Coronavirus

Modelling behind lockdown was an unreliable buggy mess, claim experts

Data that predicted 500,000 could die in UK unless extreme measures were taken are impossible to replicate, say scientific teams

Science

By **Hannah Boland**
and **Ellie Zolfagharifard**

THE Covid-19 modelling that sent Britain into lockdown, shutting the economy and leaving millions out of work has been criticised by experts.

Prof Neil Ferguson's Imperial College computer coding was derided as "totally unreliable" by leading figures, who warned it was "something you wouldn't stake your life on".

The model, credited with forcing the Government to U-turn and introduce a nationwide lockdown, is a "buggy mess, which looks more like a bowl of angel hair pasta than a finely tuned piece of programming", said David Richards, the co-founder of British data technology company WANdisco.

"In our commercial reality, we would fire anyone for developing code like this and any business that relied on it to produce software for sale would likely go bust."

The comments are likely to reignite a row over whether the UK was right to go into lockdown, with conflicting models suggesting people may have already acquired substantial herd immunity and Covid-19 may have hit Britain earlier than first thought.

Scientists have also been split on the fatality rate of Covid-19, which has resulted in vastly different models.

Up until now, significant weight has been attached to Imperial's model, which placed the fatality rate higher than others and predicted 510,000 in the UK could die without a lockdown.

It was said to have prompted a dramatic change in government policy, causing businesses, schools and restaurants to be shut immediately in March. The Bank of England has predicted that the economy could take a year to return to normal, after its worst recession in more than three centuries.

The Imperial model works by using code to simulate transport links, population size, social networks and healthcare provisions to predict how coronavirus would spread. However, questions have emerged over whether the model is accurate, after researchers released its code, which in its original form was "thousands of lines" developed over more than 13 years.

In its initial form the code was unreadable, developers claimed, with some parts looking "like they were machine translated from Fortran", an old coding language, according to John Carmack, a US developer, who helped clean the code before it was published.

Yet, the problems appear to go much deeper than messy coding. Many have claimed that it is almost impossible to reproduce the same results from the same data, using the same code.

Scientists from the University of Edinburgh said they got different results when they used different machines, and even in some cases using the same machines. "There appears to be a bug in either the creation or re-use of the network file. If we attempt two completely identical runs, only varying in that the second should use the network file produced by the first, the results are quite different," the Edinburgh researchers wrote on the Github online forum. After a discussion with a Github developer, a fix was provided.

It is said to be one of a number of bugs discovered within the system. Github developers said that the model was "stochastic" (random), and "multiple runs with different seeds should be undertaken to see average behaviour".

It has prompted questions from specialists, who say "models must be capable of passing the basic scientific test of producing the same results given the same initial set of parameters ... otherwise, there is simply no way of knowing whether they will be reliable." It comes amid a wider debate over whether the Government should have relied more heavily on numerous models before making policy decisions.

Writing for the *Telegraph* online, Prof Sir Nigel Shadbolt, the principal of Jesus College, Oxford, and chairman of the Open Data Institute – which he co-founded with World Wide Web inventor Sir Tim Berners-Lee – said: "Having a diverse variety of models, particularly those that enable policymakers to explore predictions under different assumptions, and with different interventions, is incredibly powerful."

Like the Imperial code, a rival model by Prof Sunetra Gupta at the University of Oxford works on a so-called "SIR approach" in which the population is divided into those that are susceptible, infected and recorded. However, while Prof Gupta assumed that 0.1 per cent of infected people would die, Prof Ferguson worked on 0.9 per cent. That led to a dramatic reversal in government policy from attempting to build "herd immunity" to a full-on lockdown.

Concerns over Prof Ferguson's model have been raised, with Dr Konstantin Boudnik, the VP of architecture at WANdisco, saying his track record did not inspire confidence. In the early 2000s, Prof Ferguson's models incorrectly predicted up to 136,000 mad cow disease deaths, 200million from bird flu and 65,000 from swine flu.

"The facts from the early 2000s are just yet another confirmation that their modelling approach was flawed to the core," says Dr Boudnik. "We don't know for sure if the same model/code was used, but we clearly see their methodology wasn't rigorous then and surely hasn't improved now."

A spokesman for Imperial's Covid-19 team said: "The Government has never relied on a single disease model to inform decision-making. As has been repeatedly stated, decision-making around lockdown was based on a consensus view of the scientific evidence, including several modelling studies by different academic groups.

"Multiple groups using different models concluded that the pandemic would overwhelm the NHS and cause unacceptably high mortality in the absence of extreme social distancing measures. Within the Imperial team, we use several models of differing levels of complexity, all of which produce consistent results. We are working with ... legitimate academic groups and technology companies to develop, test and further document the simulation code referred to. However, we reject the partisan reviews of a few clearly ideologically motivated commentators.

