact tracing in place experienced death rates that were approximately 50% and 70% lower (respectively) than those of the countries that did not. Moreover, contact tracing programmes should not be based on voluntary disclosure. Information systems and technologies that enable real-time, objective data collection are needed. Czech Republic, Cuba, Paraguay, and Vietnam all took early action with widespread testing and comprehensive contact tracing. As a result, their infection and death rates were remarkably low. Additionally, China, Taiwan, and South Korea took aggressive action with both the control of their populations as well as widespread testing and comprehensive contact tracing. China and South Korea monitored their citizens with massive surveillance capabilities to focus on infections and quarantine. All three of these countries recorded less than 1 death per 100 K people (Johns Hopkins University, 2020). In contrast, countries that waited to respond and did not have robust testing and contact tracing in place (Sweden and USA, for example) have experienced high infection and death rates. As we look towards the future and how to continue to respond to COVID-19, additional waves of the virus, and potentially new pandemics, how can technology be used to help protect the world's population? Surveillance by countries like China and South Korea helped greatly. However, their actions were authoritative in nature and not something that other countries might allow. One possible aid would be contact tracing by smartphone. As was seen in the USA, conducting limited contact tracing and relying on voluntary disclosures of who infected patients had been in contact with was not a sound strategy. Regardless of how we gather the data (by widespread surveillance or comprehensive contact tracing by smartphone apps), we must improve data collection and sharing in order to improve the models, assessments, and effects of intervention policies designed to curtail the spread of infectious disease. A summary of our recommendations for governments is highlighted in Table 1.

John Snow's analysis clearly showed people *how* to address the cause of the outbreak. Our problem is difficult because of population growth and expanded travel since the 1850s. The COVID-19 pandemic has affected every country in the world as a result. As economic strain compels governments to be less stringent with the measures they take (and relax those they have already taken), we need to have clarity on which measures work and which do not. As detailed in the preceding discussion, governments must develop and refine information systems to systematically track, analyse, and disseminate data about the status and spread of infectious disease. This information needs to be shared widely and includes action-oriented data so that citizens around the world understand the importance of intervention policy compliance. Further, the intervention policies governments take should be swift. Our analysis shows that those countries which moved quickly had the best outcomes. Finally, any mitigation measures must include widespread testing and comprehensive contact tracing. The rate at which our data systems can track and analyse the spread of infectious disease needs to be at least on par with rate at which population grows and people travel. If not, we will continue to be chasing John Snow.

Table 1. Summary of recommendations for governments.

Recommendation 1: Develop and refine information systems to systematically track, analyse and disseminate disease status data that spans all compartmental epidemiology model dimensions: population, infections (and testing), deaths, recoveries, infection rates, deaths rates and recovery rates.
Recommendation 2: Incorporate more action-oriented data that communicate the timing and stringency of response measures taken in order to give citizens a clear understanding of why compliance to governmental intervention policies is needed. Consistently present these data, along with disease status data, in a single dashboard (or similar data report).
Recommendation 3: Take swift action with intervention policies, particularly cancelling public events nationwide, restricting public gatherings and limiting international travel, when a resurgence or another outbreak is imminent.
Recommendation 4: Conduct widespread testing and comprehensive contact tracing, and develop information systems and technologies to enable real-time, objective data collection and analysis.

Notes

1. This statement and similar statements later in this paper are based on the statistical analysis presented earlier in this section. There may be (and likely are) other unobserved factors that explain the remainder of the difference in outcomes between the countries we considered.
2. Early Stringency is defined as the sum of daily stringency scores until the day of the 10th reported case of COVID-19.
3. A country's adoption of a measure is assessed as high when it has the equivalent of the most stringent level of a measure in place for most time through the 60-day period following the date of its 10th reported case.
4. Widespread testing is defined as testing anyone with COVID-19 symptoms (or higher, including open public testing) and Comprehensive contact tracing is defined as completing contact tracing for all identified cases at any time during the pandemic.

Disclosure statement

No potential conflict of interest was reported by the authors. The views expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the US Government. Distribution A: Approved for Public Release, Distribution Unlimited. USAFA-DF-2020-247.

