

What is Governor Polis Not Revealing about COVID-19 Modeling?

16 April 2020

Dave Kisker, Ph.D.

People United for Responsible Government

Over the last several weeks, the Colorado Governor, Jared Polis, has paralyzed businesses and terrified our citizens with his claims of a doomsday scenario as a result of the dire predictions of the impact of the COVID-19 virus in Colorado. The “stay at home” order that was issued on March 26, 2020 was not simply a recommendation for how to achieve “social distancing” which is an accepted method of slowing the spread of a virus-based pandemic threat. Rather, it was an order which has resulted in a dire impact on countless small businesses, costing the state hundreds of millions if not billions of dollars of economic activity. Cities and Counties are threatening fundamental rights guaranteed by the US Constitution, including barring transit across their boundaries. Citizens are implementing a type of “new McCarthyism” as they report their neighbors for perceived infractions. Thousands of additional hospital beds and ICU units have been prepared based on the supposed need that has been predicted by the Governor’s COVID-19 team.

It is not the intention of this writing to understate the importance and significance of the COVID-19 pandemic. Based on events in other western countries such as Italy, Spain, France and the UK, it’s clear that taking the threat seriously is totally and completely appropriate. And, indeed, the presence of “hotspots” here in Colorado in locations such as nursing homes confirms the reality of COVID-19. Instead, the purpose of this writing is to investigate and report the validity of the model apparently being used by the Colorado COVID-19 team.

These concerns occur primarily because of the unwillingness of the Polis group to embrace transparency and provide complete details of the modelling procedures they have been using to justify the extreme, negative economic impact on that the State of Colorado. This policy is particularly aggravated because it has been essentially impossible to achieve any evaluation of the accuracy of the models for disease progression that is being used. At the same time, the members of the Colorado modeling team have disparaged other modeling efforts by the Institute for Health Metrics and Evaluation (IHME) housed at the University of Washington.

Without an ongoing evaluation of the validity of a mathematical model, there is no possibility of accountability or confirmation of the appropriateness of the economic actions that have paralyzed our state.

In this document, we will present an analysis of the Polis group’s COVID-19 model as well as the IHME model as it applies to Colorado.

Background

Considerable media reporting has been done regarding the initial incidence and progression of the COVID-19 pandemic, not only around the world but also in the US. Both at the national level and the State level in Colorado, estimates of the risk and impact of COVID-19 have been based on mathematical models of the progression of the disease. Factors such as the “case fatality rate”, the period of infectiousness, the likelihood of a patient needing ICU care as well as the “reproduction number”, known as R_0 , are built into the models, either directly or indirectly. However, continual evaluation of the validity of such a model is absolutely critical to determine whether the predictions are valid. In a case such as COVID-19, which has stimulated such extreme actions by the Governor, this evaluation must be continual and transparent so that the citizens’ confidence in the analyses being presented is maintained.

Although some limited actions had been previously taken in Colorado, in his 3/27/2020 press conference, Governor Polis showed the information presented in Figure 1. He claimed that as many as 33,000 Coloradans would die due to COVID-19, if nothing was done, with a best case scenario of around 12,000 fatalities.

