ART/DATA/HEALTH
data as creative material
for health and wellbeing

University of Brighton

UNDERSTANDING HEALTH AND WELLBEING DATA

An ART/DATA/HEALTH resource

UKRI Arts and Humanities Research Council

Published by *ART/DATA/HEALTH: data as creative material for health and wellbeing*, University of Brighton, Brighton, UK. July 2020.

A full-text PDF of this resource is available from: https://artdatahealth.org/

Suggested citation:Aristea Fotopoulou, Harriett Barratt and Elodie Marandet.2020. *Understanding Health and Wellbeing Data*. Brighton, UK: ART/DATA/HEALTH: data as creative material for health and wellbeing.

CONTENTS

1. About the project 03

2. Health and wellbeing data 04

3. Health and wellbeing data during the COVID-19 pandemic 11

4. Resources 16

1

ABOUT THE PROJECT

This ART/DATA/HEALTH resource has been developed by researchers at the University of Brighton. The project 'ART/DATA/HEALTH: data as creative material forhealth and wellbeing' is funded by the Arts and Humanities Research Council(UKRI-AHRC Innovation Fellowship) and led by Dr Aristea Fotopoulou at the University of Brighton.

The academic research project 'ART/DATA/HEALTH: data as creative material for health and wellbeing' creates an innovative and interdisciplinary process that offers new tools, at the intersections of data science with art practice, to approach two key issues: digital skills and health literacy.

The project uses health and wellbeing data as the source of experiential stories and as the source material for creative expression. In a series of exploratory workshops, a community of artists, service workers, academics and people living in Brighton and Hove used a combination of creative media, storytelling and data analytics to explore evidence around health and wellbeing. They are co-producing creative work that takes various forms, using both anonymised personal and open statistical health and wellbeing data. The project was adapted mid-way through to reflect the extraordinary circumstances of the Corona virus pandemic.

The artists involved in the project are: Anna Dumitriu, working with staff from the domestic abuse charity RISE on the rise of domestic abuse during C19; Kate Genevievve; Hydrocracker; and Caroline Beavon, focusing specifically on C19 data.

Key partners are:

Brighton & Hove City Council

BRIGHTON DIGITALFESTIVAL

NHS Brighton and Hove Clinical Commissioning Group

BRIGHTON FRINGE

Freedom from Abuse and Violence Rise

To find out more about how to work with data, download the resource *Learning to Work with Data: A Data Skills Toolkit*. Published by ART/DATA/HEALTH: data as creative material for health and wellbeing, University of Brighton, Brighton, UK. July 2020. A full-text PDF of this resource is available from: https://artdatahealth.org/

UKRI Arts and Humanities Research Council

The Arts and Humanities investigate the values and beliefs which underpin both who we are as individuals and how we undertake our responsibilities to our society and to humanity globally.

2

HEALTH AND WELLBEING DATA

TYPES OF HEALTH AND WELLBEING DATA

Health and wellbeing data take lots of different forms, including:

- **Patient data:** information on someone's own health and wellbeing collected by a health body such as a GP surgery or hospital. This might include medication records, or blood pressure readings, for example. Some health insurers also collect this information.
- **Individually collected data,** for example, through wearable devices and mobile phone apps that track heart rate or exercise patterns.
- **Data collected by third parties,** for example by a third sector or community group, by employers, and by companies.

WHY WORK WITH HEALTH AND WELLBEING DATA?

Working with health and wellbeing data can give individuals, researchers, and organisations important insights. For example, data might allow governments to identify links between health and poverty, or help individuals to notice what they could change to live more healthily.

For third sector organisations, data are the materials that can fuel strategic planning and advocacy work. They not only highlight problems, but also act as the evidence to underpin clear, engaging stories.

THINGS TO BE AWARE OF WHEN WORKING WITH HEALTH AND WELLBEING DATA

Anyone working with data needs to understand how to treat them so that they can be used and shared safely and ethically. There are several ways you can anonymise data to protect the individuals concerned:

- **Anonymised data** are data from which it is not possible to identify individuals either directly or indirectly. Anonymised data can be **aggregated**. This means that statistical data about several individuals has been combined to show general trends or values, without identifying individuals.
- **Pseudonymisation** means using a unique identifier to distinguish individuals within a data set. This means that their 'real world' – most commonly their name – is not evident, but that their records can still be tracked and linked together.
- Data can also be **de-personalised**. This means that the information does not identify an individual, because identifiers or identifiable data have been scrambled or removed from that single record. However, the information is still about an individual person, so it needs to be protected.

Some individuals have special protection under law. **Special category personal data** are data defined under the Data Protection Act (1998) as data that identify a living individual regarding his or her:

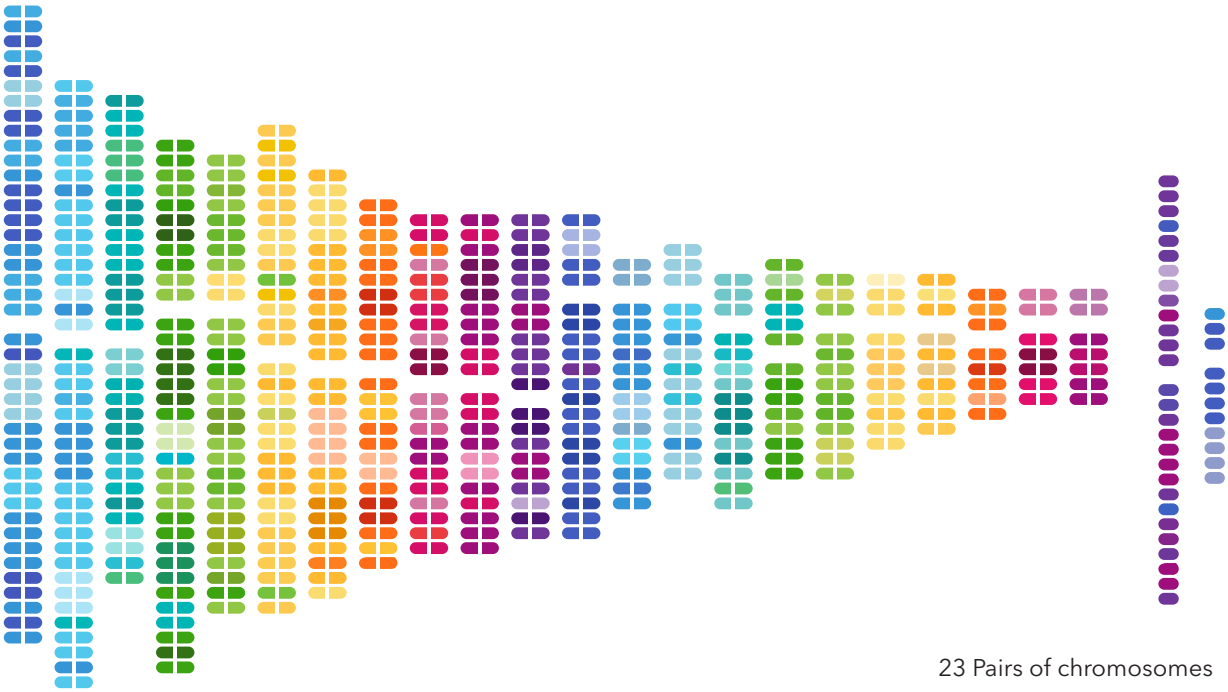
- racial or ethnic origin,
- political opinions,
- religious beliefs or other beliefs of a similar nature,
- membership of a trade union,
- physical or mental health or condition,
- sexual life,
- generic data and biometric data convictions,
- legal proceedings against the individual or allegations of offences committed by the individual.

