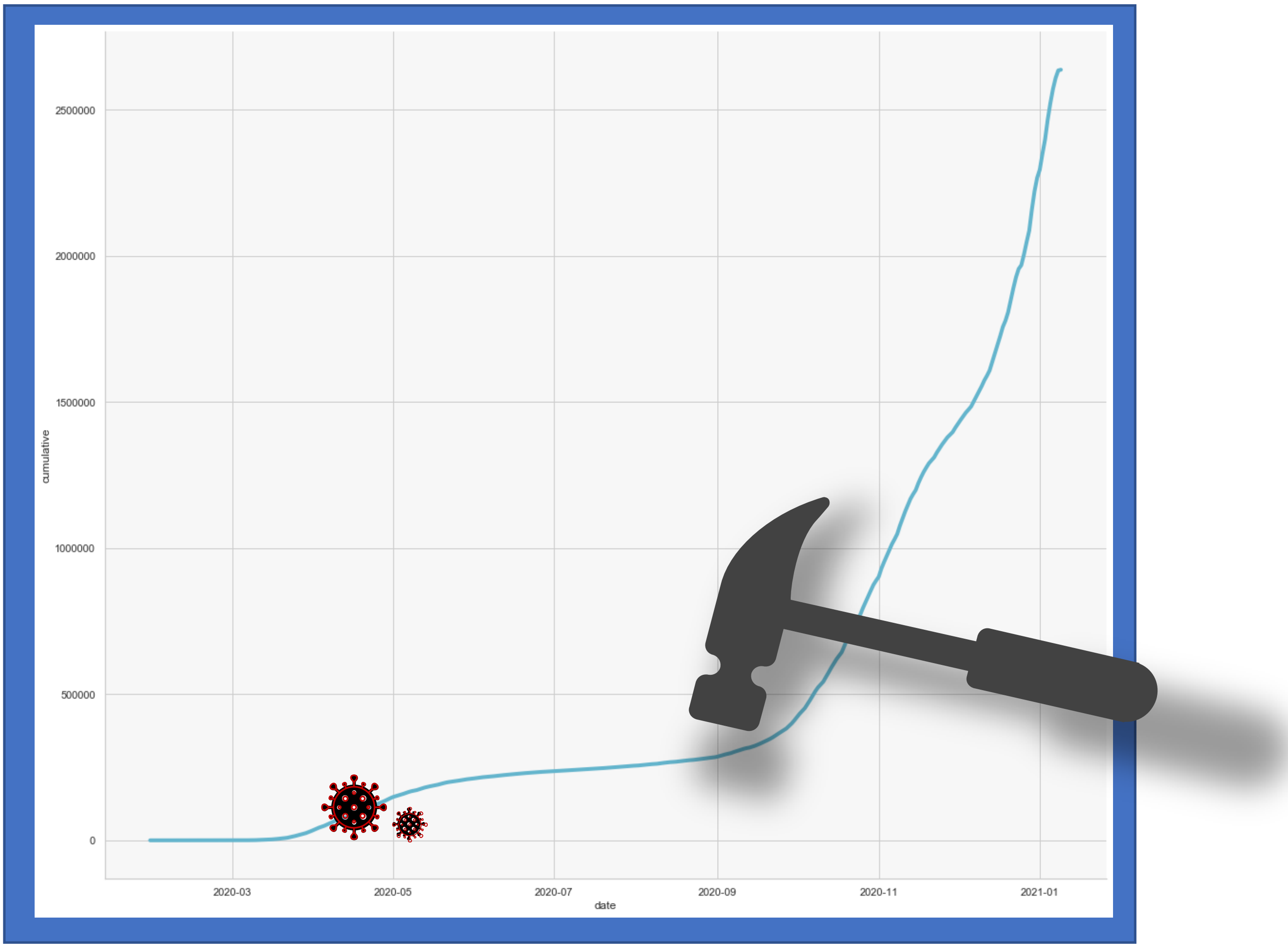


FLATTENING THE CURVE:

Is lockdown in UK effective?

TABLE OF CONTENTS

FALTTENING THE CURVE: IS LOCKDOWN IN UK EFFECTIVE?