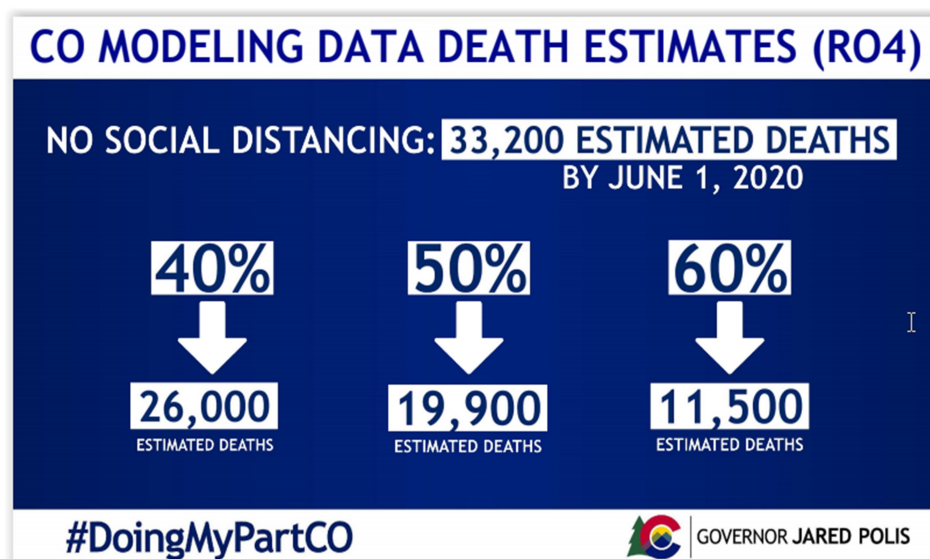


Figure 1. Slide from Polis press conference, 3/27/2020, in which the catastrophic threat presented by COVID-19 was asserted.

During the ensuing discussion, Polis explained that the reason for this very large potential fatality rate was the “ R_0 ” value of 4. He claimed that for some vague reason, Colorado’s R_0 value was thought to be between 3 and 4, despite the fact that considerable evidence from other locations estimated that the value was between 1.5 and 2.0.

There were immediate questions regarding the model because the details had not been released, and for unknown reasons, the Polis team was reluctant to release the details despite repeated requests from

the media.¹ CDPHE's records manager told Colorado Politics on Friday [4/3/2020] that they needed to delay the release of the records because of "extenuating circumstances." However, Jeff Roberts, the executive director of the Colorado Freedom of Information Coalition, said the state health department should, especially in times when providing information can mean the health of Coloradans, provide access to reports that are compiled and complete. And they should provide them proactively, he said.²

Finally, on 4/6/2020, the Polis group released a report³ that purported to explain the modeling approach that was being used to justify the actions the Governor had taken to require citizens to "stay at home". However, as we shall show below, in fact, that report does not address the key question which is, "what is the source of the claims of tens of thousands of Colorado fatalities?" As a result of the incomplete story that the Polis team persists in telling, the skepticism is growing.⁴

Using Mathematics to Model Physical, Chemical and Biological Processes

Although there are those who would attempt to wrap the use of mathematics to model/estimate the natural phenomena in our world in a veil of mystery, the fact is that while the specific equations may be quite complex, in reality the process is not nearly as mysterious as many think. Similarly, the fundamental techniques are not as flawed as some would claim, as long as the range of applicability of the model is made clear.

In fact, the great mathematician, John Von Neumann spoke of this process in this manner:

"The sciences do not try to explain, they hardly even try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work-that is, correctly to describe phenomena from a reasonably wide area."

The fact that a particular mathematical model of natural phenomena is able to correctly describe that phenomena is very useful in science. And, in a case such as the current COVID-19 pandemic, modeling may well be the best method to understand the disease progression, the implications, and to assist in developing public policy, especially in the early phases. *However, in order for this to be the case, the model must be tested so that 1) it's range of applicability is understood, and, 2) its accuracy can be validated.*

¹ https://coloradosun.com/2020/03/27/coronavirus-covid-epidemiology-curve-data/?fbclid=IwAR1jpcBoLwg10clJ0QrxGtJB0Li3dQpr9S8Y88X6t818ee2Bp_uKs65Y_hE

² https://gazette.com/news/local/state-refuses-to-let-public-see-local-virus-spread-model/article_bdb8045d-696d-53da-b078-f570e4a710b5.html?fbclid=IwAR2RePJcMDHw3FSezGbYzZYvBjeULEB8R3o-43q5TUJdR_Hnqx0Aj5hZ_-A

³ <https://drive.google.com/file/d/19ST3tRnntUGV3h7YI0tQHJDnCoL5dRE7/view>

⁴ http://freethetheconomy.com/why-are-the-state-of-colorado-coronavirus-statistics-so-far-off/?fbclid=IwAR3g52AAmOOYZYBdGjdg_thxEj89RJhgGDzD4GNbJ5tXsmXmOR3rnXisjZs

The primary way that scientists validate any mathematical model is that a portion of the data which is used to develop the model is also used to validate it. This is done by first evaluating the model's accuracy when used to describe the existing results. The second step is to withhold a portion of the measurements and use those to test how well the model performs on "out of sample" data—that is, the data points that were not used to develop the model in the first place. This is especially important when the model is used to predict future results, as is the case for COVID-19 predictions. Of course, if the input data is of poor quality because it is inaccurate or incomplete, the model itself may be of questionable validity, thus the phrase "garbage in, garbage out".

