REPORTING FROM A STATISTICAL CHAOS:
JOURNALISTIC LESSONS FROM THE FIRST YEAR OF COVID-19 DATA AND SCIENCE IN THE NEWS

An Nguyen
Xin Zhao
Brendan Lawson
Dan Jackson

Contents

Introduction ***1***

The chaos of Covid-19 data and statistics ***2***

Challenges to journalism ***4***

The rise to challenges ***6***

A positive long-term legacy? ***12***

Report's authors ***14***

Acknowledgements ***14***

Presenters and Respondents at the Symposium ***15***

This report is based primarily on presentations and discussions at the symposium, *Coronavirus, Statistical Chaos and the News: Preliminary Reflections from Journalists and Scholars*, on 4 December, 2020.

The one-day symposium was a joint initiative of Bournemouth University (host), the Royal Statistical Society and the Association of British Science Writers.

All video recordings of the day are hosted on the symposium's YouTube channel at https://www.youtube.com/channel/UCC8f6WkyIFeYHAiuEltww5w.

Date of publication: January 2021

Introduction

Covid-19 has brought data and statistics to the centre of daily life like never before: everything we do in response to the virus at individual, organisational and societal levels depends literally on what the numbers tell us. In March 2020, as the exponential virus loomed large into an existential threat, an influx of numbers that would normally stay within the domain of specialist expertise suddenly occupied the physical and cultural space of the lockdown family. "Scary" concepts – R-naught, infection rate, transmission rate, death rate, excess deaths, false positive, false negative, relative risk, absolute risk, random sampling, statistical modelling and so on – abounded everywhere, from the TV and computer screens to husband-wife or even parent-child musings.

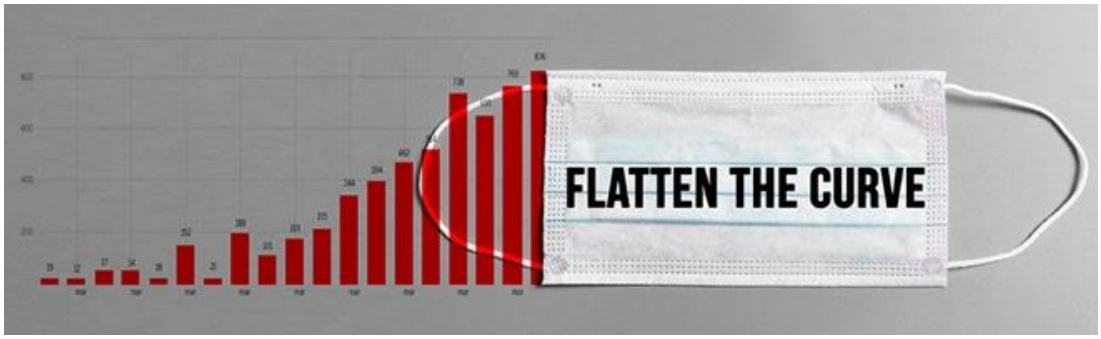
With that came a statistical chaos that continues, albeit to a lesser extent, today. Associated with such numbers is an unprecedented level of complexity and uncertainty due to the novelty of the virus. But as they become so crucial, Covid-19 data have been subject to a rather fierce battle between different frames and narratives, in which scientists dispute with each other as well as compete – not always successfully – with religion, culture and, most importantly, politics. Amidst much public confusion, anxiety and fear, numerical misinformation and disinformation seem to be everywhere on social media.

What does all this mean to journalism, a profession that is rarely commended for their ability to engage and deal with numbers? What are the major challenges? How have journalists performed in in questioning, scrutinising and communicating Covid-19 data, including debunking statistical "lies and damn lies"? What methods, techniques and platforms do they use to obtain, unpack, portray and deliver Covid-19 data and statistics to help people make sense of the pandemic? To what extent can their work change hearts or alter minds? What can the media learn from the first year of news reporting of the virus?

On 4 December, 2020, twenty-four senior science, health and data journalists, scientists, statisticians and media scholars gathered for to reflect on the above at a one-day conference jointly organised by Bournemouth University, the Royal Statistical Society and the Association of British Science Writers. The event, called Coronavirus, Statistical Chaos and the News Symposium, was a platform for journalists and scholars to share ideas and exchange experience around the above questions. This report summarises key issues, challenges and lessons learnt from the first year of British news reporting of the pandemic.

The chaos of Covid-19 data and statistics

The stakes were high in the last week of March 2020, when the UK went into the first lockdown and, according to Ofcom research, virtually every UK adult (99%) accessed news about Covid-19 at least once a day. The maximal thirst for answers of a fear-stricken public, however, met with a minimal understanding of scientists about the novel virus, its symptoms, transmission pattern, vulnerable hosts, incubation periods and so on. As Ann Hemingway, Professor of Public Health and Wellbeing at Bournemouth University, observed, scientists became "experts with no evidence," advising politicians and people primarily on the basis of their knowledge of previous viruses.