HOW TO WORK RESPONSIBLY WITH HEALTH AND WELLBEING DATA

When working with data, it's vital to think about what problem is being explored, who the stakeholders are, and who holds the power to decide what is collected and how it is used.

The ART/DATA/HEALTH project puts emphasis on working responsibly with partners and data, and communicating clearly and openly. Organizations can work with health and wellbeing data in ways that improve public trust and understanding. Some key ways in to do this are to:

- Engage community organisations in research design and data analysis
- Design your consent processes with the user at the center: make sure they are properly informed about how their health and wellbeing data may be used
- Focus on complex social issues, particularly the factors which contribute to unequal access to healthcare – some of which are hugely stigmatised
- Explore how individual stories can bring shared experiences and themes to light.
- Develop skills in data analysis and security.
- Be conscious about implicit bias or structural inequalities that your work may be perpetuating, particularly in relation to protected characteristics.



23 Pairs of chromosomes

PATIENT DATA

Patient data are data collected about a patient whenever they go to a doctor or receive social care. This can include details about the patient's health, medication record, health and social care needs, and other information. These data are recorded and stored in a care record.

Patient data use has allowed researchers and health professionals to provide **better care and improve health,** by:Improving individual care, for example by

- improving newborn screening;
- Making more accurate diagnoses, for instance by researching the earliest changes in the development of dementia;
- Enhancing treatment and prevention, e.g. by learning more about blood cancer;
- Increasing patient safety, for example by demonstrating the safety of a treatment for heart failure;
- Improving the planning of NHS services, e.g. by analysing the local variations in stroke prevention;
- Evaluating policy e.g. extending the flu vaccination programme to children; and
- Better understanding diseases, for instance by following up people discharged from hospital after a heart attack. (Understanding Patient Data, 2018)

Digital data approaches and tools have changed many **treatment practices**. For example, NHS.uk's e-prescribing system is widely used, as is their online advice repository, which had 550 million visits in 2016 (NHS, 2017). In some countries and areas of practice, the use of digital tools is now mandatory: for example, the US provided financial incentives for pharmacies switching to e-prescribing for chemotherapy drugs in the late 2000s, given how severe the impact of human error in dosage measurements would be (Zadeh and Tremblay, 2016).

In the UK, The Department for Health and Social Care (2018) is trying to make the use of health and wellbeing data more secure, open and inclusive. For example, the NHS Long Term Plan (2019) lays out an ambition for developers to create new software to improve how local health and care records (LHCRs) are recorded. Five LHCR regions, together covering 40% of the population of England, have been awarded £7.5 million each to carry out this work (Castle-Clarke and Hutchings, 2019).

Many health sector research funders (such as the Wellcome Trust, the Medical Research Council, the Biotechnology and Biological Sciences Research Council, and Cancer Research UK) either encourage or require their funded researchers to share data (UK Data Services, n.d.).

Meanwhile, technical policy initiatives such as FHIR standards (Fast Healthcare Interoperability Resources, an international standard for the exchange of electronic health records) have emerged to enable and promote data sharing.

However, there are some big challenges for both individuals and organisations around how patient data is collected, used and shared. For many patients, privacy is a huge concern. Organisations, meanwhile, need to make sure that they are fully adhering to legal, technical and ethical standards across the whole process of working with health and wellbeing data, from collection to use to storage.

There are **strict controls around how and why companies can access and use patient data**. National bodies play a key role in raising awareness of these controls. However, a 2019 report from the Nuffield Trust found that there was often a lack of **national-level strategic thinking and leadership** in regard to digital standards and data sharing (Castle-Clarke and Hutchings, 2019). As a result, there are examples across the UK of providers refusing to engage with local data sharing efforts (ibid.).

A 2019 study found that the majority of studies into data privacy reported that people are **generally worried about the security of data** and fear 'data leakage' (for example, through hacking or by companies selling data) (Skovgaard, Wadmann, & Hoeyer, 2019). The sharing of **sensitive data** (including information on mental health, sexual health, sexual preference, and religion) holds greater concern than other types of data, as does the use of data by commercial companies (Understanding Patient Data, 2018). More transparent information and communication, and space for dialogue around data use and data sharing, are key to securing users' trust.

There are also **demographic differences** in relation to attitude to data sharing. Young people are more supportive of electronic patient records and more likely to think of the benefits of patient data sharing (Understanding Patient Data, 2018), while ethnic minorities may be less confident about their data being secure if used for health research (Hunn, 2017). There may also be a risk of the emergence of a two-tier model of access, under which only those with financial means are able to afford to protect their own data (Sharon and Lucivero, 2019).