Finally, since any mathematical model requires that several parameters must be adjusted so that the mathematical description fits reality, the best models are those that have had their parameter estimates varied to see how "sensitive" the predictions are to those adjustments. This procedure helps to determine the range of applicability, and therefore, the confidence that we can have in the model.

Because it always possible to represent a set of results from a natural phenomenon if enough adjustments can be done, it's important to use as few fitting parameters as possible. Turning again to Von Neumann, who said:

"With four parameters I can fit an elephant, and with five I can make him wiggle his trunk."

Examples of COVID-19 Models

It's beyond the scope of this writing to examine the fundamental principles of the particular epidemiological models being used by researchers to explain and predict the behavior of COVID-19 as it proceeds through a population, and as various attempts are made to develop public policy based on these models. ***Rather, in this writing, we are attempting to answer the question posed by von Neumann: Does a particular model work?***

US National Model for COVID-19 deaths.

The first example which we evaluate is the COVID-19 model from the IHME, located at the University of Washington.⁵ This is apparently the model used by the US CDC to study the progression of the pandemic as it has developed in the US. According to the IHME, the model for the US is regularly adjusted using new data⁶ as it becomes available. This is also the model that was recently projecting that between 100,000 and 200,000 Americans would succumb to COVID-19. As the various limitations imposed by social distancing around the country have taken effect, the model parameters have been adjusted, and now, the same approach is predicting that a mean estimate of fatalities in the US would be about 69,000. The uncertainty range is still quite large, between 30,000 and 176,000, but these estimates were previously even higher.

⁵ <https://covid19.healthdata.org/united-states-of-america>

⁶ <http://www.healthdata.org/covid/updates>

To demonstrate the use of “out of sample” data, we have been tracking an earlier version of the IHME model that was first published on 4/2/2020. Because we have new results each day, it’s instructive to see how successfully an existing version can predict future data. In Figure 2, the results from the IHME model that was published on April 2 along with the fatality results over the last 15 days are shown. The ability of this model to project into the future by as much as 15 days with good accuracy, is promising. While it’s clearly important to continue to track the performance, this model quality is quite useful. Of course, as we see in the last few days, the 4/2/20 model is beginning to show signs of being slightly too pessimistic—it predicted a total fatality rate of about 80,000 by around 5/10/2020. Current results

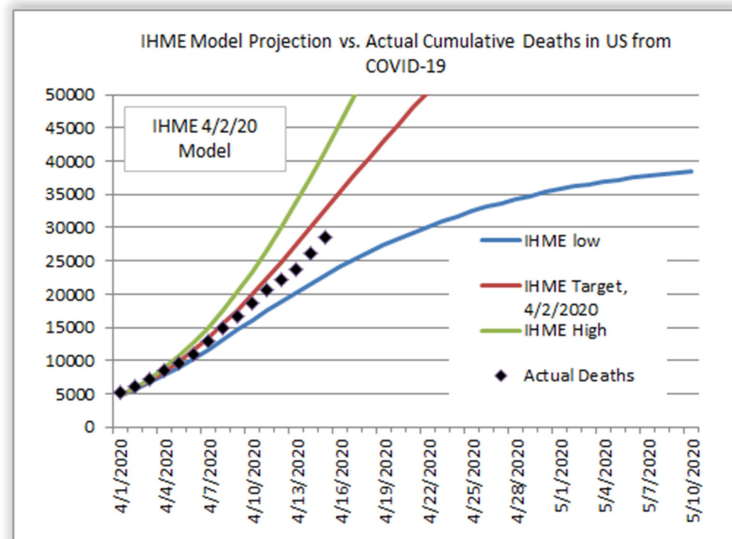


Figure 2. COVID-19 model of fatalities in the US. Released by IHME on 4/2/2020.

suggest that total number of fatalities will be lower as mentioned above⁷. But, this demonstration of “out of sample” validity gives us confidence that this approach to estimating the COVID-19 impacts will continue to be useful.

IHME Model for Colorado COVID-19 Fatalities

Turning now to the question of assessing the reality of the doomsday scenario predicted by Polis on 3/27/2020, we must first note that as of this writing, the Polis group has not released sufficient details of their modeling efforts to determine if and how well that effort reflects reality. Although there is some information in the 4/6/2020 release, the key missing information is the timing of the fatalities that are anticipated in our state. Remember that Polis threatened that as many as 30,000 Coloradans could succumb if we “don’t do anything” and it could be as many as 11,000, even if we accomplish 60% “social distancing”. However, since the Colorado model is not available, we turn instead to the efforts by IHME for the state of Colorado.