"Epidemiology is not a branch of computer science and the conclusions around lockdown rely not on any mathematical model but on the scientific consensus that Covid-19 is highly transmissible with an infection fatality ratio exceeding 0.5 per cent in the UK."

'In our commercial reality, we would fire anyone for developing code like this'

'Any business that relied on it to produce software for sale would likely go bust'

'It looks more like a bowl of angel hair pasta than a finely tuned piece of programming'

'The early 2000s were yet another confirmation that their modelling approach was flawed to the core'

Imperial's programming could go down as the most devastating software mistake of all time

Opinion

By **David Richards**
and **Konstantin Boudnik**

In the history of software mistakes, Mariner 1 was probably the most notorious. The unmanned spacecraft was destroyed seconds after launch from Cape Canaveral in 1962 when it veered dangerously off-course due to a line of dodgy code. But nobody died and the only hits were to Nasa's budget and pride.

Imperial College's modelling of non-pharmaceutical interventions for Covid-19 which helped persuade the UK and other countries to bring in draconian lockdowns could supersede the failed Venus space probe to go down in history as the most devastating software mistake of all time, in terms of economic costs and lives lost.

Since publication of Imperial's microsimulation model, those of us with a professional and personal interest in software development have studied the code on which policymakers based their fateful decision to mothball our multitrillion-pound economy and plunge millions of people into poverty and hardship.

And we were profoundly disturbed at what we discovered. The model appears to be totally unreliable and you wouldn't stake your life on it.

First though, a few words on our credentials.

I am David Richards, the founder and chief executive of WANdisco, a global leader in Big Data software which is jointly headquartered in Silicon Valley and Sheffield.

My co-author is Dr Konstantin "Cos" Boudnik, the VP of architecture at WANdisco, and author of 17 US patents in distributed computing and a veteran developer of the Apache Hadoop framework that allows computers to solve problems using vast amounts of data.

Imperial's model appears to be based on a programming language called Fortran, which was old news 20 years ago and – guess what? – was the code used for Mariner 1.

This outdated language contains inherent problems with its grammar and the way it assigns values, which can give way to multiple design flaws and numerical inaccuracies.

One file alone in the Imperial model contained 15,000 lines of code. Try unravelling that tangled, buggy mess, which looks more like a bowl of angel hair pasta than a finely tuned piece of programming. Industry best practice would have 500 separate files instead.

In our commercial reality, we would fire anyone for developing code like this and any business that relied on it to produce software for sale would likely go bust. The approach ignores widely accepted computer science principles known as "separation of concerns", which date back to the early Seventies and are essential to the design and architecture of successful software systems.

The principles guard against what developers call CACE: Changing Anything Changes Everything.

Without this separation, it is impossible to carry out rigorous testing of individual parts to ensure full working order of the whole.

Testing allows for guarantees. It is what you do on a conveyer belt in a car factory. Each and every component is tested for integrity in order to pass strict quality controls. Only then is the car deemed safe to go on the road.

We were disturbed. The model appears to be totally unreliable and you wouldn't stake your life on it

As a result, Imperial's model is vulnerable to producing wildly different and conflicting outputs based on the same initial set of parameters.

Run it on different computers and you would likely get different results. In other words, it is non-deterministic. As such, it is fundamentally unreliable. It screams the question as to why our government did not get a second opinion before swallowing Imperial's prescription.

Ultimately, this is a computer science problem and where are the computer scientists in the room? Our leaders did not have the grounding in computer science to challenge the ideas and so were susceptible to the academics.

I suspect the Government saw what was happening in Italy with its overwhelmed hospitals and panicked. It chose a blunt instrument instead of a scalpel and now there is going to be a huge strain on society.

Defenders of the Imperial model argue that because the problem – a global pandemic – is dynamic, then the solution should share the same stochastic, non-deterministic quality. We disagree.

Models must be capable of passing the basic scientific test of producing the same results given the same initial set of parameters. Otherwise, there is simply no way of knowing whether they will be reliable.

Indeed, many global industries successfully use deterministic models that factor in randomness.

No surgeon would put a pacemaker into a cardiac patient knowing it was based on such an arguably unpredictable approach for fear of jeopardising the Hippocratic oath.

Why on earth would the Government place its trust in the same when the entire wellbeing of our nation is at stake?

David Richards is founder and chief executive of WANdisco and Dr Konstantin Boudnik is the company's vice-president of architecture

Tread carefully when predicting fatality rates

Comment

By **Prof Sunetra Gupta**

We experience the world through our senses, and strive to build our understanding by constructing some kind of replica in our minds of what is actually out there.

