

Predicting the rate of CoVid-19 infected cases by Learning Theory

Romney B Duffey
Idaho Falls, Idaho, USA

Enrico Zio
Centre de Recherche sur les Risques et les Crises (CRC), MINES ParisTech/PSL Université Paris, Sophia Antipolis, France.
Department of Energy, Politecnico di Milano, Milan, Italy.
Eminent Scholar at the Department of Nuclear Engineering, Kyung Hee University, Seoul, South Korea.

We predict the evolution of the daily number of infected cases from CoVid-19 virus. We use the theory of learning from errors, adapted to the problem of virus containment by protective measures such as testing, isolation and social distancing. The theory is consistent with the findings of cognitive psychology on how humans address the solution of errors. Application of these measures leads to the infection rate declining, after reaching a peak. We use publicly available data to predict the recovery rate curve and the time still needed to reach minimum infection rates.

Keywords: CoVid-19, Infection rate, Recovery curve, Learning Theory, Resilience.

1. Introduction

Currently we are experiencing the pandemic of the Coronavirus, which is being battled mainly with extensive or targeted testing and different levels of social isolation. These measures of control and confinement are affecting many industrial, social and cultural activities, which have suspended or significantly reduced their production and services, with future global effects yet to be seen.

For effective pandemic control and confinement, and for the planning phase of restart of our globalized World, it is important to be able to predict the evolution of the virus spread. To this aim, we apply learning theory for predicting the rate of infections following the pandemic outbreak. This theory, previously applied to outcome, accident and event data from multiple socio-technological systems (Duffey and Saull, 2008; Duffey & Ha, 2010; Fiondella & Duffey, 2015; Duffey, 2017a), postulates that humans learn from experience in correcting their mistakes and errors, as they gain knowledge on the problem and skill for addressing it. The theory is consistent with the models and data in cognitive psychology of how humans behave and the brain operates

(Ohlsson, 1996; Fiondella and Duffey, 2015; Anderson, 1990; Duffey, 2017b).

The theory is based on the fact that human learning demonstrably reduces error rates (Ohlsson, 1996): wisdom is gained after an accident. In simple words, humans learn from their mistakes and reduce outcomes in such a way that the rate of decrease of the outcome (in the present case of interest, the infection rate, R) with the rate of accumulated experience, ε (in the present case the advancement in the knowledge of the virus mechanisms, the contagion spreading dynamics, the effects of the countermeasures), is proportional to the rate R itself. Thus, very simply, the differential equation that describes the accident and outcome data with learning or forgetting describes the proportionality between the rate of change of the learning *rate*, R , and the learning rate itself (Duffey and Saull, 2008; Duffey, 2017b):

$$\frac{dR}{d\varepsilon} = -k(R - R_m) \quad (1)$$

where ε is the measure of the risk exposure, learning opportunity or experience/knowledge gained; k is the learning rate (positive for a learning/improving situation and negative for no learning/improving, e.g. because of no

effectiveness countermeasures) and R_m is the lowest or minimum achievable error rate, which is never zero as the process of error-making and cognitive rule revision always continues. Physically, k is related to the non-detection or error rate in unconscious memory scanning for recall and recognition, manifesting itself in the conscious external actions, decisions and judgments.

By integrating (1), the error rate evolution is:

$$R(\varepsilon) = R_m + (R_0 - R_m)e^{-k(\varepsilon - \varepsilon_0)} \quad (2)$$

where, R_0 is the initial rate at the beginning or start of the problem when the level of experience/knowledge on it is $\varepsilon = \varepsilon_0$. Different data sets are characterized by different values of the learning constant.

For data comparison, it is opportune to render non-dimensional the quantities of interest by considering the non-dimensional error rate $E^* = (R(\varepsilon) - R_m) / (R_0 - R_m)$ as a function of the non-dimensional experience/knowledge or risk exposure $\varepsilon^* = \varepsilon / \varepsilon_T$, where ε_T is the maximum accumulated experience or risk exposure thanks to which the error is recovered (the problem is considered under control), its rate having reached the lowest or minimum achievable value R_m . From (2),

$$E^* = \frac{R(\varepsilon) - R_m}{R_0 - R_m} = e^{-k\varepsilon^*} \quad (3)$$

This expression has already been shown to represent the learning trends for outcome rate data from industrial, surgical, transportation, mining, manufacturing, maintenance, and a multitude of other activities and systems (Duffey and Saull, 2008; Duffey and Ha, 2013; Fiondella and Duffey, 2015), and it is called Universal Learning Curve (ULC).