It should be noted that in the last few days, there have been some criticisms of the IHME models⁸ for the states, but we believe that is a limitation of that analysis rather than of the IHME model.

⁷ During the preparation of this manuscript, the CDC has announced new guidelines for reporting COVID-19 deaths. This has resulted in a major shift in the daily report of cumulative deaths in the US. An updated figure is shown in the Appendix.

Similar to the US model, IHME regularly updates their models for the anticipated fatality rates in the states. Below, we show two examples of the IHME approach. Figure 3 shows the IHME projection for Colorado based on their analysis through 4/2/2020. Two things can be observed: First, we see that this model predicted a much lower number of fatalities than even the 11,000 presented by Polis. Instead, the IHME model projected a total of around 2,000 fatalities. It's interesting to note that when questioned about this discrepancy, the Polis group attempted to discredit the IHME study, saying that because it "relied too much on data from Wuhan, China, it was invalid". Of course, this completely ignores the von Neumann principle that the most important characteristic of a mathematical model is that "it works"!!!

Of course, by around 4/7/2020, the IHME model from 4/2/2020 appeared to resulting in a significantly higher prediction than the Colorado reality was suggesting. While the daily results were fitting the

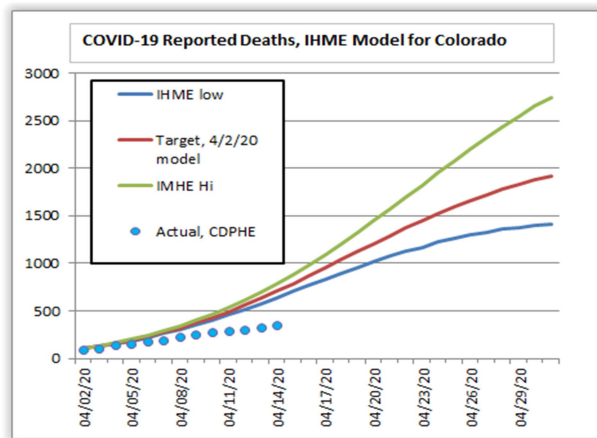


Figure 3. IHME model for Colorado fatalities. Model published on 4/2/2020. Total estimated fatalities are about 2000.

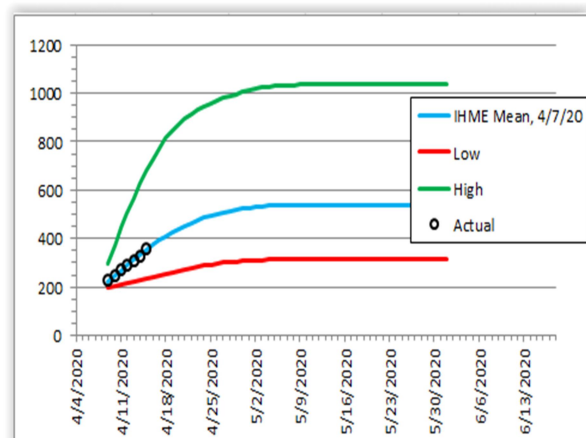


Figure 4. IHME model for Colorado fatalities, update on 4/7/2020. Total estimated fatalities are about 550.

general trend, they were falling outside the bounds of estimated uncertainty. At that point it was desirable to update the model. Given the generally good representation, IHME updated their model, and that new result is now shown in Figure 4. With about a week of "out of sample" data, this revised IHME model continues to appear very promising for projecting Colorado's COVID-19 fatalities.