Science had to compete – not always successfully – with non-science forces, especially politics, during the pandemic (Photograph: Shutterstock)

The urgency led to an unprecedented number of online preprints (research papers that have not been peer-reviewed), with contradictory conclusions and wide-ranging qualities (see Box 1). Quite a few "armchair epidemiologists" would stray far beyond their comfort zone to speculate about the virus and its disease. As scientists contested each other with diverse views on what was happening and what would be happening next, science was somewhat relegated to the status of opinionated debate.

To make matters worse, Covid-19 data have been subject to a rather fierce battle between different frames and narratives, in which science had to compete – not always successfully – with religion, culture and, most importantly, politics. In the midst of confusion, anxieties and fears, the public found mis/disinformation not only on social media but also, and rather deplorably, in press briefings and interviews by ministers and MPs.

The government keeps saying they following the science, but there is not such a thing as *the* science.
- Kevin McConway -

Sir David Spiegelhalter, Chair of the Winton Centre for Risk and Evidence Communication at Cambridge University and Co-chair of RSS Covid-19 Taskforce, recalled telling an Andrew Marr show in May that the daily press briefings at No

Box 1. "Preprints, preprints everywhere"

Kevin McConway's keynote speech highlighted the challenges that pervasive research preprints about the pandemic have posed to journalists. First, preprints received much more attention during the pandemic than normal times because they are professionally press-released, shared and/or discussed about on a large scale on blogs and social platforms. Second, although there are some very good studies that are done rapidly, there are also "terrible rubbish." For journalists under pressure, it is quite a daunting task to tell the quality of preprints and to decide what information about the newly released studies to include in their reporting.

McConway cited the example of "An analysis of SARS-CoV-2 viral load by patient age", a preprint by a group of renowned virologists in Germany that used statistical findings to make policy recommendations on the re-opening of schools and kindergartens. McConway and Spiegelhalter found important statistical flaws in the preprint and wrote immediately about it in a blog on Medium. It caused quite a stir on the German tabloid newspaper *Bild* and on Twitter, with quite diverse and polarised debates. Although the Bild intervention ended well with the preprint being revised and improved, McConway pointed to a larger issue: What if that preprint was not picked up by any academic peer like himself and Spiegelhalter? And what if the journalists covering this topic had not known whose information should be trusted in evaluating the preprint – the authors, the critics, or some others' voices?

10 were a "completely embarrassing […] number theatre". In his keynote, Kevin McConway, Emeritus Professor of Applied Statistics at Open University and a frequent statistical news commentator, added "chart cinema" to that theatre metaphor.

The "drama" comes from two main sources. One is *the poor communication of uncertainty by politicians.* McConway referred to an opinion piece in the British Medical Journal stating that the UK government communications about major policies, such as the "Guidance on social distancing for everyone in the UK" in March 2020, did not involve the messages of uncertainties of evidence. "The government keeps saying they follow the science, but there is not such a thing as *the* science," said McConway. Such practices can add another layer of difficulty for journalists, especially those with no science, health and data reporting background, forcing them to spend much more efforts to judge the validity and quality of the information they were provided with.

The other source is the way in which politicians manipulate data to make their point. In opening the symposium, An Nguyen brought into the Covid-19 context what Tim Harford called "statistical bullshit" – or "the casual slinging around of numbers [by politicians] not because they are true, or false, but to sell a message". McConway pointed to Lawson and Lovatt's study which found that politicians, rather than treating data as being objective or value-free, can treat statistics as a rhetorical device through selective, contestable, and strategic ways of interpretations.

Radio 4 science presenter and RSS fellow Timandra Harkness was one of the most infuriated about the weaponization of Covid-19 data. She cited the crucial November press conference in which misleading visuals were used to justify the second lockdown. Even though the verbal messaging contained some nuance, the visuals – which represented scenarios, not forecasts – were too striking to tell their own story, especially to the lay viewer. However good the intention is, Harkness said, the use of statistics to invoke emotions of fear or anger to influence public behaviour is very questionable. "It risks eroding the usefulness of data as information, and public trust in statistics as a reliable tool for understanding".

Challenges to journalism

All that constitutes an unsettling reporting situation that journalists have probably never faced before the pandemic. Ross Lydall, the *Evening Standard*'s Health Editor, called it "the most complex story of our careers".

On one hand is an epistemic and cultural challenge. There is no specific number to pin down to tell the story, which goes somewhat against the journalistic mindset. As pioneer data journalist Paul Bradshaw noted, "journalists are trained and culturally set up to prefer the specific and concrete", not to deal with the abstraction and vagueness that characterise Covid-19 uncertainty.