The UK's controversial **care.data NHS project** encapsulates many of these issues. Announced in 2013, the programme sought to extract data from GP surgeries into a central database. It was positioned as a new chance to make real inroads into harnessing the potential benefits of sharing health data, from longitudinal research to the monitoring of public health populations. However, fears around risk management were publicly voiced by ethicists and clinicians from the start (Sheather and Brannan, 2013). These concerns ranged from the lack of clarity around what patients were told about consent and potential risks (McCartney, 2014), to the deliverability of the programme's scope, budget and schedule, to the lack of sufficient ethical protocols and safeguarding (Mann, 2014). The programme was postponed, suspended and restarted several times in different locations between 2014 and 2016, when it was eventually abandoned. The care.data project has been described as an example of the failure to consider the need for a **'social licence'** (Carter et al, 2015). In the end, it failed to secure public trust, created a rupture in traditional understandings of the role of the general practitioner, and failed to articulate itself as a 'public good'.

DATA COLLECTED BY INDIVIDUALS, USING MOBILE PHONE APPS, WEARABLES

Individuals collect data by using mobile phone apps and wearable devices such as Fitbit. The number of connected wearable devices worldwide has more than doubled over the last three years, increasing from 325 million in 2016 to 722 million in 2019, and is forecast to reach more than one billion by 2022 (Statista, 2020). Using wearables to collect health and wellbeing data can help individuals to understand and make changes to their behaviour patterns, from diet and exercise to sleep. This is why self-tracking using wearables and other digital technologies is often understood to empower people.

However, there are many privacy and security issues associated with self-tracking devices, as well as questions such as: who retains rights of the data, how do companies use them, where do they end up, and who profits from selling these data?

Wearables and apps collect a broad range of data, including biological processes such as heart rate and sleep patterns, lifestyle factors such as diet, distance and speed travelled, behavioural and postural data such as walking gait (to prevent falls), and much more besides. They are collected through a combination of automated data collection (for example, a pedometer built into a phone or a wristband) and self-generated data (such as those data provided to health apps on mood, symptoms and physical condition). The format of the data is primarily quantitative, allowing easy exporting and comparison, though some apps also use qualitative input such as free text. Many offer customisable features – for example, the 'Clue' menstruation and fertility tracking app allows users to create 'non-standard' symptoms to be tracked on top of those provided (see Fotopoulou 2017). Although many commercially available wearables seek to position themselves as fashion items, some seek to be as unobtrusive as possible. For example, the 'Hospital Without Walls' project in Australia piloted a portable radio which included a discreet accelerometer to detect falls amongst an elderly care population (Wu and Luo, 2019). Some biosensors, act at the level of cells, enzymes and biological material. For example, blood glucose monitors used at home by diabetics account for 85% of the global market (Mehrotra, 2016).

This has also led to patient-led, 'DIY' movements. The 'We Are Not Waiting' campaign (**https://openaps.org/**) is a group of patients with diabetes who have developed their own glucose monitoring system more quickly than commercial companies have been able to. The explosion in the global use of wearables has led to a new sense of how we see ourselves, known as the 'quantified self'. This term refers both to the phenomenon of self-tracking, and to a community of users – known as Quantified Self or QS – who swap tips and resources.

In this way, self-tracking might embed a sense of an 'ideal' self – a fit, healthy and busy self whose every move is logged. This ideal is a normative one (Fotopoulou and O'Riordan, 2016) – it does not allow for widespread differences in bodies, health conditions or lifestyles. Self-tracking also shifts responsibility onto the individual for their own care. In this way, it can be seen as a symptom of a more general phenomenon: a shift towards healthcare systems where the state plays a smaller part than the commercial sector.

As well as apps and wearables which upload personal data to centralised databases, users are increasingly encouraged to share their data with other users, organisations, and reasearchers. The drive towards 'data-sharing' – embodied by phrases such as 'sharing is caring' – even takes on moral undertones (Fotopoulou, 2018).

Some employers are also now offering wearables as part of their employee benefits scheme. However, questions around the use of data generated by wearable devices are complex in a workplace setting. Research carried out by PwC in the UK in 2015 found that, while two thirds of workers (65%) wanted their employer to take an active role in their health and wellbeing, only 46% would accept a free piece of wearable technology if their employers had access to the data recorded (PwC, 2016). Four in 10 said that they didn't fully trust their employer to use it for their benefit, while 37% didn't trust their employer not to use the data against them in some way. Workers who would be happy to use a wearable device at work are most likely to want to trade their personal data in exchange for flexible working hours, free health screening and health and fitness incentives.

While users are offered a summary of their data, it is often difficult to find and extract the data sets themselves (most data standards recommend that data are exportable in a csv format to allow easy analysis – see the **Five-star deployment scheme** for open data licensing, built by Tim Berners-Lee, and the Open Data Institute's **Data Certificates**). In conjunction with the difficulty of user access, many apps retain the rights to sell on the data they collect in an anonymised form. It may also often be the case that the full data sets are only legible to health professionals – for example, devices which measure the haemoglobin concentration in the blood, or the severity of a hand tremor, produce data which are not easily read by patients, and need to be interpreted.

Self-tracking has also led to concerns about over-diagnosis (Li et al, 2017): is it always the best thing to have access to lots of quantitative data? How are they 'read' or prioritised alongside hands-on examinations, or people's everyday experiences of their own health?

DATA COLLECTED BY THIRD PARTIES

Many kinds of health and wellbeing data are collected by organisations whose main activity lies outside of traditional healthcare. This includes charities, local government, employers, community groups, commercial bodies and many more.

The research fieldwork of the ART/DATA/HEALTH project involved working with community groups that collected health and wellbeing data of both qualitative and quantitative nature. These related to experiences of domestic abuse; data which tracked someone's journey through a support service, including rates of re-engagement; details of volunteers supporting a 'befriending' service with isolated people, and many more. Each set of data is used by these organisations in specific ways – for example, for communications, lobbying, and bid-writing purposes, as well as to inform patients' treatment options. Some of these data are entered into a national, anonymous database, while others are kept locally by a small number of staff members.