ORCID

Jesse Pietz http://orcid.org/0000-0003-2517-7577

References

Ågerfalk, P. J., Conboy, K., & Myers, M. D. (2020). Information systems in the age of pandemics: COVID-19 and beyond. *European Journal of Information Systems. Advance Online Publication*. https://doi.org/10.1080/0960085X.2020.1771968

Ahituv, N. (1980). A systematic approach toward assessing the value of an information system. *MIS Quarterly*, *4*(4), 61–75. https://doi.org/10.2307/248961

Angelou, G. N., & Economides, A. A. (2008). A decision analysis framework for prioritizing a portfolio of ICT infrastructure projects. *IEEE Transactions on Engineering Management*, *55*(3), 479–495. https://doi.org/10.1109/TEM.2008.922649

Bertsimas, D., Boussioux, L., Wright, R. C., Delarue, A., Kitane, D. L., Lukin, G., Li, M. L., Mingardi, L., Orfanoudaki, A., Papalexopoulos, T., Paskov, I., Pauphilet, J., Lami, O. S., Stellato, B., Tazi, H., Carballo, K. V., Wiberg, H., Fazel-Zarandi, M., Jacuillat, A., & Zeng, C. (2020, May 29). *COVID Analytics*. MIT Operations Research Center. Retrieved May 19, 2020, from https://www.covidanalytics.io/

Branas, C. C., Rundle, A., Pei, S., Tang, W., Carr, B. G., Sims, S., Zebrowski, A., Doorley, R., Schluger, N., Quinn, J. W., & Shaman, J. (2020, May 29). *COVID-19 Projection in the US*. Retrieved May 29, 2020, from https://columbia.maps.arcgis.com/apps/webappviewer/index.html?id=ade6ba85450c4325a12a5b9c09ba796c

Brennan, D. (2020, March 18). Fines, jail time and sackings: What happens when people break coronavirus quarantines around the world. Newsweek. https://www.newsweek.com/fines-jail-time-sackings-what-happens-people-break-coronavirus-quarantines-around-world-1492947

California Health and Human Services. (2020, May 29). *California health and human services open data portal*. Retrieved May 29, 2020, from. https://data.chhs.ca.gov/

Cameron, D., & Jones, I. G. (1983). John Snow, the Broad Street pump and modern epidemiology. *International Journal of Epidemiology*, *12*(4), 393–396. https://doi.org/10.1093/ije/12.4.393

Centers for Disease Control and Prevention. (2020a, May 28). Cases in the U.S. Retrieved May 29, 2020, from. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html

Centers for Disease Control and Prevention. (2020b, May 29). Cases in the U.S. Retrieved May 29, 2020, from. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html

Colorado Department of Public Health & Environment. (2020, May 29). Case data. Retrieved May 19, 29, 2020, from. https://covid19.colorado.gov/data/case-data

CoronaBoard. (2020, May 29). *COVID-19 dashboard*. Retrieved May 29, 2020, from. https://coronaboard.com/

European Centre for Disease Prevention and Control. (2020, May 29). *COVID-19*. Retrieved May 29, 2020, from. https://www.ecdc.europa.eu/en/covid-19-pandemic

FiveThirtyEight. (2020, May 30). *Where the latest COVID-19 models think we're headed — And why they disagree*. Retrieved May 30, 2020, from. https://projects.fivethirtyeight.com/covid-forecasts/

Flahault, A., Manetti, E., Simonson, T., Lee, G., Obozinski, G., Krymova, E., Haro, B. B., Thanou, D., Sun, T., & Choirat, C. (2020, May 29). *COVID-19 daily epidemic forecasting*. Retrieved May 29, 2020, from. https://renkulab.shinyapps.io/COVID-19-Epidemic-Forecasting/

Forgionne, G. A. (1999). An AHP model of DSS effectiveness. *European Journal of Information Systems*, *8*(2), 95–106. https://doi.org/10.1057/palgrave.ejis.3000322

Google News. (2020, May 29). *Coronavirus (COVID-19)*. Retrieved May 29, 2020, from. https://news.google.com/covid19/map

Gu, Y. (2020, May 29). *COVID-19 projections using machine learning*. Retrieved May 29, 2020, from. https://covid19-projections.com/

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, *42*(4), 599–653. https://doi.org/10.1137/S0036144500371907

Huber, G. P. (1974). Multi-attribute utility models: A review of field and field-like studies. *Management Science*, *20*(10), 1393–1402. https://doi.org/10.1287/mnsc.20.10.1393