We just don't have a United Kingdom of Covid-19 data.

- Pamela Duncan -

On the other is a series of practical problems. Amidst uncertainty, there is a myriad of competing concepts and measurements that journalists have to face and handle. To depict the virus's progress, for instance, should they use the number of new cases, the number of newly reported deaths, the R-number or something else? And what is a "Covid death" anyway? There are at least three considerably different sources of daily death counts by DHSC, NHS England, and the Office of National Statistics. Even the R-number could vary according the chosen calculation method. "We just don't have a United Kingdom of Covid-19 data," said Pamela Duncan, data journalist at Guardian News and Media. "There are a lot of holes in data that we are still struggling to grapple with," Duncan said.

In that context, journalism – a profession that is not only "allergic to numbers" (in Duncan's words) but also has seen many specialist science, health and data reporting roles slashed under immense economic pressures in the past decades – was thrown into the deep end. Jane Kirby remembered very well the early days of dealing with confusing data and "floundering about, trying to work out what the data meant and what on earth to do with them". The heightened importance of data-based stories makes matters worse by imposing substantial pressure on journalists to deal with data and their sources as quickly as possible. For afternoon newspapers like the *Evening Standard*, the pressure to "get the morning news on the street quickly" was permanent, sometimes with only "minutes to work out what the data means", said Ross Lydall, its Health Editor, who cited its copy deadline is 10 am. "How do you make the right call when you don't have much time?" is the constant question that his team would contemplate, especially when there was such uncertainty in the data.

Access to data is another issue, especially at local levels. Claire Miller, Head of Data Journalism at Reach PLC, explained the many problems her team faced, including the unavailability of Google Mobility data in the early days and the fact that the lack certain datasets – such as care home data from local authorities – were initially recorded and then stopped being published. In some cases, data are not made available for journalists to vet politicians' messages. Miller mentioned the case of the stalemate debate over tier

restrictions in Greater Manchester in October, in which government ministers and local leaders such as Andy Burnham cited all kinds of hospital figures to make arguments. But nobody on either side was willing to release hospital-level data for public scrutiny. Reach's *Manchester Evening News* made a series of adamant attempts to collect the data, including direct requests to NHS England, Greater Manchester's Health and Social Care Partnership, all seven hospital trusts in the conurbation, DHSC, the mayor's office and No 10. They ended up with very little more than generic, irrelevant data from only a minority of those who received the requests. Some hospital trusts, they found, were instructed not to provide data to the media. Facing no other option, they decided to tell the readers in an investigation piece that concludes as follows:

> People here are watching on in anxiety, wondering what will happen and what this winter will really mean for them and the NHS. In the meantime, the war of words over what happens next to Greater Manchester continues, while each part of the English system sits on the numbers behind their arguments.

As they turn to rely on scientists to triangulate facts and figures, securing access to the right expert sources has proved to be tricky at times. In some cases, sources are too busy dealing with a huge amount of work, including unprecedented numbers of media enquiries. Jane Kirby, Health Editor at PA Media, remembered a situation when Oxford University and AstraZeneca released their Phase III findings about their vaccine via a press release. She had some confusion regarding the two numbers regarding efficacy: 70% (which was contextualised with clinical data) and 90% (no underpinning clinical data). Under the pressure to file the story, she conservatively opted for the 70% figure with contextualised data, although her mind was occupied by the other, much more impressive, and hence more newsworthy, 90% efficacy rate. It was "a few uncomfortable hours" before she could get hold of someone at Oxford to gather more data and then publish that 90% figure and its basis.

With the increasing involvement of political journalists, there was the real danger that some top scientists would turn down interviews for the fear of being pitched against each other at the expense of scientific nuance.

- Fiona Lethbridge -

In other cases, as the pandemic was increasingly politicised, some with the genuine expertise started to be reluctant to participate in media discussions. "With the increasing involvement of political journalists, there was the real danger that some top scientists would turn down interviews for the fear of being pitched against each other at the expense of scientific nuance," said Fiona Lethbridge, Senior Press Officer of the Science Media Centre, which exists to act a bridge between journalism and experts.