Polis Model for Total COVID-19 Cases in Colorado

As mentioned earlier, an evaluation of the modeling approach that is being used by Polis et al. cannot be compared on the same basis with the IHME models since the Polis model has not published any time series data that addresses the cumulative number of fatalities from COVID-19. However, in the 4/6/2020 report, an estimate of the total number of COVID-19 cases that can be expected to occur in Colorado is included. The calculation is presented for various results as a function of what is referred to

⁸ A group from Australia has asserted that the state models accurately predict the fatalities in only a few cases. In fact, the Colorado situation is one that they claim is not a good result. However, their approach relied on very short term data, which is extremely difficult to assess and often fluctuates wildly. In this work, we look at the longer term trend, using out of sample data.

as “social distancing”, including values from 0% to 80%. However, there is no information in the report about how this parameter is incorporated into the model. Nonetheless, Governor Polis has stated that the goal of the “Stay at Home” lock down policy is to achieve 80% social distancing. The implication is that the model will inform any decision to relax the lockdown. Therefore, by tracking this “Total COVID-19 Cases” model, an estimate of whether or not the lockdown has been effective should be possible,

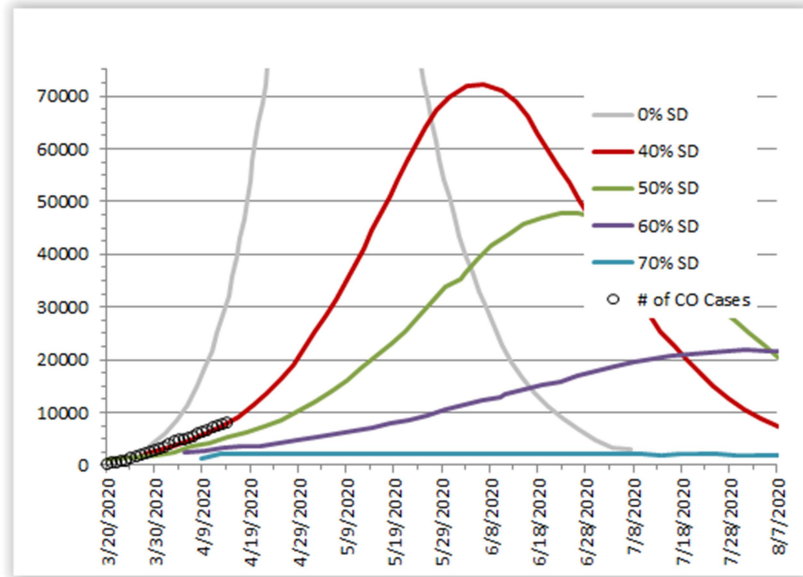


Figure 5. Comparison of actual reported COVID-19 cases with the calculated predictions from the 4/6/2020 Polis report (Figure 6 of that report). Note the apparent lack of any change in trend that would indicate a major effect of social distancing.

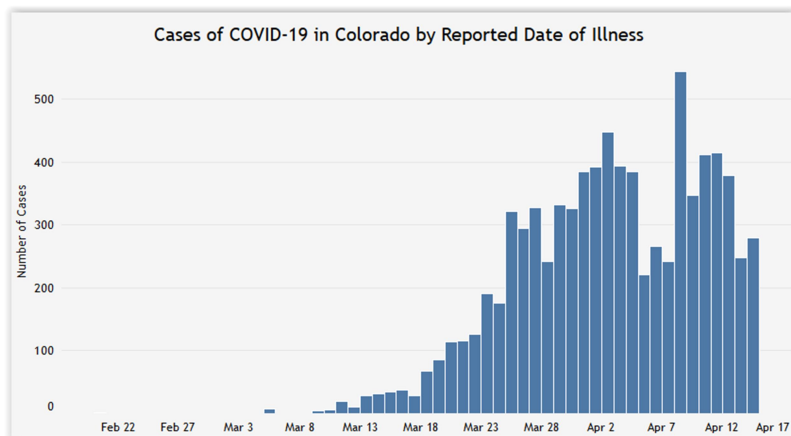


Figure 6. Daily new cases of COVID-19 reported by CDPHE.

to reflect any indication of the impact of social distancing, nor why the results seem to only be consistent with the initial results reported by Polis after the first phase was instituted—45%. However, the results clearly demonstrate the difficulty of achieving any sort of scrutiny or accountability when the entire story is not being revealed.

assuming that the model is valid. This result, reproduced from Figure 6 of the 4/6/2020 report, along with the reported number of cases is shown in Figure 5.

Interestingly, the number of actual reported cases follows the calculation for an SD value of 40%.

Furthermore, there is no evidence of any change in the trend in the reported data, which suggests that this model is not able to reflect the actual effect of social distancing. At the same time, as has been reported by both Polis and the media, the number of new cases of COVID-19 in Colorado has generally flattened resulting in a greatly reduced cumulative growth rate. This effect is apparent in the 4/15/2020 data reported by CDPHE⁹, reproduced in Figure 6.