INTRODUCTION

Amid the 2019 covid-19 pandemic, the virus ability to fast-transmission has put stress on medical resources in many countries including the UK (Sheposh, 2021). Government authorities around the world were forced to imposed lockdown measures to reduce the spread of the coronavirus.

On March 16th, UK Prime Minster Boris Johnson (2020) pledged to "*flatten the peak of the epidemic*" by "*bringing the right measures at the right tim*e". In the UK, two national level lockdowns were introduced before 2020 Christmas where people were forced to work from home. However, it has been suggested that people have been defensive about the lockdown. Smith et al. (2020) discovered that people has been reluctant to adhere to lockdown policies which makes flattening the curve more difficult.

So how has the curve has been "flattened" in the UK, did UK left restriction too soon? The prime purpose of this report is to analysis if the non-vaccine interventions in the UK (in specific, England) is effective in "*flatten the curves"* . The report aims to accomplished in three steps:

- One, slicing data into two periods. One when lockdown was implemented in full speed ("*full on lockdown mode*") and one when some restrictions were lifted ("*loose restrictions*"). Within each *period*, subsets **to training sets** and **validation sets**.
- Two, modelling: do statistical modelling for each of these periods, choose the best model and validate.
- Three, compare parameters of these models based and contrast two policies before and after lockdown.

In addition, data analytics wishes to uncover the uncovered (Provost and Fawcett, 2013), so alongside this the main quest, we wish to disclose new findings along the journey.

METHOD

DATA SELECTION

The data used for this analysis were obtained from The Official UK Government Website for data and insight on Coivd-19. Those data were processed in R and visualised in Python and were directly called through API (Application Programming Interface). Those data traced for as early as March 2nd, 2020 and are live updating.

We want to find out if lockdown measures in UK is effective in restricting amount of people infected by covid-19. This is going to be problematic if reported infection rate is inaccurate. Obviously, the amount of test conducted is going to influence amount of people who were tested positive. In fact, UK government has not introduced mass testing program(pillar 2 test) until early July. Before July, one can only find out who has covid by test those who checked into hospital (pillar 1 test). Those facts give amount of people who were infected with covid a very inaccurate estimate, a common challenge of estimate covid cases in any other countries(). The officials themselves acknowledges that they were not able to report accurate figure until 27th of October (2020).

Cumulative covid death was used as a poxy to how many people were actually infected to avoid the nuisance of inaccurate data during early pandemic. The premise of using death case as proxy is that covid-19 death case is in parallel with reported covid-19 infection cases. This assumes covid-19 death is independent over time and the death rate is constant.

SLICING DATA

The first slice of data ranges from Feb 23rd to Jun 15th before government starts to lifting restrictions. The second slice ranges from July 13th to October 31st before which the second lockdown ends. The reason for slicing data this way is because restrictions was not fully lifted until July 13th. For the purpose of this analysis, we want to isolate policy making

Provost (2013) suggests one should set up **training sets** and **validation sets** to avoid overfitting. Hence, within each slice, we set up 20 days period for validation sets.

MODELLING

Organic growth usually has sigmoid growth pattern, or what is known as S-curve. Micro-organisms are likely to infect hosts in similar pattern. Organisms grow geometrically without constraints to resources (Karkash, 2006). When constraints of resources were introduced, organic growth follows a sigmoid pattern, that is it starts off growing geometrically but then plateau as if it reaches a "ceiling". Biological growth pattern gives bases to model covid.

Miro-organisms such as virus are likely to grow in sigmoid patterns. In the case of covid, population in UK can be considered as a resource. Non-vaccine-based interventions (such as wearing masks or reduce social contact) does not "kill" the virus but it restricts permeation of the virus by restricting resources.