Related to that is the fact that internal newsroom politics, particularly the power relationship between science, health and data journalists and those in politics and other more popular news beats, came to the fore in the early days. As newsworthiness still

dictates journalists' work, said Lethbridge, there was tension over things such as what types of Covid-19 stories are to be prioritised. In a webinar in June 2020, Dorothy Byrne, Editor-at-large at Channel 4, observed that "broadcasters sometimes sent the wrong journalists to press conferences (because) the heavyweight political editors pushed out the less well-known health stalwarts." Although this antagonism was mentioned a few times at the symposium, however, it did not emerge as seriously as Byrne's early reflection or Lethbridge's observation. Kirby, who acknowledged that there is that fight for frontpage stories between politics and science journalists at some national news outlets, did not see it as a problem at PA Media. She cited cases in which she worked effectively alongside political correspondents at No 10's press conferences. Ross Lydall, who wears two hats as Health Editor and Town Hall Editor, had no experience of difficulty in the working relationship between science and politics colleagues at the *Evening Standard.*

The rise to challenges

As they tread hard waters to grasp the nature of the virus and the confusing and ever-changing influx of data about its severity and progress, mistakes have inevitably been made. The scientists and journalists at the symposium cited a number of examples and cases in which crucial data have been dangerously – or hilariously – misinterpreted in the news. In April, for instance, Spiegelhalter published a graph showing the risk of dying from being infected by the virus is equivalent to the total risk that an average individual is exposed to in an entire year. Essentially, it means the annual risk of dying would double if one catches the virus. It was, however, misreported all over the news that Covid-19 risk was no greater than the average annual risk. It became a godsend gift for influential anti-lockdown, anti-vaccine figures, such as the far-right Katie Hopkins (now banned on Twitter), to share with millions of followers on social media.

There is much more to do, but we have come a long way

- Jane Kirby -

That said, however, even the most critical at the symposium recognised that journalism has been a crucial positive force in guiding the public through Covid-19 data and science. By August, a study by the Reuters Institute for the Study of Journalism at Oxford University found that 86% of the UK understood what R-naught means and 77% knew what an antibody test is. In the absence of other usual science communication channels, such increase in public understanding about the virus must be attributed primarily to the "pretty good" job of the media, especially science/health/data journalists, said McConway. Spiegelhalter himself, despite having "suffered enough" from many tricky incidents, described his overall experience of journalists' performance during the pandemic as "very good". "There is much more to do, but we have come a long way," said Kirby.

Box 2: Data journalists as data scientists at *The Economist*

In March, *The Economist* received emails about discrepancies between officially reported and actual numbers of deaths in Italy. They decided to embark on their own research journey, turning to other types of alternative data that had not been used at the time, such as official excess deaths and numbers of burials. By May, *The Economist* became "the first news organisation to publish an interactive international tracker of excess deaths across the Western world," said James Tozer, a key journalist behind this project. "We made the raw data available for others to use and since then the tracker has been used widely by other newsrooms as well as cited in over 100 academic studies."

This was only one of many complicated Covid-19 projects in which *The Economist*'s data journalists have effectively turned themselves into data scientists. To deal with the short supply of good data during the first lockdown, for example, the team pooled together Google Map data to measure people's foot movement from nearly 30 cities around the world. Their purpose was journalistic – to tell how bustling cities had become "ghost towns" as people responded to the fledgling pandemic and lockdown restrictions – but the development of their own concepts, tools and methods was scientific. This was, according to data correspondent James Fransham, long before any big tech firm offered their mobility data service. Google eventually widened public access to its mobility data but that might not have happened without initial (unanswered) requests from journalists like Fransham.

Another was their seroprevalence project to deal with the question that scientists were grappling with: how many people have actually been infected by Covid-19 all over the world? Their model was based on 6000 data cells of antibody tests, deaths, cases and testing capacities that they collected from academic papers, preprints and government websites in 19 countries. It estimated that about 630 million people had been infected by mid-September (which was about twenty times the number of diagnosed at the time) as well as predicted when and where undiagnosed infections took place. Senior data journalist Sondre Solstad noted that a couple of weeks after this was published, WHO came up with a very similar estimate, with the same ratio of diagnosed to undiagnosed cases as that of *The Economist*'s.

Kirby and the other journalists at the symposium provided many cases and examples that show energy, resilience and innovation in the way they navigate uncertain waters, detect consistent holes in data, keep politicians and scientists on toes, learn to communicate the "known unknowns", make data relevant and meaningful to audiences, and even do the job of scientists to create their own rigorous pandemic data. Sometimes the success goes far beyond what one expects of journalism. *The Economist* presented three highly complicated, creative Covid-19 data projects that have not only benefited audiences and newsrooms around the world but also been applauded, used and/or followed by hundreds of academic studies, WHO and tech firms (see Box 2).

Beyond these "blockbuster" cases is much independent thinking power, resilience and creativity in the way journalists detect patterns out of the chaos and to create an order out of the disorder on that basis. Here are some of the key take-aways for journalists in their daily micro dealings with pandemic statistics and beyond:

- ***Be sophisticated with data.*** Numbers do not simply speak for themselves and should not be taken at face value. Beneath the surface of each daily Covid-19 death count, for instance, is a very different counting method, each with its own limitations. As Tom Whipple, Science Editor of *The Times*, advised from his own experience of "learning not to trust the daily counts", journalists need to avoid an overreliance on simplistic, even naïve interpretation of data just to tell the story.