Imperial College London. (2020, May 29). *Short-term forecasts of COVID-19 deaths in multiple countries*. Retrieved May 29, 2020, from. https://mrc-ide.github.io/covid19-short-term-forecasts/index.html

Information is Beautiful. (2020, May 29). *COVID-19 #Coronavirus infographic datapack*. Retrieved May 29, 2020, from. https://informationisbeautiful.net/visualizations/covid-19-coronavirus-infographic-datapack/

The Institute for Health Metrics and Evaluation. (2020, May 29). *IHME measuring what matters*. University of Washington. Retrieved May 29, 2020, from. http://www.healthdata.org/

International Air Transport Association. (2020). *IATA*. Retrieved May 30, 2020, from. https://www.iata.org/

Ji, C., & Jiang, D. (2014). Threshold behaviour of a stochastic SIR model. *Applied Mathematical Modelling*, *38*(21–22), 5067–5079. https://doi.org/10.1016/j.apm.2014.03.037

Johns Hopkins University. (2020, May 29). *COVID-19 dashboard by the center for systems science and engineering (CSSE) at Johns Hopkins University (JHU)*. Retrieved May 29, 2020, from. https://coronavirus.jhu.edu/map.html

Kim, S. R., Kung, T., & Abdelmalek, M. (2020, May 1). *Trust, testing and tracing: How South Korea succeeded where the US stumbled in coronavirus response*. ABC News. Retrieved May 29, 2020, from. https://abcnews.go.com/Health/trust-testing-tracing-south-korea-succeeded-us-stumbled/story?id=70433504

Laato, S., Islam, A. N., Islam, M. N., & Whelan, E. (2020). What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? European Journal of Information Systems. *Advance Online Publication*. https://doi.org/10.1080/0960085X.2020.1770632

Lee, K. M., & Jung, K. (2019). Factors influencing the response to infectious diseases: Focusing on the case of SARS and MERS in South Korea. *International Journal of Environmental Research and Public Health*, *16*(8), 1432. https://doi.org/10.3390/ijerph16081432

Liebowitz, J. (1986). Useful approach for evaluating expert systems. *Expert Systems*, *3*(2), 86–96. https://doi.org/10.1111/j.1468-0394.1986.tb00198.x

Los Alamos National Laboratory. (2020, May 29). *COVID-19 Confirmed and Forecasted Case Data*. Retrieved May 29, 2020, from. https://covid-19.bsvgateway.org/

Mamoon, N., & Rasskin, G. (2020, May 29). *COVID-19*. Retrieved May 29, 2020, from. https://www.covidvisualizer.com/

Ministério da Saúde do Brasil. (2020, May 29). *Coronavírus Brasil*. Retrieved May 29, 2020, from. https://covid.saude.gov.br/

New York State Department of Health. (2020, May 29). *New York State Department of Health COVID-19 tracker*. Retrieved May 29, 2020, from. https://covid19tracker.

health.ny.gov/views/NYS-COVID19-Tracker/NYSDOHCOVID-19Tracker-Map

The New York Times. (2020a, May 29). *Coronavirus Map: Tracking the Global Outbreak*. Retrieved May 29, 2020, from. https://www.nytimes.com/interactive/2020/world/coronavirus-maps.html

The New York Times. (2020b, May 29). *New York City Coronavirus Map and Case Count*. Retrieved May 29, 2020, from https://www.nytimes.com/interactive/2020/nyregion/new-york-city-coronavirus-cases.html

Osemwinyen, A., & Diakhaby, A. (2015). Mathematical modelling of the transmission dynamics of ebola virus. *Applied and Computational Mathematics*, *4*(4), 313–320. https://doi.org/10.11648/j.acm.20150404.19

Our World in Data. (2020, May 29). *COVID-19: Government response stringency index*. Global Change Data Lab. Retrieved May 29, 2020, from. https://ourworldindata.org/grapher/covid-stringency-index

Phillips-Wren, G., Hahn, E., & Forgionne, G. A. (2004). A multiple-criteria framework for evaluation of decision support systems. *Omega*, *32*(4), 323–332. https://doi.org/10.1016/j.omega.2004.01.003