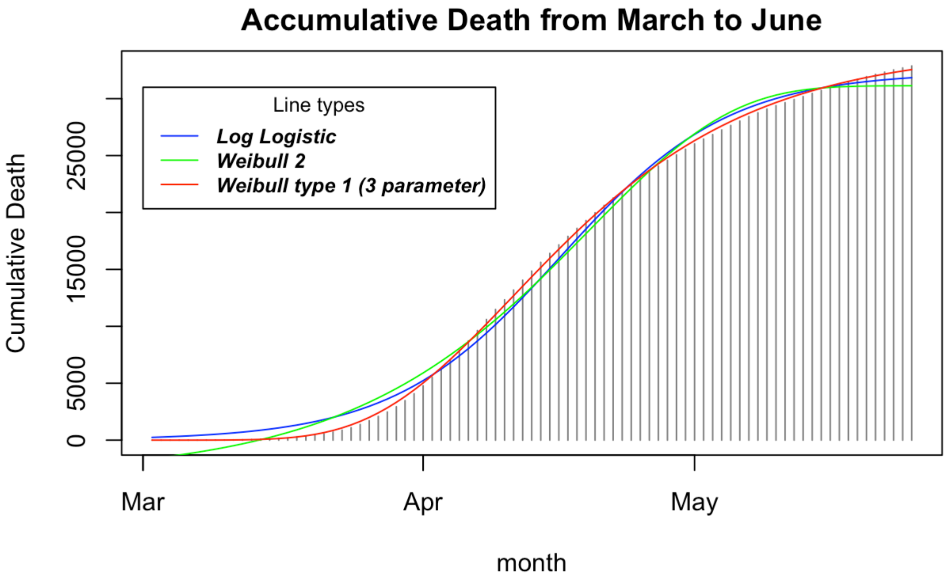
Parametrization of *beta*, the ceiling of growth or upper asymptotic value, is usually used non-determinate growth model (Karkash, 2006). Modelling growth of covid overtime makes estimating beta possible. One can compare beta value before and after lowdown to see how effective lockdown is.

Those models were built in R using package "drc" (R Core Team, 2020) to model non-linear regression.

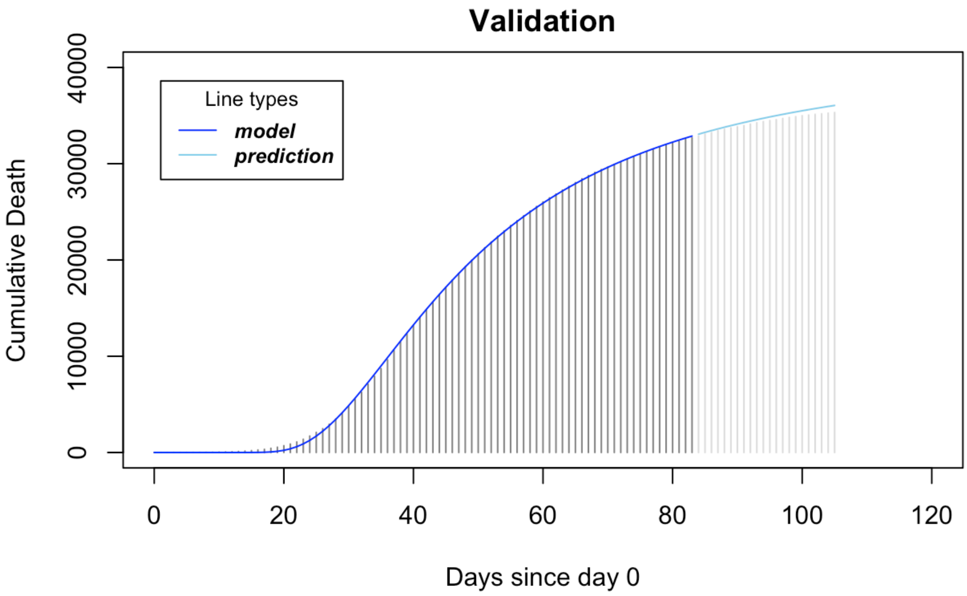
RESULTS

Modelling

Several models have been used in order to find a fit for the data. Model found the best described by a 3 parameter Weibull Type 1 curves (orange line in the graph below).