One example was the way PA Media navigated uncertainty within data in the early days. As Kirby noted, there was a general confusion because the date something (a positive test case or a death) happens is not the date it is reported. This lack of real-time data makes the reporting of increases or decreases of daily cases rather meaningless. The difference became even starker during the Autumn months, when the pressure increased on the testing system and results were taking even longer to come back and be recorded. In careful discussion with Public Health England, PA Media finally decided to eliminate local authority-level data of the previous four days from their rolling average case rates as they are so incomplete. The BBC later adopted this approach to maintain consistency around their reporting of case rates.

- ***Respect your audience.*** Public receptibility of data is determined by multiple socio-cognitive factors that are beyond the data themselves, as pointed out by Professor Helen Kennedy of the University of Sheffield, Dr Jon Roozenbeek of Cambridge University and Professor Philip Schlesinger of Glasgow University (see Box 3). But don't assume that the public at large is not capable of understanding data and uncertainty. Much science communication research has shown that people can and will learn to deal with uncertain science when *and if* they feel the need to do so. A recent study finds that communicating uncertainty in the news doesn't substantially reduce public trust in data and science. In fact, there have been backlashes from readers and viewers are likely to come if journalists do not do well.

In practical terms, that means journalists need to be transparent in data reporting and be ready to admit uncertainty where required. "Covid-19 is particularly difficult to be communicated because it involves uncertainty about both the present and the future, but no science communication should ignore it," said McConway. If there are different measures for the same thing (e.g., two R-nought figures, three death counts), it is best to report all and explain the limitations behind each of them. That would also mean journalists need to downplay their constant desire for the definite, specific and concrete. Further, Paul Bradshaw called on journalists to expend more efforts on employing innovative interactive tools and

Box 3. Complexity around public receptibility of Covid-19 data

Journalists face the complexity in public receptibility of data. Helen Kennedy argued that the widely-used "data literacy" term is a potentially pejorative term that does not capture many beyond-data determinants of such receptibility. Her research highlights three key elements:

- How people receive data and data visualisations depends on their *emotional engagements* data, such as pleasure, frustration, and confusion.
- People's feelings of trust/distrust in data practices are aligned with their perceptions of the organisation that is responsible for the data practices (e.g., BBC), rather than the data practices themselves.
- Trust in data is intimately related to social justice and structural inequalities, and therefore not something that can be addressed without this starting point.

In line with Kennedy's observations, Jon Roozenbeek's recent study finds that numeracy was only of many factors that determines public susceptibility to Covid-19 mis/disinformation, alongside demographics (including self-perceived minority status) and trust in science.

Philip Schlesinger noted the "major tension between conceptions of rational behaviours and reasoning". Data literacy, he said, needs to be addressed together with media literacy, public trust in mainstream media, and the uneasy relationship between the expert and the lay.

However, not all journalists are aware of, or could incorporate, such factors into their data and science reporting.

platforms – such as data animation – to explain uncertainty and educate audiences about data and science along the way. One example is the attempt by the *New York Times* to engage audiences with statistics by inviting them to draw a line showing what they think the figures are, before displaying the actual ones. Roozenbeek presented early research findings that shows information inoculation through playful gaming could improve public susceptibility to Covid-19 mis/disinformation.

➡ ***Put numbers in context.*** Giving people a comprehensive picture behind the data, so that they can make up their own mind, is another way to show respect to the audience. "There is always an extra layer of complexity," said award-winning freelance science journalist Tom Chivers, and journalists need to bring context to every important number being reported. That requires journalists to be clear about not only what the data tell them, but also what the data do *not* tell them. When it comes to risk, for instance, Professor Stuart Allan of Cardiff University noted that journalists need to explain which risk is acceptable, which is worth taking, and which should be avoided.

The insightful and well-thought-through approach to the pandemic of Tom Whipple and his team at *The Times*, which was praised as a benchmark by both statisticians and journalists at the event, would provide other excellent examples of data contextualisation (see Box 4).

Box 4. How *The Times* kept the second wave in perspective

Tom Whipple was particularly proud of the way in which *The Times* covered the second wave. By plotting key data of both waves – such as outbreaks in care homes, the percentage of patients surviving and other metrics – against each other in the same graph, he explained, they could emphasise the similar yet markedly different challenges that the second wave faced.

For example, by plotting hospital admissions between March and August against those from 1st September onwards, *The Times* highlighted the second wave involved rising hospital admissions yet at a less steep gradient than the first wave.