Phillips-Wren, G., Mora, M., Forgionne, G. A., & Gupta, J. N. (2009). An integrative evaluation framework for intelligent decision support systems. *European Journal of Operational Research*, *195*(3), 642–652. https://doi.org/10.1016/j.ejor.2007.11.001

SafeGraph. (2020). *The source of truth for POI data & business listings*. Retrieved May 30, 2020, from. https://www.safegraph.com/

Shanahan, E. (2020, July 1). Party guests wouldn't talk after 9 tested positive. Then subpoenas came. The New York Times. Retrieved July 3, 2020, from. https://www.nytimes.com/2020/07/01/nyregion/rockland-coronavirus-party.html

Snow, J. (1855). *On the mode of communication of cholera*. John Churchill.

Statistical Machine Learning Lab at UCLA. (2020, May 29). *COVID-19 Cases in the United States*. Retrieved May 29, 2020, from. https://covid19.uclaml.org/

Tufte, E. R. (2001). *The visual display of quantitative information (Vol. 2)*. Graphics press.

The University of Texas COVID-19 Modeling Consortium. (2020, May 29). *COVID-19 mortality projections for US States and metropolitan areas*. Retrieved May 29, 2020, from. https://covid-19.tacc.utexas.edu/projections/

Vespignani, A., Chinazzi, M., Davis, J. T., Mu, K., Pastore Y Piontti, A., Samay, N., Xiong, X., Halloran, M. E., Longini, I. M., Jr., Dean, N. E., Viboud, C., Sun, K., Litvinova, M., Gioannini, C., Rossi, L., & Ajelli, M. (2020, May 24). *COVID modeling Unites States*. GLEAM Project. Retrieved May 29, 2020, from. https://covid19.gleamproject.org/

Virginia Department of Health. (2020, May 29). *COVID-19 in Virginia*. Retrieved May 29, 2020, from. https://www.vdh.virginia.gov/coronavirus/

WHO. (2020, May 24). This statement is updated on an ongoing basis, in response to evolving events and common media queries. Retrieved May 24, 2020, from. https://www.who.int/news-room/detail/27-04-2020-who-timeline—covid-19

Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, *395*(10225), 676. https://doi.org/10.1016/S0140-6736(20)30461-X

Appendix

Table A1. Summary of different data visualisation dashboards and models.

Label	Reference	Sector	Dimensions	Model
BRA	Ministério da Saúde do Brasil (2020)	Government	geography, time, cases, deaths, per capita data	descriptive
Cal	California Health and Human Services (2020)	Government	geography, time, demographics, medical capabilities, cases, tests	descriptive
CB	CoronaBoard (2020)	Independent	geography, time, cases, deaths, recoveries, population	descriptive
CDC	Centers for Disease Control and Prevention (2020b)	Government	geography, time, demographics, cases, deaths, per capita data	descriptive
CO	Colorado Department of Public Health & Environment (2020)	Government	geography, time, demographics, medical capabilities, public sentiment, cases, deaths, tests, per capita data, outbreaks	descriptive
CP	Gu (2020)	Independent (endorsed by US CDC)	geography, time, cases, deaths, reproduction rate	predictive (SIR with ML)
CU	Branas et al. (2020)	Academia/Research (endorsed by US CDC)	geography, time, medical equipment, intervention policies, cases	predictive (SIR)
CV	Mamoon and Rasskin (2020)	Independent	geography, time, cases, deaths, recoveries	descriptive
ECDC	European Centre for Disease Prevention and Control (2020)	Government	geography, time, demographics, cases, deaths, incidence report, per capita data	descriptive
GN	Google News (2020)	News	geography, time, cases, deaths, recoveries, per capita data	descriptive
ICL	Imperial College London (2020)	Academia/Research (endorsed by US CDC)	geography, time, deaths, cases, reporting ratio, reproduction number	predictive (SIR ensemble)
IHME	The Institute for Health Metrics and Evaluation (2020)	Academia/Research (endorsed by US CDC)	geography, time, medical equipment, intervention policies, cases, deaths, tests, mobility, per capita data	predictive (SIR with curve fitting)
IIB	Information is Beautiful (2020)	Academia/Research	geography, time, health conditions, demographics, other diseases, mask materials, cases, deaths, per capita data, infection severity, reproduction rate, incubation period, media traffic	descriptive
JHU	Johns Hopkins University (2020)	Academia/Research (endorsed by US CDC)	geography, time, demographics, medical capabilities, cases, deaths, recoveries, population	predictive (SIR)
LANL	Los Alamos National Laboratory (2020)	Academia/Research (endorsed by US CDC)	geography, time, cases, deaths, model performance	predictive (growth model with forecasting)
MIT	Bertsimas et al. (2020)	Academia/Research (endorsed by US CDC)	geography, time, per capita data, health conditions, demographics, policies, medical capabilities, intervention policies, cases, deaths	predictive (SIR)
MOBS	Vespignani et al. (2020)	Academia/Research (endorsed by US CDC)	geography, time, intervention policies, medical capabilities, cases, deaths, mobility	predictive (individual-based, stochastic, and spatial epidemic model)
NY	New York State Department of Health (2020)	Government	geography, time, health conditions, demographics, cases, deaths, population, tests	descriptive
NYTa	The New York Times (2020a)	News	geography, time, demographics, cases, deaths, population	descriptive
NYTb	The New York Times (2020b)	News	geography, time, cases, deaths, population	descriptive
OWiD	Our World in Data (2020)	Academia/Research	geography, time, intervention policies, cases, deaths, tests, population	descriptive
UCLA	Statistical Machine Learning Lab at UCLA (2020)	Academia/Research (endorsed by US CDC)	geography, time, cases, deaths, recoveries, reproduction rate	predictive (SIR)
UG	Flahault et al. (2020)	Academia/Research (endorsed by US CDC)	geography, time, cases, deaths, reproduction rate	predictive (GLM)
UT	The University of Texas COVID-19 Modeling Consortium (2020)	Academia/Research (endorsed by US CDC)	geography, time, deaths	predictive (curve-fitting)
VA	Virginia Department of Health (2020)	Government	geography, time, demographics, cases, hospitalisations, deaths, population, tests, outbreaks	descriptive