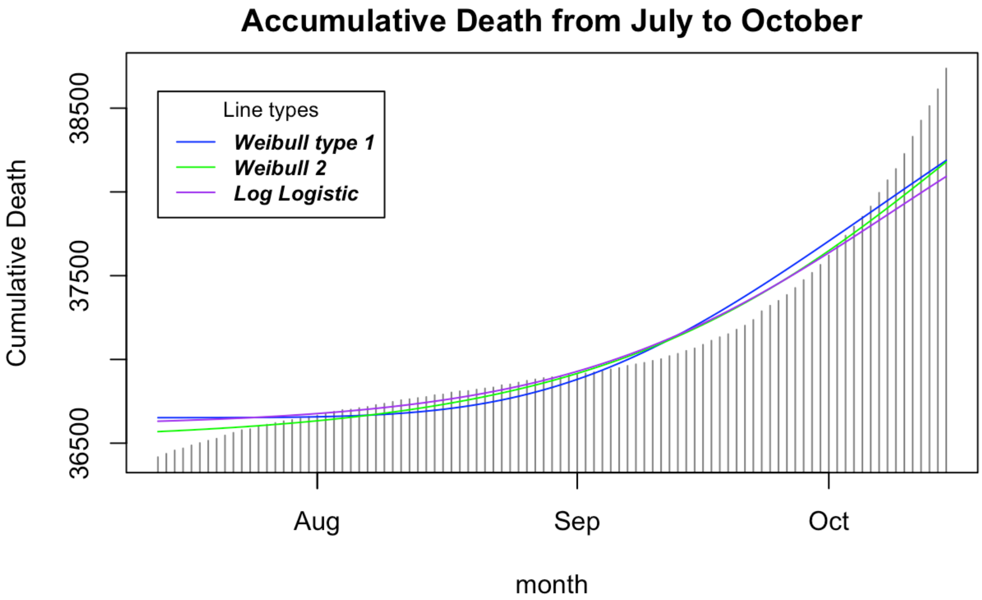


By Intuition, type 2 Weibull curve seems to best represent training set as it fit is sigmoid property as well as fact that its asymmetric at the inflection point. For the validation, this model has been applied to the *holdout set* (Validation set) we reserved previously.



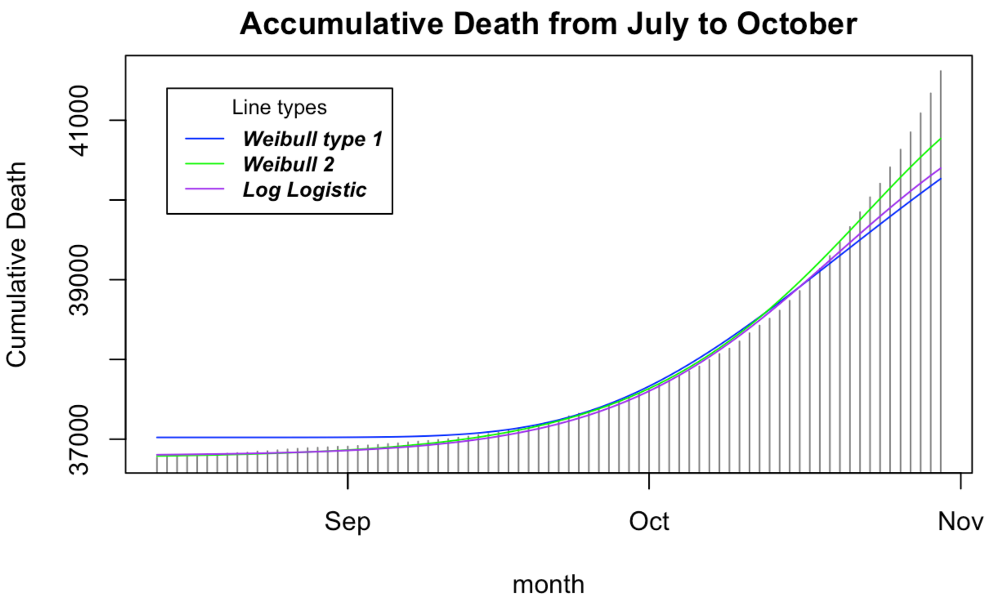
Generalisation seems good only by intuition. The model seems to have roughly predicted the validation set however this prediction is biased. The model tends to predict value than actual.