2. CoVid-19 infection rate prediction

To look at pandemic recovery, we look at the rate of infections. The risk of an infectious disease is not controlled unless the infection rate slows down, so-called ‘flattening the curve’. Once the rate peaks, then, it should decrease if countermeasures are successful (whatever they are). So, the next question is: what happens next and how do we know how effective the control

measures are, and how long before they can be relaxed or maintained?

Data from Johns Hopkins University website coronavirus.jhu.edu/map show that in China, the actual infection rate (increase per day) rose to a peak of nearly 4000 a day in about 17 to 20 days, and, then, fell away steadily to about 50 or so a day by another 30 days. As a further consideration for a different country, S Korea had a different peak rate of about 1000 per day in ten days, falling to a low rate in another 10 or so days. In Italy, after some delay in implementing countermeasures, the infection rate seemingly peaked at about 6000 per day in about 30 or so days. As of writing, (early April) this rate has decreased to nearly 4000 per day.

The situation for those mentioned countries that are currently showing some form of recovery are reported in Figure 1, which plots E^* , the non-dimensional infection rate normalized to the initial peak value, versus $N^* = \varepsilon / \varepsilon_T$, the non-dimensional elapsed time of experience/knowledge or risk exposure after the rate has peaked (number of days after peak/day of peak). In relation to the equations (1) and (2) of learning theory above, the infection rate takes the role of the error rate R , the risk exposure time, ε , corresponds to the accumulated experience/knowledge from which we learn and is measured in days, the time of peak ε_T , is the time for the rate to approach its achievable minimum value, R_m (the lowest or minimum achievable error rate, in equation (2)). Based on the available data for China and S Korea, using countermeasures, the overall recovery timescale is about 20-30 days to attain the minimum infection rate of about 50 per day.

For further direct comparison, we plot also the reduction curve of the world pulmonary disease death rate per day for 1870-1970 (Source: Human Nature, T.McKeown, April, 1978 as given in Horwitz and Ferleger "Statistics for Social Change"). We can simply think of this overall World data over the years, and its reduction trend, as resulting from many pandemics and multiple outbreaks of influenzas and differing virus strains, that have been more and more successfully treated as we have learned to better control/reduce

infections and improved effect recovery, thus steadily reducing the rate.

It is seen that despite the differences in timescales the evolution trend is simply the exponential ULC of equation (3) above, given by $E^* = \exp(-KN)$. The different datasets considered follow the same decreasing trajectory and, furthermore, the learning curve is nearly the same ($K \sim 3$). We claim that this trend decline due to the effectiveness of control measures of the pandemic is direct evidence of risk reduction and call it the Universal Recovery Curve.

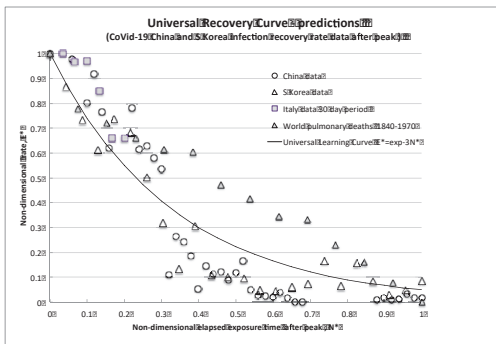


Figure 1 Predicted infections rate curve (URC, solid line) compared to various CoVid-19 infection data: China (data circles), S. Korea (open triangles), Italy (squares), based on about a 30-day timescale.

Note that Italy's trend is still evolving and has not yet reached its minimum, as for China and S. Korea. The infection reduction/recovery rate for Italy was estimated to occur in about 30 days, by assuming a timeframe similar to that of China. In other words, the prediction is that there is a 30-day recovery timeframe for Italy's infection rate to get down to a minimum achievable level of say 50 per day (as for China) and the current data match the prediction reasonably well. Furthermore, the prediction can be effectively monitored and should be updated as more data is collected. In fact, at present, only very few days have elapsed since the seemingly reached peak rate in Italy and sensitivity analysis can be used to verify it further, based on objective data, and also to control whether or not any change or improvement in countermeasures is warranted, and by how much. To do this, one can simply plot the data changing the denominator of the non-dimensional peak elapsed time N^* for estimated

reduction to the minimum (i.e. 50-100 per day), from the 30 days (27 March to 27 April) assumed in analogy to China, to say 20 days (27 March to 6 April) or 40 days (27 March to 6 May).

To further confirm the URC general theoretical correlation, we next compare to the projections for medical resource loads made by complex computer modeling of infections and deaths in the USA (IHME, 2020). As a reasonable surrogate measure, the number of required hospital beds has been assumed to be proportional to the number of infections, which daily values were directly transcribed from the website graph (available at covid19.healthdata.org/united-states-of-america). The interval available is a projection from a peak resource use on April 15th out to July 1st, 2020, so to be consistent with the actual available country data. The infection rate per day, R , was calculated until attaining an assumed but realistic minimum rate, R_m , of 50 per day on 10th June (55 days later). Figure 2 shows the results.