Further, by comparing how the first and second wave compared in terms of patients surviving COVID-19 in hospitals, their visualisations pointed to the improvement in hospital care during the second wave.

All this, as Whipple was keen to stress, was to provide context for public understanding and in no way to "downplay" the significance of the second wave.

➡ ***"Geo-customise" data to make them personally relatable.*** The pandemic is as global as it is local. Scaling the data down to users' local settings and making them easy to use have helped local and national news outlets, including the BBC. An excellent example is the initiatives by Reach PLC, which has a big stable of local news titles, in updating and localising ONS and government data through interactive maps and a postcode checker. The checker could provide local-level data about a range of topics – deaths, cases, hospitalisations, hospital capacity, food bank use and economic growth – as well as specific lockdown rules. The latter were specified down to the ward level to take into account specific rules that vary across different areas within a city/shire/borough council. The ultimate aim is to make them "accessible, useful and relevant", explained Claire Miller.

➡ ***Humanise the data.*** Numbers and graphs are informative, but cannot tell the whole story. Data and science reporters need to listen to other modes of pandemic explanation and to leave room for emotion, empathy and persuasion. Otherwise, as Whipple said, "one death is a tragedy and one million deaths is just a statistic." Similarly, Anna Feigenbaum,

Associate Professor of Digital Media and Communication at Bournemouth University, argued that the process of turning human lives into numbers and graphs can risk taking on an authorial (and overly technocratic) perspective, sanitise experience and emotion and reduce complexity into simple graphics. In response to this, she called on journalists to look to expand into new ways of telling stories with data, such as using data comics (see Box 5).

- ***Treat scientists as scientists.*** Scientists are the first go-to place for journalists when uncertainty is ripe and the data are quickly politicised. But be sure to put them in the right seat. There are two aspects to this.

First, they should be allowed to talk only about things that are clearly within their specialist realm (no "armchair epidemiologists"). Once that is established, let them talk to their point and with their evidence, with your inquisitive mind. Build trust through calmness, humility and transparency, but don't lose your healthy habit to ask hard questions when necessary.

Second, do not "force" scientists, consciously or subconsciously, into taking side. Scientists generally want to talk from the position of facts and figures, not from a value-laden stance. To give an indication of the importance of this to scientists, two of fourteen media-work recommendations that McConway and Spiegelhalter offered statisticians in the early stage

Box 5: Humanising data through comics: a new innovation for journalism?

Anna Feigenbaum drew on Arthur Frank's *The Wounded Storyteller*, which identifies three main types of illness narratives. Most applicable to a conference on statistical chaos, she outlined, is Frank's concept of "chaos narrative" – a discourse of health that centres on a lack of control, hopelessness, spiralling and anti-resolutions. She argues that this "chaos narrative" encapsulates many of our experiences of COVID-19. In this context, she advocates for an illness narrative perspective to telling COVID-19 stories with data comics.

Drawing on a range of examples - from Katy Doughty's work in the Nib and Mona Chalabi's excellent illustrations to her own work in collaboration with other artists - Feigenbaum argues that health comics humanise data in three important ways.

- The annotative style allows for interactions with graphics that escape the technological or authorial voice. Such practices can help increase data literacy and comprehension.
- Comics can humanise what is counted and what is not counted through illustrative (rather than iconic) imagery. This can create more empathy for the viewer.
- Through their work with participants, these comics allow for agency from those who engage with data visualisations.

All three aspects encourage us to expand beyond sleek, technocratic data visualisations and towards ones that place the people being quantified, and the people interacting with the visualised data, at its centre."

In a follow-up interview, Feigenbaum said that news outlets, as leaders in shaping the field of infographics and data visualisations, have the resources to further explore these potentials for humanising data. While annotation can be found in existing news visuals, more can be done to expand on conventions and experiment with illustrative forms. Features found in weekend edition magazines, for example, might be brought into scientific and statistical reporting to experiment with generating audience empathy.

Future projects are another existing form of data visualisation that could be further explored to help build data literacy, Feigenbaum added. Using simple animated illustrations, such as those designed by cartoonist Toby Morris and Dr Siouxsie Wiles, can help bring future projections to life, inviting users to put themselves in the story. This also helps develop agency, offering an opportunity for people to see themselves as an active contributor to a shared public health goal.

of the pandemic are "Don't be pulled into someone else's arguments" and "Stick to explaining".

- ***Make more use of local health professionals' expertise.*** Local health protection teams routinely work with recordable diseases of different types and provide specialist support to prevent and reduce the effect of infectious diseases. But they receive neither enough financial support from the government nor sufficient media coverage, said Hemingway. She recommended journalists to more frequently garner local health teams' superb expertise in testing, tracing, isolating and caring for patients.