New technological practices – such as the use of data storage in the 'cloud' – mean that the health and wellbeing sectors start to adapt to a **decentralised, interconnected structure**. This gives an opportunity for decision-making to be spread across a wide range of individuals, researchers and organisations, and for health research to more fully take into account a wider range of people, from participants to third sector organisations. But enhanced connectivity also means that there is a need for clear, shared approaches to ethics, consent and governance of health and wellbeing data, which inevitably differ from those that apply to commercial data (Bot, Wilbanks, and Mangravite, 2019).

There is a **global market in healthcare data** beyond the local community level. Third parties gather de-identified data from healthcare systems, pharmacies and other sources, and sell it on to buyers interested in analysing large data sets. These are known as **'data brokers' or 'information brokers'**; their customers include government bodies, marketing agencies, and research organisations. For example, the US health information technology firm Iqvia (formerly IMS Health), one of the biggest companies in the health data market, generated $8 billion in revenue in 2017 (Arndt, 2018).

These companies' access to health and wellbeing data differs from country to country. In the UK, there are strict controls around how and why companies can access and use national patient data, as outlined above (Understanding Patient Data, 2018). However, models of data ownership and sharing differ hugely across the private companies offering internet solutions, wearables and app technologies. For example, a 2018 study (Binns et al, 2018) found that the majority of 959,000 apps from the US and UK Google Play stores transferred data to third parties, and that many of these operated on a transnational basis (and therefore not necessarily in adherence to the legal system of the country of use).

In the EU, the General Data Protection Regulation (GDPR) legislation established in 2018 seeks to answer some of these concerns. However, while consent forms may be technically compliant with GDPR, this may not mean that they adhere to high standards of **informed consent**. When dealing with health and wellbeing data, it is important to place the user at the centre of your consent processes and to make sure that they know exactly how their information will be used (Schairer, Rubanovich, and Bloss, 2018).

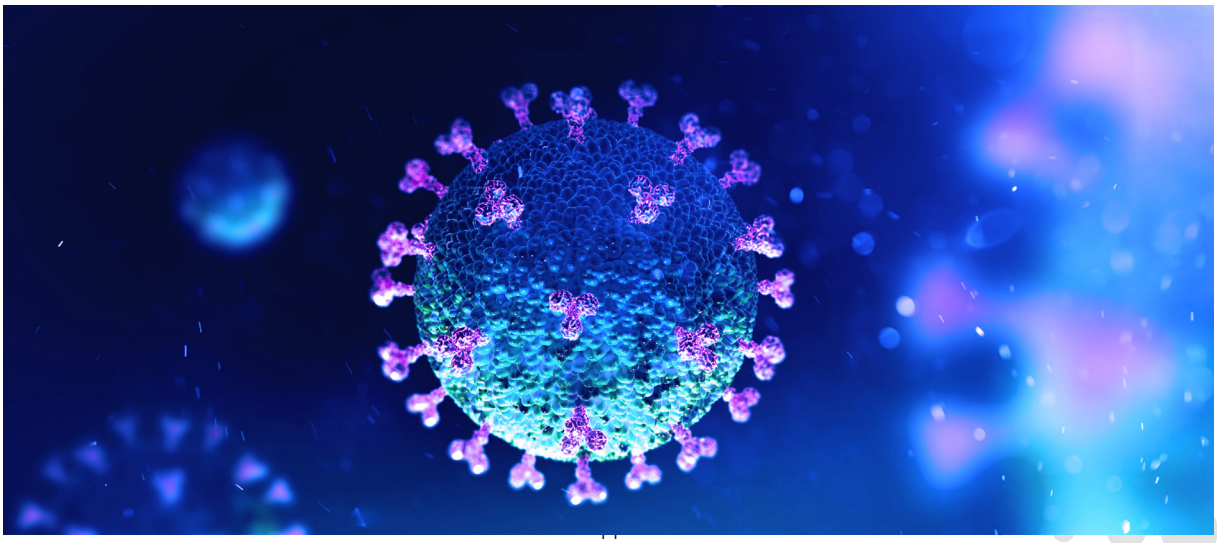
3

HEALTH AND WELLBEING DATA DURING THE COVID-19 PANDEMIC

The COVID-19 pandemic has turned public attention to the crucial role played by data in enabling us to understand rapid developments worldwide and to take appropriate measures. However, the crisis has also highlighted that these data do not speak for themselves; their collection, use and presentation to the public are complicated.

Take for example the number of deaths from COVID-19. Controversies about these numbers have shown that counting the victims is more complex than expected. In the UK, the daily announcements in the first weeks of the pandemic only included deaths in hospital of those who tested positive for COVID-19. Even then, those who had recently passed away may not be counted as there was generally a delay of a few days in hospital reports, potentially leading to considerable discrepancies. For example, while on 27 March 2020 the government announced that 926 COVID-19 deaths had so far taken place in English hospitals, NHS England now reports that the true figure was 1,649 (Richardson and Spiegelhalter, 2020).

A more reliable number is the one collated by the Office for National Statistics from death certificates issued by local authorities. This would include deaths occurring in the community and care homes. However, deaths can be reported for up to thirteen days after a person passed away, so numbers are likely to lag even further behind. In addition, in the absence of systematic testing, it is likely that some deaths caused by COVID-19 may not be registered as such, leading to an underestimation of mortality rates.
A more accurate number still – but suffering even further delays – is arrived at by looking at excess of deaths i.e. the number of extra deaths recorded during a certain period (compared to a similar period in previous years). This excess can therefore be attributed to either COVID-19 or to the new context. But new issues crop up: how to ascertain how many deaths are down to new policies such as lockdowns, and their effects (e.g. reduced medical cover, increase in domestic violence, poverty, mental health issues)? Or how many lives have been saved by these changes (e.g. reduced travel, increased air quality)?
What is more, it is likely that some of those who lost their lives to COVID-19 would have passed away later that year, effectively bringing their



loss earlier rather than adding to the overall numbers. However, this figure may provide a better approximation of the truth. For example, on 22 April 2020, the Financial Times published extrapolations showing that the likely number of "excess deaths" since the start of the pandemic in the UK could be in the region of 41,000, rather than the official 17,337 fatalities officially recorded (Giles, 2020).