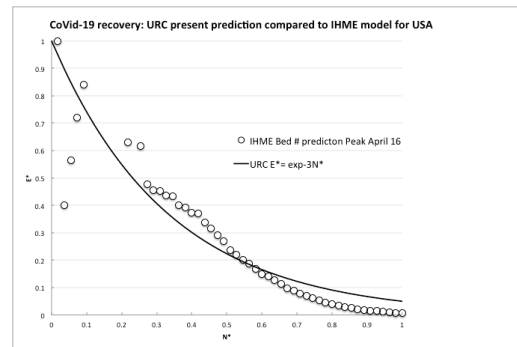


Figure 2 Comparison of URC to IHME model predictions of required hospital beds

3. Conclusions

In this paper, we have originally proposed Learning Theory for predicting the decline of CoVid-19 pandemic infections. A key point is to look at infection rate, as a measure of error outcome, and time, as a measure of experience/knowledge or risk exposure which allows learning. The analyses of the currently available data show that the CoVid-19 infection rate data follow, after peaking, almost exactly the Universal Learning Curve describing the decreasing trajectory of many other instances where humans learn to apply effective

countermeasures. More specifically, the learning curve is nearly the same (with universal constant $K \sim 3$) as for any learning experience reducing outcomes, accidents and events for other modern technological system operated by humans.

We claim that this trend decline due to learning is direct evidence of learning about risk reduction, also in this case of the pandemic, and call it the Universal Recovery Curve. It can be used to predict the expected time at which the pandemic will be under control, in terms of minimum achievable infection rate, and to test and demonstrate the relative effectiveness of the adopted countermeasures. As such, it is a fundamental tool for risk handling during the development of a pandemic.

From the publicly available data, China, Italy and S Korea are showing to have indeed learned how to control the spread of a viral pandemic. All other countries/systems/people have to do to predict the infection rate evolution is to check to follow the same trend after first reaching their rate peak.

A word of caution is necessary, however, these numbers cannot be exact and are not meant to be exact. These are just calculated risk estimates, which are subject to uncertainty related to all the many endogenous factors related to the virus spreading, and the actuation and respect of the measures implemented. The numbers provide guidance to thinking about the absolute risk and the best approach to take.

The strong message here is that the rational and logical approach to dealing with the risk of the occurring pandemic (as with any other risk, for that matter) is to limit own personal and potential exposure, and to minimize both the size and scale of the potentially exposed population. This is precisely what governments and contagious disease experts have been saying all along- but is also what any individual should be doing anyway while exposed to the risks of “normal” life. A sort of ethics of resilience (B. Rajaonah and E. Zio, 2020).

References

- Anderson, J.R., 1990, *Cognitive Psychology and its Implications*, W.H. Freeman, 3rd edition. ISBN 0-7167-2085-X.
- Duffey, R.B., 2017a, “Dynamic theory of losses in wars and conflicts”, *Eu. J. Op. Res.*, 261, pp1013-1027,
- Duffey, R.B., 2017b, “Search theory: the probabilistic physics and psychology of memory and learning”, *J.Sci. Res. and Studies*, Vol 4, No 5, pp107-120.
- Duffey, R.B., 2017c, “Submarine Warfare and Intelligence in the Atlantic and Pacific in the Second World War: Comparison and Lessons Learned for Two Opponents”, *J. Maritime Res.*, Vol 19, No 2, pp143-167,
- Duffey, R.B. and Ha, T., 2010, “Human reliability: benchmark and prediction”, *Proc. Part O: J. Risk and Reliability*, *IMechE* Vol. 224, p 185-190.
- Duffey, R.B. and Saull, J.W., 2008, *Managing Risk: the human element*, J. Wiley and sons, UK, ISBN 978-0-470-69976-8.
- Fiondella, L. and Duffey, R.B., 2015, “Software and Human Reliability: error reduction and prediction”, *Proc. PSA 2015*, paper 12221, *Am. Nucl. Soc.*, April 21-26, Sun Valley, Idaho.
- IHME COVID-19 health service utilization forecasting team. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator days and deaths by US state in the next 4 months. *MedRxiv*. 26 March 2020. doi:10.1101/2020.03.27.20043752.
- Ohlsson, 1996, “Learning from performance errors”, *Psychological Review*, 103 (2):241-262.
- Rajaonah B., Zio E.. *Contributing to Disaster Management as an Individual Member of a*

Collectivity: Resilient Ethics and Ethics of
Resilience. 2020. hal-02533290.