Mathematical modelling is an extension of this fundamental activity. The Covid-19 pandemic has, at its heart, a very simple dynamic which can be captured by the fundamental SIR model, known as a powerful tool for dissecting the spread of diseases which have short "infectious" periods and where "recovery" renders an individual unavailable for infection.

Members of a population can be allocated to the "I" and "R" bins, with the remainder being left in the "susceptible" bin; hence SIR.

The behaviour of these SIR systems is fully specified by two factors: how long an individual is infectious, and the fundamental transmission potential of the pathogen. Fortunately, both of these can be estimated at a fairly early stage of the epidemic. What is harder to estimate at the early stage is the infection fatality rate.

Two extreme scenarios fit the data. In one, as assumed by the Imperial College modellers, the fatality rate is in the ball park of 1pc, which would imply the epidemic took off towards the end of February in the UK and less than 10pc of the population had been exposed by time of lockdown. In the other, the fatality rate could be as low as 0.01pc. This would require the virus to have been introduced about a month earlier and for over 50pc of individuals to be exposed by March 23.

In essence, the Covid-19 situation is one that can be served by the canonical SIR model. Elaborate computer models could be used to simulate specific activities such as standing two metres apart in a queue but this is contingent on the validity of the underlying assumptions.

Simple models induce you to consider a range of possibilities of the key parameters, while complicated models tend to "fit" parameters to the available data. Most modellers have fitted their models to very limited data available at the time of lockdown on cases and deaths as reported. This corresponds to a particular solution of the general SIR model with a high fatality rate – the worst-case scenario.

But how plausible is a fatality rate of 1pc? A more robust approach would be to measure the proportion already exposed to the virus. The rapid refinement of antibody detection methods over the last couple of months will enable this, allowing us to determine the true fatality rate.

Sunetra Gupta is a professor of theoretical epidemiology at the University of Oxford

Funeral firms told to expect up to 85,000 extra deaths in summer

Analysis

By **Steve Bird**

FUNERAL directors have been told to expect anywhere between 33,000 and 85,000 extra deaths by the end of the September, according to new analysis.

It emerged after the National Association of Funeral Directors (NAFD) commissioned research to predict what effect easing restrictions could have on the country's mortality rate.

The modelling also found even if the Government controls the spread of the virus at a national level some regions could experience a "significant second peak". But if opening up is not managed well, that second spike could be worse than the first in some areas. Andrew Lilico, of Europe Economics, who led the research, predicted what could happen if the lockdown was firstly successful – where those with Covid-19 infected fewer than one person each.

He then analysed what could transpire if easing restrictions was less effective and the national rate of infection rose to 1.5 people being infected by each carrier.

"The modelling suggests that 85,000 people could die post-lockdown if the Government fails to keep the rate of infection below one," Mr Lilico told *The Sunday Telegraph*. "But if lifting lockdown was successful, then about 33,000 people could still die."

He analysed published data and forecasts from Oxford University, Imperial College London and the London School of Hygiene & Tropical Medicine to predict how coronavirus behaves, its infection fatality rate, as well as the effect herd immunity and antibodies have on protecting the public.

Jon Levett, NAFD chief executive, said: "What is striking about the modelling is that different parts of the UK and different regions within England would see significant second peaks with even a successful lifting of lockdown – most notably Scotland, Wales, and the North East and Yorkshire.

"If lockdown is lifted less successfully, the impact is even more striking, with the potential for those regions to see a second peak higher than the first. The North West and South West would also be significantly affected by a less than successful lifting of lockdown, while London, the East of England, Midlands and Northern Ireland are predicted to have a far lower impact."

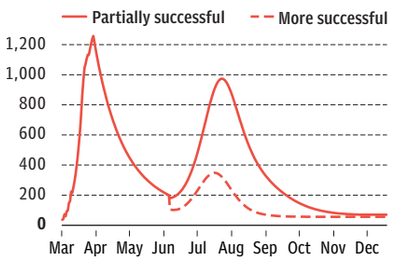
The association has repeatedly asked the Government to release its modelling data on deaths so it can better prepare for the demand for funerals.

However, ministers have not provided that data, despite funeral directors saying they are struggling to cope with the recent rise in funerals.

Mr Lilico added: "We hope the models will assist funeral directors in making their business decision and in providing the best service to bereaved people over the next few months."

Fatality forecasts as lockdown is lifted

Evolution of daily COVID-19 deaths across the UK if the government is successful in keeping R under control post-lockdown



SOURCE: NATIONAL ASSOCIATION OF FUNERAL DIRECTORS AND EUROPE ECONOMICS

Total deaths in England, all causes