Without mass testing, and given the wide range and variable severity of the symptoms, ascertaining how many people have had COVID-19 is also far from simple and is best done in retrospect. For example, on 21 April 2020, a study by researchers at Hong Kong University's School of Public Health published in The Lancet found that more than 232,000 people might have been infected in the first wave of COVID-19 in mainland China (Tsang et al, 2020). In an article in March 2020, Tomas Pueyo showed how 22 COVID-related deaths in Washington state probably translated into 3,000 cases state-wide, using approximate mortality rate, average length of the illness before death, predicted transmission rate, and looking for any large clusters than would skew results (Pueyo, 2020).

Reliable data is key in ensuring that any responses to the pandemic addresses inequalities. Data journalism has been crucial in shedding light on the fact that Black, Asian and minority ethnic (BAME) populations may be suffering disproportionately from the pandemic, with deaths of BAME healthcare workers especially high. See this simple but striking data visualisation by Bristol-based artist Niki Groom, using statistics cited in Sathnam Sanghera's article, 'Coronavirus and ethnicity: black and Asian NHS medics on the front line'.

This issue has also exposed **gaps in data gathering**. For example, in the UK, ethnicity is not registered on death certificates (Barr et al., 2020). In the US, neither the Centers for Disease Control and Prevention's information site, nor the Johns Hopkins University database and the COVID Tracking Project offer data desegregated by ethnicity (Kendi, 2020a). But data from hospital deaths suggests that, up to 19 April 2020, 19% of those who died in hospital in England were from BAME backgrounds when BAME residents make up 15% of the country's population (Barr et al, 2020). The release by the Office for National Statistics of data for COVID-19-related deaths sorted by local authority was key in enabling analysts to take this further. It allowed them to establish correlations between COVID-19 deaths and local authority data such as population and environmental characteristics and revealed that a high proportion of BAME residents was the strongest predictor of a high COVID-19 death rate. For every 10% increase in ethnic minority residents there were 2.9 more COVID-19 deaths per 100,000 people (Barr et al, 2020). These correlations were echoed in the US after the New York Times released data by zip code (Kendi, 2020b).

DATA VISUALISATION DURING COVID-19

As the COVID-19 pandemic progressed in 2019 and 2020, effective visualisation has been particularly important to communicate ideas and facts about the situation.

Good data visualisation can be incredibly useful in conveying information in a format that is easy to understand and help shape policy debates. For example, this data visualisation by Financial Times journalist Bob Haslett compared the US and China in terms of how effectively the spread of the virus was contained. The graphics in this New York Times article contrast confirmed infection with unconfirmed infections, using a subtly creative approach to mimic the spread of the virus through the air. It shows clearly the different between the the very small number of confirmed cases of COVID19 in five major US cities on 1st March 2020 (23 cases) with what researchers now suspect to have been 30,000 unconfirmed infections.

COMPUTER MODELLING

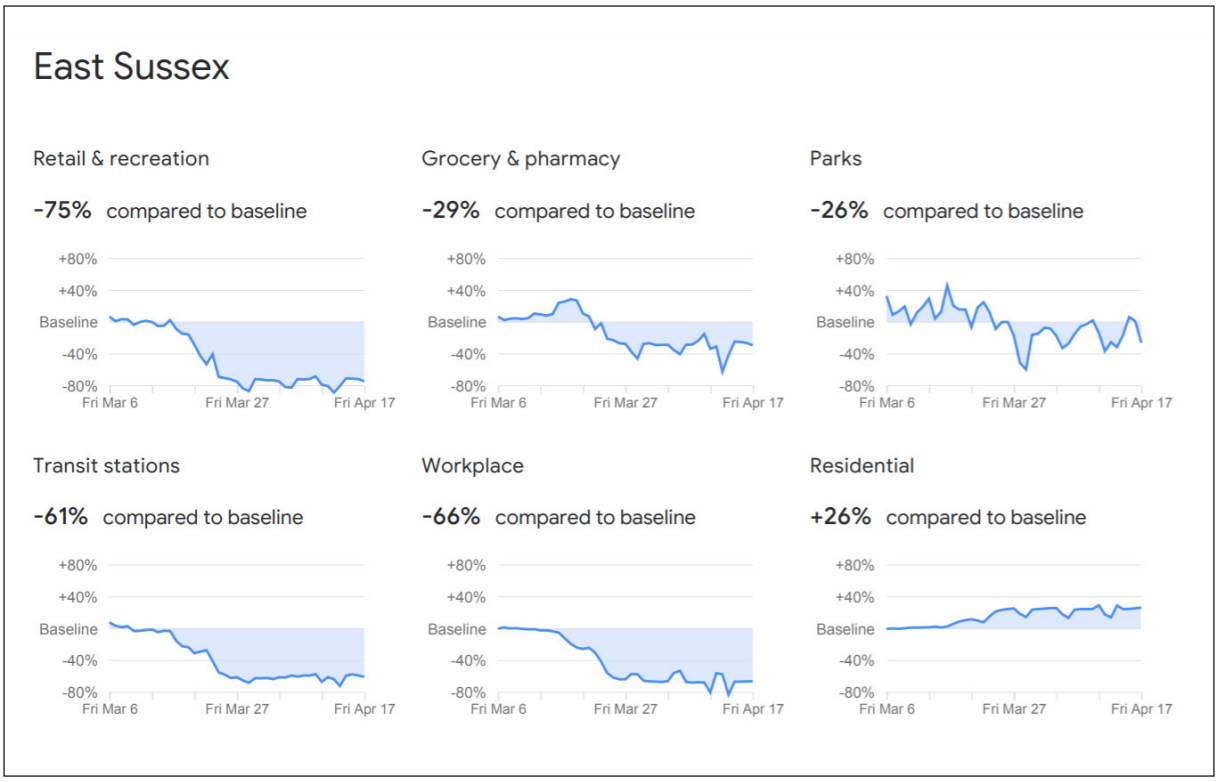
Computer modelling of the COVID-19 pandemic has also been key in discussing and evaluation ways forward. Caitlin Rivers, an infectious diseases modeller with the Johns Hopkins Center for Health Security argues: 'modeling plays a really important role in understanding how an outbreak is unfolding, where it might be going, and what we should be thinking through'.

A Washington Post model visualisation (Stevens 14 March 2020) – shared extensively in social media as the key to understanding social distancing – shows a simulation of people depicted as dots, which move around and thus 'infect' one another. It shows changes of count of the recovered, healthy and sick over time, but interestingly it does not depict deaths.