Initially, for the second set of data, training data failed to yield good fits. This is due to an issue that has not been foreseen previously. We has used cumulative death as a proxy to infection rate, one can expect it take times before virus infect to results in death of its host. In fact, the virus has to incubate in its host before it reaches to a detectable level. World Health Organisation WHO (2020) suggests it takes 10 - 14 days before the virus will show symptoms. It takes about another 28 days before the virus kills a host as far as data concerns. Accumulative death in theory will always lagged behind in terms of true infection number. It may take ups to 6 weeks before the data can capture impact of lifting restrictions.

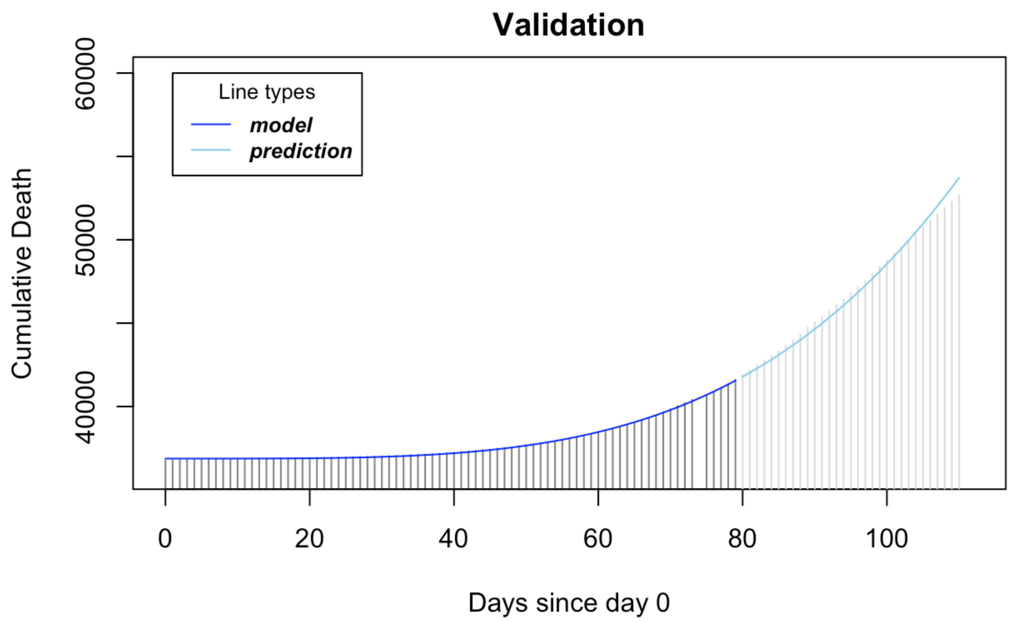


In case of this model, death occurred between July and August were in fact comes from infections in last month in lockdown. So, the in the first few days, data is still taking effect from last period, which explains why the curve flatten first before it grows exponentially.

On our second attempt, we shift this *period* one month forward to remove data the caused from previous period and include death caused before lockdown.



The model clearly improves itself after shifting *period*. The best fit for the model is 4 parameter Weibull type 2 curves (illustrated in green) a variant of Weibull type 1. Sigmoid property of Weibull curves nicely suffices biology growth assumptions.



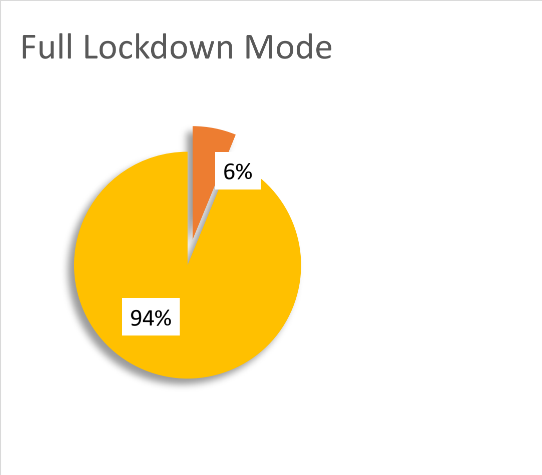
Model seems to fit into the validation set.