Of course, it is not clear why the model reported in the 4/6/2020 report does not seem

⁹ <https://covid19.colorado.gov/case-data>

In the next section, we will briefly consider new research that has just been reported that may help to explain the apparent inconsistencies between the results predicted by the Polis model, and other models, in particular, the IMHE model which seemingly fits the Colorado results so well.

What can go wrong?

While it's far beyond the scope of this brief writing to try to analyze the reasons that lead to these observations, a recent report by Korolev¹⁰ may shed some light on a key issue. In that work, it is shown that "There are many sets of parameters that are observationally equivalent in the short run but lead to markedly different long run forecasts." For the current purpose, we only reproduce a pair of the figures from that work which are illustrative.

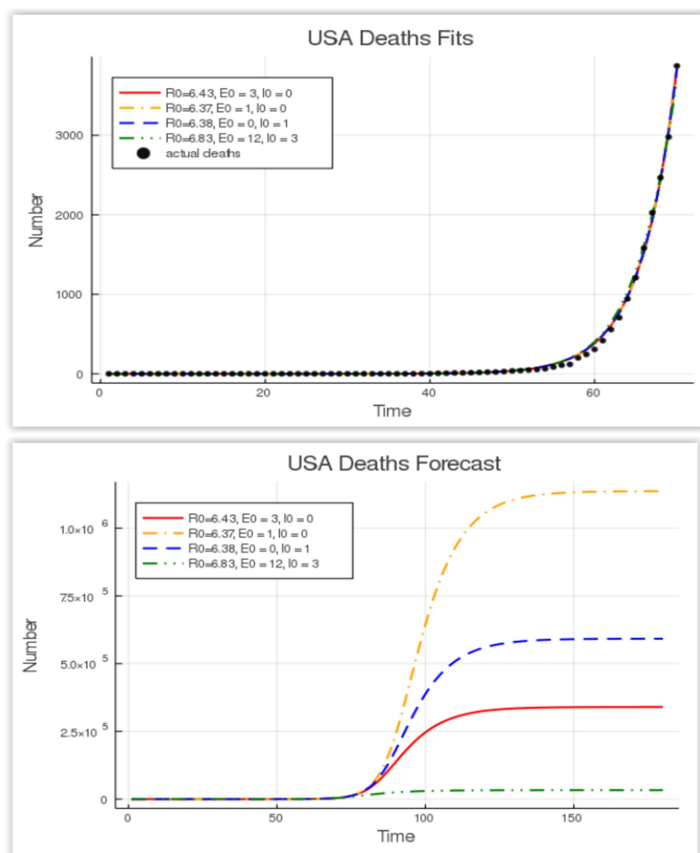


Figure 7 (upper). Fit of the actual deaths and reported cases by models with four different sets of initial conditions. Figure 8. (lower) Forecasts from the same four models.

In brief summary, Korolev explores the possible effect of not thoroughly investigating the long term ability of the standard "SEIRD"¹¹ model to fit "out of sample" epidemic progression data. We won't address the details of the model he is using, except to observe that he is evaluating a standard approach to this type of problem.

The next two figures, Figure 7 and 8 are taken from Figure 4 of the Korolev publication. Figure 7 shows the excellent agreement between the calculations starting with 4 different sets of parameters and the fit to actual reported fatality data for the US. Each of these fits determines values for the model parameters R0, E0 and I0 using the first 60 days of data. In Figure 8, these 4 models are then extended to forecasts out to 180 days. The predicted results show as much as more than 10X difference in the

predicted fatalities! Which brings us back to the essential view of von Neumann: **a mathematical model is useful simply because it works.**

¹⁰"Identification and Estimation of the SEIRD Epidemic Model for COVID-19", Ivan Korolev.

<https://drive.google.com/file/d/16fhm4DbYFKKkywmvCBtjWfdqjo-ZPx6E/view>

¹¹ SEIRD models generally describe epidemics using equations for the numbers of Susceptible, Exposed, Infected, Recovered and Deceased members of the population.