Table A2. Early Stringency Measure Summary Data.

Early stringency measure	In effect before 10th reported case	Avg cases per 100 K (as of May 24)	Avg deaths per 100 K (as of May 24)	Num countries
Public Event Cancellations	Yes (Nationwide)	43.21	1.66	97
	Yes (Targeted)	72.32	3.66	9
	No	243.92	11.63	66
Public Gathering Restrictions	Yes	39.35	1.31	82
	No	199.72	9.64	90
International Travel Controls	Yes	92.09	3.36	137
	No	239.09	14.27	35

Table A3. Persistent Stringency Measure Summary Data.

Persistent stringency measure	High level in effect after 10th reported case[3]	Avg cases per 100 K (as of May 24)	Avg deaths per 100 K (as of May 24)	Num countries
Public Events Cancellations	Yes	45.83	1.89	109
	No	250.12	11.85	63
Public Gathering Restrictions	Yes	54.37	2.42	48
	No	149.74	6.91	124
International Travel Controls	Yes	83.12	2.87	100
	No	176.68	9.41	72
Testing	Yes	107.38	2.99	45
	No	128.45	6.60	127
Contact Tracing	Yes	42.60	1.76	14
	No	130.25	6.01	158

Table A4. Summary of Country Comparisons.

	When initial stringency measures were taken …						On May 24 …	
Country	Stringency Score	Cases	Days before/ after (±) 10th case	Early Stringency	# of Cases when stringency score exceeded 50 (or at their highest level)	Both Widespread Testing and Comprehensive Contact Tracing[4]	Cases per 100 K	Deaths per 100 K
Zimbabwe	5.56	0	72	1685.02	2	No	0.39	0.03
Paraguay	5.56	0	55	752.05	6	Yes	12.39	0.16
Czech Republic	5.56	0	42	669.52	94	Yes	84.26	2.96
Cuba	11.11	0	51	655.56	40	Yes	17.12	0.72
Vietnam	2.78	2	11	258.33	95	Yes	0.34	0.00
Ecuador	5.56	0	39	225.15	28	No	215.14	18.19
USA	5.56	8	1	11.12	3,774	No	503.00	29.91
Iran	2.78	2	2	5.56	20,610	No	165.89	9.07
Egypt	11.11	93	−8	0	256	No	17.54	0.78
Sweden	11.11	203	−9	0	16,755	No	328.83	39.29