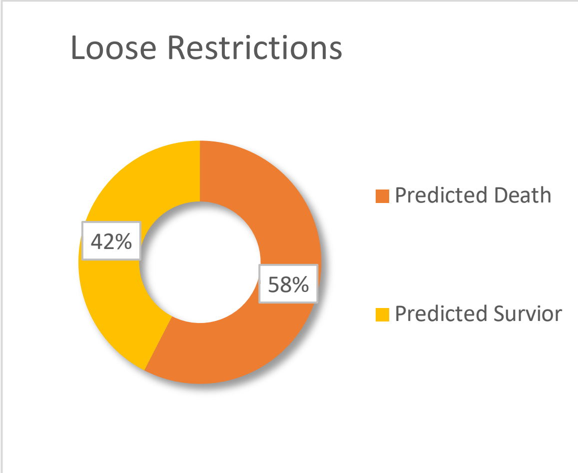
PARAMETERS

It is easy to compare parameters since both models belong to Weibull-sigmoid families. In this estimation, all parameter estimated are *significant* (indicated by *low p-value*).

```
Model fitted: Weibull (type 1) with lower limit at 0 (3 parms)

Parameter estimates:

                 Estimate  Std. Error t-value   p-value
b:(Intercept) -2.1941e+00  2.2083e-02  -99.36 < 2.2e-16 ***
d:(Intercept)  4.1340e+04  2.6134e+02  158.18 < 2.2e-16 ***
e:(Intercept)  4.2423e+01  1.4859e-01  285.50 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error:

 177.7716 (81 degrees of freedom)

Model fitted: Weibull (type 2) (4 parms)

Parameter estimates:

                Estimate Std. Error   t-value   p-value
b:(Intercept) 3.9240e+00 3.3664e-02   116.564 < 2.2e-16 ***
c:(Intercept) 3.6876e+04 6.4120e+00  5751.006 < 2.2e-16 ***
d:(Intercept) 3.8367e+05 1.0000e+01 38366.597 < 2.2e-16 ***
e:(Intercept) 2.3627e+02 2.3829e+00    99.155 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error:

 47.51953 (75 degrees of freedom)
```

For the first model, the upper limits (***d_1*** value) is about 4000 cases of death. Imperial College London estimates covid mortality rate in UK is about 1% (Elsland & Campus, 2020), that is, about 666 thousands of people if everyone is infected by covid. This marks up only 0.06% of the population who could potentially die from covid. The strict lockdown measure is very effective in protecting the covid vulnerable.





The upper limits for the second model (***d_2*** value) is almost 400 thousand. This implies that even if restrictions were much loose than full lockdown, more than 40% covid-victim could be saved from the virus. The difference between upper limits and lower limits (**$d_2 - c_2$**) is about 34 thousand, which is 8 times more than the first model (**$[d_2 - c_2] / d_1$**).

DISCUSSION AND CONCLUSION

This report aims to exam and explore effectiveness of lockdown in England by conducting statistically modelling accumulative death rates. It has been found that shape of the curve during "full-on lockdown" is similar with Rypdal & Rypdal (2020), who modelled covid-19 death case in Global data base. Instead of Gompertz sigmoid model, the data is best described by Weilbull model. The two are similar for its sigmoid pattern and are asymmetric around the inflection point. Parametrisation from this model suggests both of lock down measure are effective, however, full lockdown seems to be far more effective in restricting the spread of the virus. The difference it makes seems significant.

First, using covid death case based on the assumption that the progression of death resulted from covid, is in parallel with number of populations who were infected. This assumption is rather over-simplified. There are more sophisticated models. SEIR (Susceptible Exposed Infectious Recovered; Hethcote, 1989; cited by Ala et al, 2021) provides a more accurate estimate because this take into the consideration of the time lagged between a covid victim is being infected and dead.

Second, data was sliced based on whenever government announces lockdown. Although using domain specific knowledge is appropriate (Provost, 2013), this still very intuitive rather than objective. There are more objective ways. One approach probably by compute derivative functions and find minimum value as a cutting point.

A problem has been encountered when modelling for the second period. Our solution to this is to shift this period forward a few weeks to reflect true effect of policy change. In full acknowledgement, this method is not quite elegant, but it has the supported of domain specific knowledge (Ala et al., 2021) thereby can justify this solution (Provost, 2013).