Local health teams can provide fresh news perspectives on the ground too. The *Evening Standard* has benefitted much from these teams. As Ross Lydall explained, his team often turned to clinical staff at hospitals to assess the suitability and usefulness of data provided by the government when reporting on COVID-19. Through these relationships, they recognised that the number of ventilator beds (and their occupancy rates) does not give a picture of how many people are in Intensive Care Units (ICU). Furthermore, critical care doctors often emphasised that many patients in ICU did not have coronavirus but this did not mean the staff were not busy. Lydall emphasised the need to look beyond statistics to understand certain aspects of the virus – most notably, in this case, hospital pressures.

- ***Use science-journalism intermediaries such as the Science Media Centre.*** Since its inception, the SMC model has been a subject of debate for amongst science journalism practitioners and critics. The pandemic might be changing that. Virtually every journalist and statistician at the symposium praised the SMC's helping hand in connecting them with the right experts in an atmosphere of uncertainty and science politicisation. From January to the end of November, according to Lethbridge, SMC offered 101 briefings and 999 Covid-19 round-ups and rapid reactions to journalists, in addition to handling more than 3,200 journalistic enquiries.

- ***Collaborate with other newsrooms.*** In a highly competitive and time-pressured industry, such collaboration is not evident yet. But, as *The Economist* learnt from their data-sharing initiatives, it can only benefit everyone. It conserves newsrooms' scarce resources and amplifies the impact of the work beyond their news organisation's reach, all ultimately for the public good. Data sharing aside, there are other things that could be done. Duncan, for example, called for journalists to work together to standardise the way they report and explain key data and the uncertainty around them.

- ***Tap into the power of citizen science.*** This was not discussed by any journalist at the symposium, but was identified by Stuart Allan as a promising tool for science journalism. His research has shown that citizen science, including the crowdsourcing of big datasets, has contributed to improving the quality of data and science reporting before and during the pandemic. One example is the Covid Tracking Project launched in March from *The Atlantic*, which attracts "hundreds of volunteer data-gatherers, developers, scientists, reporters, designers, editors, and other dedicated contributors". These volunteers collect and submit testing and outcome data from 820 data points on health department websites across 56 states and territories of the US.

A positive long-term legacy?

What has emerged from the chaos of Covid-19 data and science scene calls on critics to rethink the stereotype of journalism as a number-phobic and statistically incompetent profession. Apparently, journalists can do a tremendous job in handling the messiness of frontline science and delivering complicated data to the lay public, *provided they are adequately invested to do so*.

Our symposium was a little skewed to the positive in part because most participating journalists belong to newsrooms that have maintained a more or less considerable support for science, health and data reporting expertise. *The Economist*, in addition to its renowned science journalism, has an established team of 17 data journalists and designers.

Editors' attitudes to – and tolerance of – numbers can play a decisive part. Tom Whipple remembered a moment weeks before the virus was a real public concern, when people were starting to talk about herd immunity. A senior editor at *The Times* "shouted over the newsroom, asking me what it was all about," Whipple said. As he sat down to explain things, Whipple was "prepared for [the editor] to tell me to go quite quickly". He was wrong. "I went through the R number, and he was listening," said Whipple. "I went to herd immunity, and he was still listening. I went on to the disease and epidemiology models, and he was still listening. It's extraordinary because his job was to determine if there was any public interest in the story, and the public up to that point was not interested in those stats."

I went through the R number, and he was listening. I went to herd immunity and he was listening. I went on to the disease and epidemiology models, and he was still listening. It's extraordinary.

- Tom Whipple on how a senior editor at* The Times *tolerated Covid-19 data in February 2020 -

Above all, however, things would not have been so positive without the resilient, tireless and selfless dedication of the not so many specialist science, health and data journalists out there wading through a heavy workload and constant pressures. Jane Kirby, for instance, moved from a pre-pandemic life of working on slower stories on the NHS to now chasing and turning around breaking news throughout the day. The three-strong data team at Guardian News and Media previously worked on medium to long-term projects with correspondents in other news beats, but now had to face a constant demand for very short-term stories. Duncan cited some telling personal figures: nearly one third of the 338 stories she had done in her five years at the *Guardian* was about Covid-19 between March and early December. That has not included her acting, like many other specialist journalists, as a data consultant for the news desk, where "too many figures can confuse rather than illuminate".

That leads us back to the condition above: investment. Journalism as a whole is not in a good position to deal confidently with statistics yet. The pandemic – with many simple mistakes made by those with no specialist expertise in dealing with data and science, especially political journalists in the early stage – highlights rather than belittles the need for more statistical expertise among journalists.