However, it's worth remembering that modelling is only as good as the data it is based on. Models provide indications, not accurate predictions. The quality of available COVID-19 data, in particular, is poor: "Right now the quality of the data is so uncertain that we don't know how good the models are going to be in projecting this kind of outbreak," says Marc Lipsitch, an epidemiologist at the Harvard T.H. Chan School of Public Health (Greenfieldboyce 2020).

Reponses to the COVID19 pandemic also provide a good example of the many and various ways you might choose to approach a particular topic, rather than focusing on the most obvious types of data. While many visualisations have portrayed the numbers of cases per country, visualisations showing the decrease of air pollution or the risk factors for different jobs have also been widely shared. You can read more about this here.

Google offers Community Mobility Reports which allow you to search by region to see how social distancing measures are working locally. For example, these graphs show how Brighton and Hove is performing compared to its usual baseline:



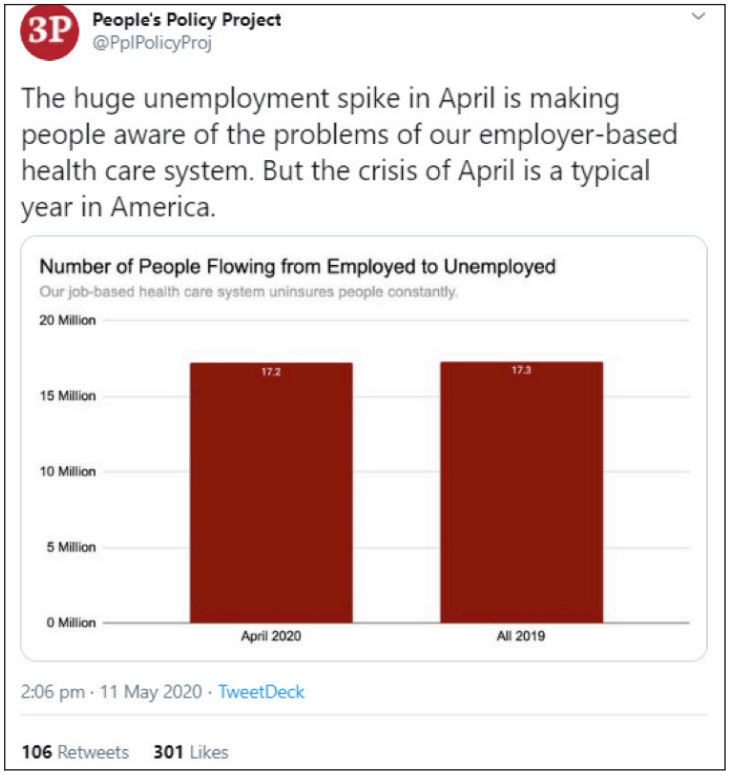
© Google Community Mobility Reports – figures accurate to 14th May 2020

You can read more about data visualisation approaches to the COVID19 pandemic here and here (resource links also appear in full in the Resources section).

TRAINING YOUR CRITICAL EYE

Be aware that many data visualisations are misleading – always take a critical view. Do the segments of a pie chart actually add up to 100%? Are the segments proportionate? Do the horizontal and vertical axes of a bar chart make sense?

For example, take a look at this supposedly comparative bar chart comparing the rate of people moving from employment to unemployment in two periods of time, tweeted by the US think tank People's Policy Project:



People's Policy Project, *The huge unemployment spike in April is making people aware of the problems of our employer-based health care system. But the crisis of April is a typical year in America*, Tweet, 11 May 2020 **https://twitter.com/PplPolicyProj/status/1259832193117020162**

WHAT DO YOU NOTICE?

At first glance, the two bars on this chart show comparable data. But the data labels within each red bar are extremely hard to read. The title 'huge unemployment spike' is not visible at all! Notice how the y-axis (running horizontally) is not an even scale: it gives the same weight to one month ('April 2020') as it does to twelve months ('All 2019'). And why does 2020 appear first, given that English speakers read left to right?

It would be more effective to use colour to make a clear case for the story it is trying to tell. A completely different type of visual graphic might have done the job. And there is a real lack of further information – who made this graphic? Where are the figures collected from? Even if these are publically available, it's good practice to clearly reference your data. You can follow the link to the original tweet (**https://twitter.com/PplPolicyProj/status/1259832193117020162**) to see how commentators have asked for a description of the sources used by the think tank, as well as clearer labelling of the graph.

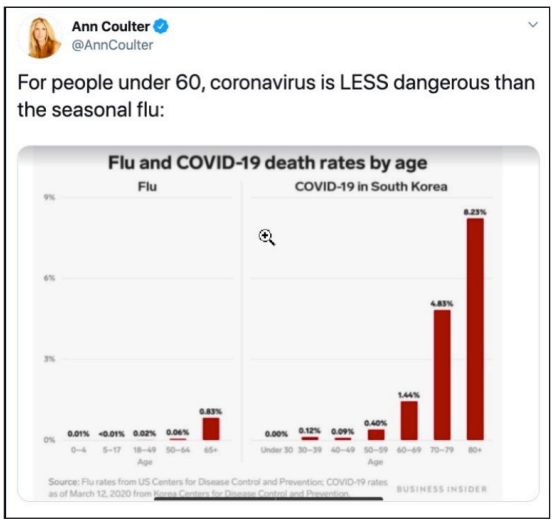
MISLEADING DATA VISUALISATIONS

Some data visualisations may be unintentionally misleading, but be alert to graphics that lead you to a conclusion that support specific agendas.

For example, this graph from business insider was tweeted by Ann Coulter, a far-right US media pundit, to minimise the impact of COVID-19. Have a look at the two graphs.

The message is *For people under 60, coronavirus is LESS dangerous than the seasonal flu*. If you notice the visualisation however, you will see how the two graphs (Flu on the left and C19 on the right) show different age categories and different countries. The two graphs compare death rates associated with seasonal flu in the US with deaths associated with COVID-19 in South Korea. The death rates due to C19 in South Korea have been one of the lowest rates recorded worldwide. The two countries have demonstrated very different ways of dealing with the spread of the virus, so comparing the two data sets without additional context is misleading, and is only used here for political purposes.

