

SPECIAL ISSUE PAPER

The cross-over of statistical thinking and practices: A pandemic catalyst

Andrew D. Garrett Scientific Operations, ICON Clinical
Research UK, Reading, UK**Correspondence**Andrew D. Garrett, Scientific Operations,
ICON Clinical Research UK, Green Park,
Reading, UK.
Email: andrew.garrett@iconplc.com**Abstract**

Written during the SARS-CoV-2 pandemic, and in recognition of Andy Grieve, the polymath, this article looks at an eclectic mix of topics where statistical thinking and practices should transcend typical dividing lines—with a particular focus on the areas of drug development, public health and social science. The case is made for embedding an experimental (or quasi-experimental) framework within clinical practice for vaccines and treatments following their marketing authorisation. A similar case is made for public health interventions—facilitated by pre-specification of effect size and by the greater use of data standards. A number of recommendations are made whilst noting that progress is being made in some areas.

KEYWORDS

COVID-19, master protocols, NPIs, pharmacovigilance, regression discontinuity design, stepped-wedge cluster randomisation design, vaccines

1 | INTRODUCTION

Despite nailing his colours to the mast as a Bayesian, Andy Grieve has always struck me as a bit of a polymath. Someone with a wide range of interests both within and outside statistics. Indeed, both within drug development and outside of it—including music, literature and the arts. This truly manifested itself in 2006 when Andy was part of the Royal Statistical Society's (RSS) team that, alongside notable others including Stephen Senn, reached the final of *University Challenge*, *The Professionals*.

My first interactions with Andy were in response to a letter that I had written to *RSS News and Notes* expressing an interest in setting up a Bayesian Discussion Group. Andy kindly responded with his support, quoted Cervantes's Don Quixote *tilting at windmills*, and as a result of some advice, Oliver Keene (from GSK, who had also responded positively) and I formed a PSI Bayesian Special Interest Group. Indeed, Andy was the guest speaker at our inaugural meeting. He was also RSS President from 2003 to 2005, a time when I also served on RSS Council, and we subsequently worked together on the RSS's Working party on statistical issues in first in man studies (2006–2007) producing an article,¹ with lead author Stephen Senn and others. As RSS President, Andy represented the broad church of statistics and statisticians with gravitas whilst also doing an exemplary job in raising the profile of pharmaceutical statistics and the pharmaceutical statistician. Most recently, our paths crossed for a short period in 2016–2017 at ICON Clinical Research where Andy was employed as an expert statistician, notably in adaptive design, in the Consultancy team.

So, in recognition of Andy Grieve, the polymath, the aim of this article is to look both within and beyond drug development and provide a somewhat eclectic mix of topics where other areas (notably social science and public health) could potentially learn from drug development whilst also identifying some areas where drug development could learn

from the practices of others. Written during the midst of the SARS-CoV-2 pandemic, the challenges observed provide the backdrop for many of the observations. The article is structured into five main sections. The first two sections, “quantification and pre-specification” and “meta data, standardisation and validation” focus on what others could learn from good statistical practices in drug development while the final three sections, “master protocols,” “quasi-experimentation and less used designs” and “post-marketing pharmacovigilance” consider areas where drug development could benefit from some fresh thinking and challenge. In each section a recommendation(s) is made with the aim of challenging current thinking and practices.

2 | QUANTIFICATION AND PRE-SPECIFICATION

The SARS-CoV-2 pandemic has been dominated by an array of daily statistics relating to infection detection (testing) and prevalence, hospitalisation and mortality—and how these have been impacted by various non pharmaceutical interventions (NPI) including, *in extremis*, lockdowns. It has also been dominated by vaccine efficacy and the subsequent impact on those daily statistics, as phase III clinical trial results emerged within an unprecedented 12 month development period and emergency use authorisation (EUA) was granted for the first COVID-19 vaccine. It had earlier been reported (12 September 2020) in the Guardian newspaper that 321 COVID-19 vaccines were in development at the time with 32 of those in clinical trials. It was stated that the typical success rate for a vaccine was 20% for those in clinical trials and 7% for those in earlier development. A swift calculation over breakfast pointed to a > 99% chance, either way, of finding at least one successful vaccine [$= 1 - (1 - 0.07)^{321}$ or $p = 1 - (1 - 0.2)^{32}$], if the past were indeed a good predictor of the future, although the future can often be a long way off. As it turned out, the first vaccine to report its phase III data demonstrated unexpectedly high efficacy and cleared the regulatory efficacy bar (criteria) with ease. Andy Grieve will have been pleased to find that this first vaccine to receive EUA had been developed by one of his former employers (Pfizer) and even more pleased to have seen a Bayesian statistical analysis undertaken to support the EUA and subsequent full approval.