The fact that the model makes more sense after shifting its period forward reveals significance of *virus incubation times* in affecting modelling. This has very critical implications for governments and policy makers. If this guess is accurate, any observation would be 4 weeks behind from when it observes. In certain cases, this could be dangerous. It could too late for policy makers to adjust their policies if. Policy makers should avoid rely on death rates to monitor and control transmissible diseases such as covid. Issues like these highlights why conducting mass testing is so important.

REFERENCE

Ala, raj, M., Majdalawieh, M., & Nizamuddin, N. (2021). Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. Infectious Disease Modelling, 6, 98–111. https://doi-org.uoelibrary.idm.oclc.org/10.1016/j.idm.2020.11.007

Amer, F., Hammoud, S., Farran, B., Boncz, I., & Endrei, D. (2020). Assessment of Countries' Preparedness and Lockdown Effectiveness in Fighting COVID-19. Disaster Medicine and Public Health Preparedness, 1-8. doi:10.1017/dmp.2020.217

Elsland, D., & Campus, S. (2021). COVID-19 deaths: Infection fatality ratio is about 1% says new report | Imperial News | Imperial College London. Retrieved 11 January 2021, from https://www.imperial.ac.uk/news/207273/covid-19-deaths-infection-fatality-ratio-about/

Izadparast, A. H., & Niedzwecki, J. M. (2013). Four-parameter Weibull probability distribution model for weakly non-linear random variables. Probabilistic Engineering Mechanics, 32, 31–38. https://doi-org.uoelibrary.idm.oclc.org/10.1016/j.probengmech.2012.12.007

Mortality Analyses - Johns Hopkins Coronavirus Resource Center. (2021). Retrieved 11 January 2021, from https://coronavirus.jhu.edu/data/mortality

Prime Minister's statement on coronavirus (COVID-19): 16 March 2020. (2021). Retrieved 11 January 2021, from https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-16-march-2020

Provost, F., Fawcett, T., 2013. Data Science for Business. O'Reilly Media LLC.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richards, F. J. (1959). A flexible growth function for empirical use. Journal of experimental Botany, 10(2), 290-301.

Rypdal, K., & Rypdal, M. (2020). A Parsimonious Description and Cross-Country Analysis of COVID-19 Epidemic Curves. International Journal of Environmental Research and Public Health, 17(18). https://doi-org.uoelibrary.idm.oclc.org/10.3390/ijerph17186487

Sedmák, R., & Scheer, L. (2015). Properties and prediction accuracy of a sigmoid function of time-determinate growth. IForest - Biogeosciences & Forestry, 8(5), e1–e7. https://doi-org.uoelibrary.idm.oclc.org/10.3832/ifor1243-007

Sheposh, R. (2021). Coronavirus Disease 2019 (COVID-19). Salem Press Encyclopedia of Health.

Smith, L. E., Amlôt, R., Lambert, H., Oliver, I., Robin, C., Yardley, L., & Rubin, G. J. (2020). Factors associated with adherence to self-isolation and lockdown measures in the UK: a cross-sectional survey. Public Health, 187, 41–52. https://doi-org.uoelibrary.idm.oclc.org/10.1016/j.puhe.2020.07.024

Vrána, J., Remeš, V., Matysioková, B., Tjørve, K. M. C., & Tjørve, E. (2019). Choosing the right sigmoid growth function using the unified - models approach. Ibis, 161(1), 13–26. https://doi-org.uoelibrary.idm.oclc.org/10.1111/ibi.12592

Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of applied mechanics, 103(730), 293-297.

Wise, J. (2020). Covid-19: New coronavirus variant is identified in UK. BMJ (Clinical Research Ed.), 371, m4857. https://doi-org.uoelibrary.idm.oclc.org/10.1136/bmj.m4857

YIN XINYOU, GOUDRIAAN JAN, LANTINGA EGBERT A., VOS JAN, & SPIERTZ HUUB J. (2003). A Flexible Sigmoid Function of Determinate Growth. Annals of Botany, 91(3), 361–371.