One of Covid-19's positive legacies might be a growing appetite for data and science, with all their uncertain properties, among both audiences and journalists. Bradshaw was optimistic that the combination of the pandemic and the recent US election will stress the need for journalism to deal with data. Coupled with recent demands for statistical skills set by training bodies such as the BJTC, Bradshaw said, such heightened need will push newsrooms into that direction. On part of the public, Allan hoped that many citizens would come out of the pandemic with the realisation that they need good journalism to deal with complicated reality, including the messiness of frontline science, and that it needs to be paid for. As for science and statistics institutions such as the RSS, there is a clear willingness to collaborate with the media, which this joint symposium itself represents. All this, we hope, will constitute a good professional and commercial case for newsroom executives to have a strategic commitment to strengthening the crucial but grossly overlooked specialist expertise in data and science reporting.

Thinking of the future, Pamela Duncan borrowed a quote from a Radio 4 More or Less show to express her hope that Covid-19 will bring data journalists from a peripheral position to the status of "rock stars of the newsroom". We would wish the same for science and health journalists.

Report's authors

- Dr An Nguyen is Associate Professor of Journalism at Bournemouth University. A former science editor from Vietnam, he has published research in the broad area of science, health and data journalism and communication, including *News, Numbers and Public Opinion in a Data-Driven World* (Bloomsbury, 2018).

- Dr Xin Zhao is Lecturer in Communication and Journalism at Bournemouth University. A former intern science journalist from China, she has published research about journalism practices in journals such as *Journalism* and *Asian Journal of Communication*.

- Dr Brendan Lawson is University Teacher in Media and Communication at Loughborough University. He was recently awarded his PhD from the University of Leeds, specialising in the communication of numbers during crises. This research has been published in a range of journals, from *Media, War & Conflict* to *European Journal of Communication*.

- Dr Daniel Jackson is Associate Professor of Media and Communication at Bournemouth University. His research broadly explores the intersections of media, power and social change, including news coverage of politics and political communication. He has edited five books and is co-editor of the election analysis reports, published within ten days of major electoral events.

Acknowledgements

- The symposium was the inaugural public event of the new Science, Health and Data Communication Research Group in the Faculty of Media and Communication at Bournemouth University. It was funded from the faculty's quality-related research budget. Carole Haynes, former Education and Policy Manager at the RSS, helped to facilitate the development of this symposium between March and October, 2020.

- Khanh Hoang, PhD Candidate at Bournemouth University, provided excellent assistance throughout this symposium, including the design of this report.

- This report is built on an earlier, shorter commentary, Coronavirus, Statistical Chaos and the News, One Year On, which An Nguyen published on Journalism.co.uk on 15 January, 2021.

- Kevin McConway, Andy Extance and Thomas King provided very useful feedback to an earlier draft of this report.

- Images used in this report are sourced from Shutterstock.

Presenters and Respondents at the Symposium

Stuart Allan, Professor & Head, School of Journalism, Media and Culture, Cardiff University

Paul Bradshaw, Pioneer Data Journalist & Scholar, BBC Shared Data Unit and Birmingham City University

Tom Chivers, Award-Winning Freelance Science Writer

Timandra Harkness, Science Presenter and Writer & RSS Fellow

Andy Extance, Freelance Science Journalist and Chair, ABSW

Anna Feigenbaum, Associate Professor of Communication and Digital Media, Bournemouth University

James Fransham, Data Correspondent, *The Economist*

Pamela Duncan, Data Journalist, *Guardian News & Media*

Ann Hemingway, Professor of Public Health and Wellbeing, Bournemouth University

Helen Kennedy, Professor of Digital Society, Sheffield University

Jane Kirby, Health Editor, PA Media

Fiona Lethbridge, Senior Press Officer, Science Media Centre

Ross Lydall, Health Editor & City Hall Editor, *Evening Standard*

Claire Miller, Head of Data Journalism, Reach PLC

Kevin McConway (Keynote Speaker), Emeritus Professor of Applied Statistics, Open University & Ex-Vice President for Academic Affairs, RSS (Keynote)

Lawrence McGinty, Chair, Medical Journalists' Association & former Science and Medical Editor, ITV

An Nguyen (Conference Chair), Associate Professor of Journalism, Bournemouth University

Einar Thorsen, Professor & Executive Dean, Faculty of Media and Communication, Bournemouth University

Philip Schlesinger, FRSE, FAcSS, FRSA, Professor in Cultural Theory, University of Glasgow

David Spiegelhalter, Chair, Winton Centre for Risk and Evidence Communication, Cambridge University & Co-chair, RSS Covid-19 Taskforce

James Tozer, Data Journalist, *The Economist*

Jon Roozenbeek, Postdoctoral Research By-Fellow in Psychology, Cambridge University

Sondre Solstad, Senior Data Journalist, *The Economist*

Tom Whipple, Science Editor, *The Times*