Stop Flaunting Those Curves! Time for Stats to Get Down and Dirty With the Public

Timandra Harkness¹

¹BBC Radio, United Kingdom

Published on: Aug 03, 2020

DOI: <https://doi.org/10.1162/99608f92.caab8ba0>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Column Editor's Note: *Some level of numeracy, and especially of understanding of data, is increasingly important in the modern world. This has been brought home with painful urgency through the Covid-19 pandemic. Leading science communicator Timandra Harkness points out the benefits, and limitations, of communicating understanding of what data are telling us through the medium of graphs.*

Keywords: graphs, communicating statistics, mathematics, COVID-19

In the 1990s, a UK television drama series about a forensic pathologist had an unexpected effect. As *Silent Witness*, followed by *CSI: Miami* and its ilk, drew millions of viewers, British universities noticed increased demand from teenagers aspiring to solve crimes in white coats. The number of Forensic Science degree courses rose from 2 to 285 in under 20 years, recruiting 1,667 students in 2008 alone. *Guardian* columnist Tim Dowling (2009) commented at the time, “In order to ensure there are enough jobs to go round, more than half of them will have to retrain as serial killers.”

As today's teenagers, their schools closed, watch the news coverage of COVID-19, they must be absorbing the importance of logarithmic scales, projected curves, and R numbers for saving real, not fictional, lives. Let's hope some of them go on to study statistics and data science when education sputters back to life.

The Open University has made its [Medical Statistics course](#) (OpenLearn, 2019) free to access, but that's designed for people at (roughly) second-year undergraduate Maths level. They should be preparing to market their whole Mathematics and Statistics degree course, not only to teenagers but to their parents and grandparents. If I can gain their BSc in seven years of part-time study, so can anyone else.

Who would have thought, a year ago, that rates of increase, the reliability of data, and the failings of computer models, would become the stuff of heated arguments in whatever (real or digital) forum replaced the pub or coffee shop? For those of us who want to nurture wider public interest, understanding, and confidence in statistics and data, the problem has suddenly shifted. Nearly everyone is interested, now. The challenge is to turn that interest into better understanding and more confidence.

“One often laments that the general population does not have the time or resources to understand the data science behind social scientific concepts,” note Podkul et al. (2020), “this is a rare scenario in which most now have both.” They used the opportunity to test how well the public understands one key statistical idea. “[The Coronavirus Exponential: A Preliminary Investigation into the Public's Understanding](#)” (Podkul et al., 2020) found that most people could grasp the scale of exponential growth with visual aids and a few multiple-choice options, but struggled without.

The BBC's [More or Less](#) radio series (and podcast)(2020) are doing excellent work unpacking the deluge of numbers in the news for a lay audience. Professor Sir David Spiegelhalter is leading a vanguard of statisticians

in the UK media helping the public make sense of vital information. But there are a few barriers to better public understanding of statistics.

One barrier is mathematics. Mathematics is hard. Like science, it's often counterintuitive, which is why it helps us toward an understanding of the world beyond our senses, our instincts, and our everyday experience. The popular idea of a mathematician is somebody for whom mathematics is easy, but even mathematicians admit it is a more difficult language than the language you're reading now. It may be *less* difficult for a mathematician, but most of them will say the important difference is that they enjoy the difficulty.

Beyond being hard, mathematics often holds a particular terror for people who enjoyed it at school until a certain point, and then stopped understanding and got left behind, abandoned in a dark forest of bafflement. This terror, this assumption that numbers are an impenetrable mystery for the masses, lends an aura of authority to statistics and data that is not helpful.

Why did UK government briefings include a daily diet of graphs? Most of them were never explained, their limitations, assumptions, and context not spelled out. Scientific and medical advisors pointed at them to illustrate a core message, to drive home the point that cases are rising, that the situation is better or worse than other countries, that we are driving less or testing more. The graphs lent authority to official instructions.

Their desired effects may be elusive. Podkul et al. (2020) report that their research results “fail to demonstrate any clear connection between being able to understand exponential patterns with being worried about disease outbreak.” There's no simple, causal relationship between statistical understanding and emotional or behavioral response.

Now, I like a graph as much as...OK, I like a graph 37% more than the next person. So, I am also quick to show a visual aid that readily says “cases still rising, but not as fast” or “this disease is WAY more dangerous for the old than the young.” And one thing I would love to do is equip people to read graphs and plots more fluently, and more critically.

Little or no mathematical knowledge is needed to read a horizontal axis as time passing, while some other variable increases or decreases in level. Years of live performing have taught me that humans intuitively read space as a proxy for other variables. Simply by standing in different positions on a stage, I can get an audience to read them as past, present, and future (pro tip—this is easier if they can read past-to-future as *their* left-to-right).

Similarly, we are used to judging relative sizes, so a bar chart using a y-axis that includes zero makes sense to most lay audiences. This column is twice as tall as the previous one, so that category has twice as much of whatever it is.