This behavior of apparently similar models is a clear demonstration of why it's absolutely critical that continuous evaluation of a proposed mathematical model is critical, especially when public policy decisions depend on proper representation of the "facts on the ground". **Only through complete transparency, is it possible to credibly assert that steps such as the extreme lockdown currently in effect in Colorado are valid.**

Summary

In this brief document, we have attempted to evaluate the validity and performance of the mathematical models being used to predict and document the behavior of the COVID-19 pandemic currently progressing in Colorado, the U.S. and around the world. Because of the lack of transparency of the Polis administration, despite the extreme economic and social costs of actions such as the "Stay at Home" mandate, we find that there is currently no evidence that the doomsday scenario that has been used to justify these measures is valid. There are inconsistencies with the claims that at least 80% "social distancing" is needed before it will be safe to release Colorado to begin to revert to normal since the published Polis model doesn't indicate that a trend toward 80% is either possible or underway.

In contrast, the models for the U.S. and Colorado that have been published and documented by the IHME seem to do a much better job at predicting the deaths caused by COVID-19 and, predict much lower values of fatalities, totaling under 600 deaths by August 2020.

Hopefully this analysis is sufficiently accessible by non-experts that an honest assessment of the approach towards public policy that has been implemented by the Polis administration can be undertaken. **While it's possible that the entire scenario as presented until now has been scientifically legitimate, our current level of access to the details is not sufficient to make that determination.**

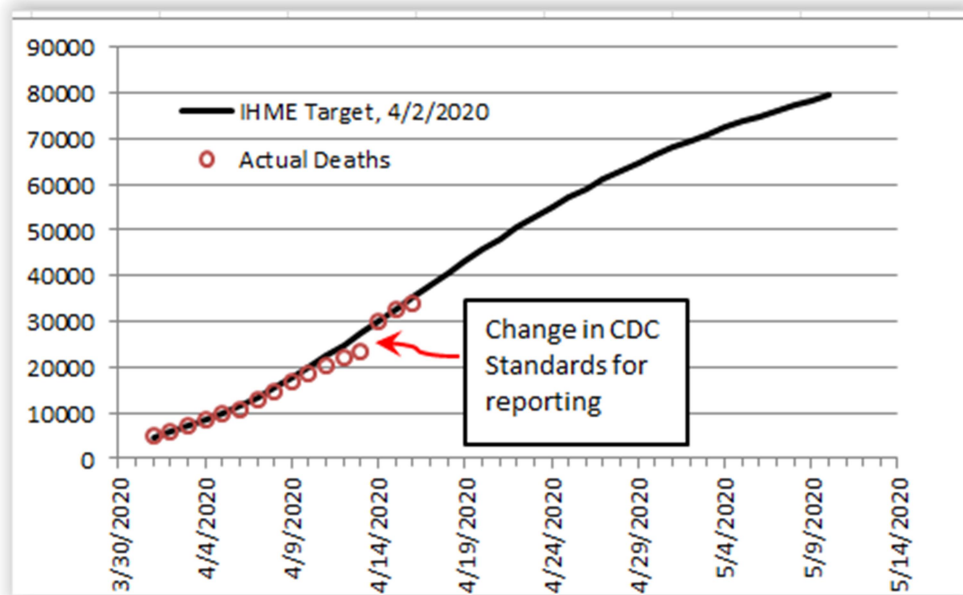
Appendix

As mentioned earlier⁷, while this document was in preparation, the US CDC has changed the standard for reporting deaths associated with the progression of COVID-19. In particular, the statement says:

"As of April 14, 2020, CDC case counts and death counts include both confirmed and probable cases and deaths. This change was made to reflect an interim COVID-19 position statement issued by the Council for State and Territorial Epidemiologists on April 5, 2020. The position statement included a case definition and made COVID-19 a nationally notifiable disease.

A confirmed case or death is defined by meeting confirmatory laboratory evidence for COVID-19. A **probable case or death** is defined by i) meeting clinical criteria AND epidemiologic evidence with no confirmatory laboratory testing performed for COVID-19; or ii) meeting presumptive laboratory evidence AND either clinical criteria OR epidemiologic evidence; or iii) meeting vital records criteria with no confirmatory laboratory testing performed for COVID19"¹²

In response to this new information, we have updated our Figure 2 with the new, most recent data. Note that the model that is shown remains the IHME model that was published on 4/2/2020. The key observation at this point is that the actual number of fatalities is represented even more accurately by the IHME model. As this issue progresses, we will update this figure.



¹² <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html>