What is most striking however is that the figures don't support the assertion made in the tweet! The graph on the left shows 0.08% death rate for flu for Americans under the age of 65. The graph on the right shows 0.61% death rate for South Koreans under 60 for COVID-19. This is over 7 times higher!



Ann Coulter, *For people under 60, coronavirus is LESS dangerous than the seasonal flu*, Tweet, 24 March 2020, **https://twitter.com/anncoulter/status/1242484117373100037?lang=en**

This example illustrates how bad visualisations are not only easily misread but also misrepresented. Read the article about Coulter's misuse of the original figures by journalist Andy Kiersz.

RESOURCE:
see **Viz.wtf** for more examples of data visualisation gone wrong.

4

RESOURCES

- Washington Post visualization showing the effects of social distancing on the spread of the Coronavirus: **https://www.washingtonpost.com/graphics/2020/world/corona-simulator/?fbclid=IwAR2hC8DXtddmohRgJijWfAuq9Xzh7XTo30FNspZ69Rv-3oTKbeodrsJ7lYY**

- 'Ten Considerations Before You Create Another Chart About COVID-19': this piece, while specific to the pandemic, is a good set of rules for assessing how appropriate any data visualisation project is: **https://medium.com/nightingale/ten-considerations-beforeyou-create-another-chart-about-covid-19-27d3bd691be8**

- 'Five questions to ask when you see acoronavirus map': **https://medium.com/@garrett.dash.nelson/five-questions-toask-when-you-see-a-coronavirus-map-8e7ec56feeac**

- 'Thirteen Things to Visualize About COVID-19 Besides Case Loads': **https://medium.com/nightingale/thirteen-things-to-visualize-aboutcovid-19-besides-case-loads-581fa90348dd**

Unless otherwise specified, all online links were accessed on 13th May 2020.

Arndt, R. Z. (2018). How third parties harvest health data from providers, payers and pharmacies. *Modern Healthcare*, 48(15), p. 18. Available at: **https://search.proquest.com/docview/2024121781**.

Barr, C., Kommenda, N., McIntyre, N. and Voce, A. (2020). Ethnic minorities dying of Covid-19 at higher rate, analysis shows. *The Guardian*. Available at: **https://www.theguardian.com/world/2020/apr/22/racial-inequality-in-britain-found-a-risk-factor-for-covid-19**.

Binns, R., Lyngs, U., Van Kleek, M., Zhao, J., Libert, T., and Shadbolt, N. (2018). Third Party Tracking in the Mobile Ecosystem. *WebSci '18: Proceedings of the 10th ACM Conference on Web Science*, pp. 23–31. Available at: **https://doi.org/10.1145/3201064.3201089**.

Bot, B. M., Wilbanks, J. T., and Mangravite, L. M. (2019). Assessing the consequences of decentralizing biomedical research. *Big Data and Society*, [online] 6(1). Available at: **https://doi.org/10.1177/2053951719853858**.

Carter, P., Laurie, G. T., Dixon-Woods, M., Carter, P., Laurie, G. T., & Dixon-woods, M. (2015). The social licence for research: why care.data ran into trouble. *Journal of Medical Ethics*, 41(5), pp. 404–409. Available at: **jstor.org/stable/44014080**.

Castle-Clarke, S., & Hutchings, R. (2019). *Achieving a digital NHS: lessons for national policy from the acute sector*. Available at: **https://www.nuffieldtrust.org.uk/research/achieving-a-digital-nhs-lessons-for-national-policy-from-the-acute-sector%0Ahttps://www.nhsx.nhs.uk/**.

Department of Health and Social Care. (2018). *The future of healthcare: our vision for digital data and technology in health and care*. Available at: **https://www.gov.uk/government/publications/the-future-of-healthcare-our-vision-for-digital-data-and-technology-in-health-and-care**.

Devlin, H. (2020). Why are people from BAME groups dying disproportionately of Covid-19? *The Guardian*. Available at: **https://www.theguardian.com/world/2020/apr/22/why-are-people-from-bame-groups-dying-disproportionately-of-covid-19**.

Fdez-Arroyabe, P., Fernandez, D. S., & Andres, J. B. (2019). Work environment and healthcare: a biometeorological approach based on wearables. In N. Dey, A. S. Ashour, S. J. Fong, and C. Bhatt (Eds.), *Wearable and Implantable Medical Devices: Applications and Challenges*, pp. 141–161. Available at: **https://doi.org/10.1016/B978-0-12-815369-7.00006-9**.

Ford, E., Boyd, A., Bowles, J. K. F., Havard, A., Aldridge, R. W., Curcin, V., and Sperrin, M. (2019). Our data, our society, our health: A vision for inclusive and transparent health data science in the United Kingdom and beyond. *Learning Health Systems*, [online] 3(3). Available at: **https://doi.org/10.1002/lrh2.10191**.

Fotopoulou, A (2020) Conceptualising critical data literacies for civil society organisations: agency, care, and social responsibility, Information Communication and Society. Available at: **https://www.tandfonline.com/eprint/MZPZRPPG4FPPQBGGGNAS/full?target=10.1080/1369118X.2020.1716041**

Fotopoulou, A. (2018). From networked to quantified self: Self-tracking and the moral economy of data. In Z. Papacharissi (Ed.), *A Networked Self: Platforms, Stories, Connections*. New York, USA: Routledge.

Fotopoulou, A., and O'Riordan, K. (2016). Training to self-care: fitness tracking, biopedagogy and the healthy consumer. *Health Sociology Review*, 26(1), pp. 54–68. Available at: **https://doi.org/10.1080/14461242.2016.1184582**

Giles, C. (2020). UK coronavirus deaths more than double official figure, according to FT study. *Financial Times*. Available at: **https://www.ft.com/content/67e6a4ee-3d05-43bc-ba03-e239799fa6ab**.

Hunn, A. (2017). *Survey of the general public: Attitudes towards health research*. Available at: **https://www.hra.nhs.uk/documents/1422/HRA_NIHR_general_public_omnibus_survey_2017_FINAL.pdf**.