But a histogram can easily mislead if it truncates the y-axis, or if it divides continuous data into unequal brackets, then displays them as equal width bars. Comparing the number of deaths and ICU admissions in the age groups 0–19 (20 years), 20–44 (25 years), and 45–54 (10 years) is not comparing like population sizes with like, yet that’s exactly what a graph published by the Centers for Disease Control and Prevention (CDC) does (2020).

Figure 1. Coronavirus disease 2019 (COVID-19) hospitalizations, intensive care unit (ICU) admissions, and deaths, by age group—United States, February 12–March 16, 2020. [Graphic published by the CDC](#), March 27, 2020.

A quick glance would even suggest that 20–44-year-olds face a *higher* risk of hospitalization than older people, which is simply not true.

Again, it takes almost no mathematics to spot this misleading flaw, but many people feel unqualified to ask these basic questions. And this undue reverence for numbers and those who use them is another barrier, not only to public understanding of statistics, but to participation in democracy.

Greater use of data by governments is to be welcomed when it informs action that is more effective toward policy goals, monitors the success and failure of those actions, and helps hold politicians to account for their promises and policies.

But recent experience shows that data and statistics are also used to convey an impression of scientific certainty where none exists. In the words of Spiegelhalter, statistics in government briefings were “[number theatre](#)” (Spiegelhalter, 2020), deployed not for deeper understanding but for dramatic effect.

Some later examples from UK government briefings (2020) have the form of a graph but no mathematical underpinning at all, such as the [curve](#) that seemed to show stages of relaxing lockdown rules as the R number decreased, but with no axis labels.

Figure 2. Steps of adjustment to current social distancing measures.
[Graphic published by UK Government](#), May 11, 2020.

There was even an ‘equation’ that contained an equal sign but made no sense either mathematically or in any other terms. Apparently COVID Alert Level = R (rate of infection) + Number of infections. For anyone who knows the difference between addition and multiplication it should come with a disturbing content warning.

These visuals are clearly aimed neither to harness nor to encourage greater public understanding of statistics or anything else. They simply invoke the authority of mathematics and data as a shield against doubt.

Perhaps the most insidious barrier against greater public understanding of statistics is the quest for certainty in an uncertain world. The desire to know the future, and to feel that somebody is in control, is so strong that even a terrible future in a world controlled by evil conspiracies holds strong appeal against rudderless and unpredictable reality.

But statistics doesn’t offer certainty, you cry. Statistics is all about engaging with *uncertainty*, trying to tease apart what we know, what we can reasonably infer, and what we must accept as unknowable, at least for now.

That is exactly why it’s so hard to turn this upsurge in public interest into genuine engagement and deeper understanding. Anyone should be able to read graphs, and to ask questions about how data was collected, what was left out, and what assumptions underlie the models used to predict the future. That should be as much a part of being an engaged citizen as knowing what policies each party stands for.

But the more the public gets stuck in asking awkward questions about the numbers, the clearer it will become that statistics and data are not gods, but tools for fallible humans making sense of the world. Computer models

are not oracles, but machines for imagining the world in more detail than one human brain can hold.

When people stop worshiping statistics from afar, awestruck by their ineffable mystery, and get close enough to smell the sweat and see the pimples, some disillusionment will set in. But only then can a real relationship start to happen, one where the public feels free not only to understand statistics but to challenge them. I hope we're ready for that.

Disclosure Statement

Timandra Harkness has no financial or non-financial disclosures to share for this article.

References

- Centers for Disease Control and Prevention. (2020). Severe outcomes among patients with coronavirus disease 2019 (COVID-19)—United States, February 12–March 16, 2020. *MMWR Morbidity and Weekly Mortality Report*, 69, 343–346. <http://doi.org/10.15585/mmwr.mm6912e2>
- Dowling, T. (2009, October 15). The grisly truth about CSI degrees. *The Guardian*. <https://www.theguardian.com/education/2009/oct/15/csi-effect-forensic-science>
- HM Government. (2020). *Steps of adjustment to current social distancing measures*. GOV.UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/884352/slides_-_11_05_2020.pdf
- Spiegelhalter, D. (2020). [Video]. *The Andrew Marr show*. BBC. <https://www.youtube.com/watch?v=QIw2l-trRXc>
- More or Less*. (2020). Harford, T. (Presenter.), & Vadon, R. (Ed.). Behind the stats [Audio podcast]. BBC. <https://www.bbc.co.uk/programmes/p02nrss1/episodes/downloads>
- OpenLearn. (2019). *Medical statistics*. <https://www.open.edu/openlearn/science-maths-technology/medical-statistics>
- Podkul, A., Vittert, L., Tranter, S., & Alduncin, A. (2020). The coronavirus exponential: A preliminary investigation into the public's understanding. *Harvard Data Science Review*, (Special Issue 1). <https://doi.org/10.1162/99608f92.fec69745>
-

©2020 Timandra Harkness. This article is licensed under a Creative Commons Attribution (CC BY 4.0) [International license](https://creativecommons.org/licenses/by/4.0/), except where otherwise indicated with respect to particular material included in the article.