IPSOS MORI for the Wellcome Trust. (2016). *The One-Way Mirror: Public attitudes to commercial access to health data*. Available at: **https://www.ipsos.com/sites/default/files/publication/5200-03/sri-wellcome-trust-commercial-access-to-health-data.pdf**.

Kendi, I. X. (2020a). Why Don't We Know Who the Coronavirus Victims Are? *The Atlantic*. Available at: **https://www.theatlantic.com/ideas/archive/2020/04/stop-looking-away-race-covid-19-victims/609250/?utm_source=twitter&utm_medium=social&utm_campaign=share**.

Kendi, I. X. (2020b). What the Racial Data Show. *The Atlantic*. Available at: **https://www.theatlantic.com/ideas/archive/2020/04/coronavirus-exposing-our-racial-divides/609526/**.

Li, X., Dunn, J., Salins, D., Zhou, G., Zhou, W., Schüssler-Fiorenza Rose, S. M., and Snyder, M. P. (2017). Digital Health: Tracking Physiomes and Activity Using Wearable Biosensors Reveals Useful Health-Related Information. *PLOS Biology*, [online] 15(1). Available at: **https://doi.org/10.1371/journal.pbio.2001402**.

Mann, N. (2014). The care.data programme: explicit detail on protocols and safeguards is needed. *BMJ: British Medical Journal*, Vol. 348. Available at: **https://doi.org/10.1136/bmj.g2201**.

McCartney, M. (2014). Care.data doesn't care enough about consent. *BMJ: British Medical Journal*, Vol. 348. Available at: **https://doi.org/10.2307/26514579**.

NHS. (2017). *Harnessing technology and innovation*. Available at: **https://www.england.nhs.uk/five-year-forward-view/next-steps-on-the-nhs-five-year-forward-view/harnessing-technology-and-innovation/**.

NHS. (2019). *The NHS Long Term Plan*. Available at: **https://www.longtermplan.nhs.uk/**.

Pateraki, M., Fysarakis, K., Sakkalis, V., Spanoudakis, G., Varlamis, I., Maniadakis, M., and Koutsouris, D. (2020). Biosensors and Internet of Things in smart healthcare applications: challenges and opportunities. In In N. Dey, A. S. Ashour, S. J. Fong, and C. Bhatt (Eds.), *Wearable and Implantable Medical Devices: Applications and Challenges*, pp. 25–53. Available at: **https://doi.org/10.1016/B978-0-12-815369-7.00002-1**.

Pueyo, Tomas. (2020). Coronavirus: Why You Must Act Now. *Medium*. Available at: **https://medium.com/@tomaspueyo/coronavirus-act-today-or-people-will-die-f4d3d9cd99ca**.

PwC. (2016). *Use of wearables in the workplace is halted by lack of trust*. Available at: **https://pwc.blogs.com/press_room/2016/06/use-of-wearables-in-the-workplace-is-halted-by-lack-of-trust-pwc-research.html**.

Richardson, S. and Spiegelhalter, D. (2020). Coronavirus statistics: what can we trust and what should we ignore? *The Guardian*. Available at: **https://www.theguardian.com/world/2020/apr/12/coronavirus-statistics-what-can-we-trust-and-what-should-we-ignore**.

Sanghera, S. (2020). 'Coronavirus and ethnicity: black and Asian NHS medics on the front line'. *The Times*. Available at: **https://www.thetimes.co.uk/edition/news/coronavirus-and-ethnicity-black-and-asian-nhs-medics-on-the-front-line-dhznrc3kc**.

Schairer, C. E., Rubanovich, C. K., & Bloss, C. S. (2018). How Could Commercial Terms of Use and Privacy Policies Undermine Informed Consent in the Age of Mobile Health? *AMA Journal of Ethics*, 20(9), pp. 864–872. Available at: **www.amajournalofethics.org**.

Sharon, T., & Lucivero, F. (2019). Introduction to the Special Theme: The expansion of the health data ecosystem – Rethinking data ethics and governance. *Big Data and Society*, Vol. 6. Available at: **https://doi.org/10.1177/2053951719852969**.

Sheather, J., & Brannan, S. (2019). Patient confidentiality in a time of care.data: The benefits of unlocking patient data are enormous; hopefully the risks can be managed. *BMJ: British Medical Journal*, 347(7935). Available at: **https://www.jstor.org/stable/43513265**.

Skovgaard, L. L., Wadmann, S., & Hoeyer, K. (2019). A review of attitudes towards the reuse of health data among people in the European Union: The primacy of purpose and the common good. *Health Policy*, Vol. 123, pp. 564–571. Available at: **https://doi.org/10.1016/j.healthpol.2019.03.012**.

Statista. (2020). *Connected wearable devices worldwide 2016-2022*. Available at: **https://www.statista.com/statistics/487291/global-connected-wearable-devices/**.

Stockdale, J., Cassell, J., & Ford, E. (2019). "Giving something back": A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland [version 2]. *Wellcome Open Research*, Vol. 3, p. 6. Available at: **https://doi.org/10.12688/wellcomeopenres.13531.2**.

Tsang, T. K., Wu, P., Lin, Y., Lau E. H. Y., Leung, G. M. and Cowling, B. J. (2020). Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. *The Lancet*, Open Access. Available at: **https://doi.org/10.1016/S2468-2667(20)30089-X, https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30089-X/fulltext**.

UK Data Service. (n.d.). *Why share data*? Available at: **https://www.ukdataservice.ac.uk/manage-data/plan/why-share.aspx**.

Understanding Patient Data. (2018). *Public attitudes to patient data use: A summary of existing research*. Available at: **https://understandingpatientdata.org.uk/sites/default/files/2019-05/Public attitudes key themes 0.pdf**.

Wu, M., & Luo, J. (2019). Wearable Technology Applications in Healthcare: A Literature Review. *Online Journal of Nursing Informatics (OJNI)*, [online] 23(3). Available at: **https://www.himss.org/resources/wearable-technology-applications-healthcare-literature-review#_ENREF_22**.

Zadeh, P. E. & Tremblay, M. C. (2016). A review of the literature and proposed classification on e-prescribing: Functions, assimilation stages, benefits, concerns, and risks. *Research in Social and Administrative Pharmacy*, 12:1, pp. 1-19.