Now, the FDA² vaccine guideline to prevent COVID-19, the disease, is prescriptive in specifying the primary efficacy criteria that dictate vaccine success in a placebo-controlled efficacy trial. Specifically, that the point estimate for the relative effect is $\geq 50\%$ and that the corresponding lower confidence limit is $> 30\%$. These criteria have historical precedent in vaccine development. (The guideline went further in anticipating the future need for active controlled studies, specifying a non-inferiority margin of 10%, that is a lower confidence limit $> -10\%$.) In this respect, Vaccine Efficacy (VE) is defined as $VE = 100 \times (1 - IRR)$, where IRR is the ratio of infection rate in the vaccine group to the control group—in effect, the infection incidence in the experimental vaccine group relative to the control group, expressed as a percentage. The guideline points to the need for alpha-adjusted confidence intervals taking into account multiplicity in terms of interim analyses and multiple endpoints.

In this respect, the guidance has all the characteristics one would expect from an ICH E9 *Statistical principles for clinical trials*³ approach, including pre-specification, sample size justification, and alpha control. Indeed, the leading development programs were all remarkably similar as a result of the prescriptive regulatory approach with some nuanced differences in terms of endpoint definition, notably in terms of symptom criteria for symptomatic infection. These phase III protocols were substantial in terms of the number of participants enrolled with over 40,000 in the Pfizer/BioNTech (C4591001) study,⁴ over 30,000 in the AZ/Oxford (AZD1222) study,⁵ and over 44,000 in the Janssen (ENSEMBLE) study,⁶ with each providing safety databases orders of magnitude larger than would typically be found in drug development for a treatment—a point returned to later in the article.

Contrast this with the investigation of so-called non-pharmaceutical interventions (NPIs) such as face masks, social distancing, school closures, and so on. Ferguson⁷ describes two NPI strategies: the first is suppression, with the aim of reducing R (the reproduction number, defined as the average number of secondary cases each case generates) to below 1 while the second is mitigation, with the aim of reducing the impact on health (i.e., severe disease) by reducing R, but not necessarily to below 1. It is interesting to note that there is no mention of an equivalent effect size like VE. Now, perhaps the point about NPIs, in the era of SARS-CoV-2, is that the effectiveness is expected to be modest in isolation (in contrast to the vaccines) and it is the cumulative effect of multiple NPIs that is expected to lead to material impact—including as an adjunct to vaccination. It is also the case that an NPI is not necessarily a definitive intervention in the same way as a vaccine. Mask wearing and social distancing can all be undertaken to degrees which blurs the intervention lines. Of course, the impact of interventions can change over time, but the same can be said of vaccines as new variants dominate and antibody levels reduce over time. Given that NPIs were first implemented prior to vaccines

being available, public policy would surely have benefited (and would benefit in the future) from giving greater thought to the size of NPI effects that warrant introduction and withdrawal.

Part of the issue is that most evidence for NPIs in SARS-CoV-2 have been generated from observational or mechanistic studies. A review of cloth face masking⁸ identified just two reported randomised controlled trials (RCT) comparing face mask intervention to no mask control with respect to the spread of COVID-19, both of which pre-specified the size of effect. One ongoing study was also identified.

One of the two reported studies randomised 6024 participants in Denmark⁹ to either wear a surgical mask outside home or not. The primary outcome was SARS-CoV-2 infection after 1 month (confirmed with a polymerase chain reaction [PCR] antibody test) in eligible participants who reported being outside home for >3 h/day without occupational mask use. The protocol included a standard power calculation (80% power, two-sided 5% significance level) to detect a 50% reduction in the assumed underlying infection rate of 2%—in effect, mirroring the FDA vaccine guideline.

The other study¹⁰ was a randomised cluster design of 342,000 adults in 600 villages in rural Bangladesh. Villages were randomised to mask wearing (intervention) or not (control) with primary outcome being symptomatic seroprevalence. The protocol detailed the sample size assumptions with regard to villages (600), households per village (200), eligible persons per household (2), inter-cluster correlation (0.02) and so on, with a minimum detectable effect of 9.2% stated likely akin to VE, albeit governed by budget and logistics constraints. However, power was not obviously stated and the sample size section is not easy to follow.

So, although based on very few RCTs, the observation holds true that pre-specification of effect size tends to be the preserve of randomised studies, where the sample size is determined and then data are prospectively generated in light of the objectives (subject to the budget and logistics constraints illustrated above) rather than for observational studies that typically report on retrospective data—that is, those data that are available. Indeed, this has been my observation elsewhere, specifically in relation to the social sciences.

Reviewing research proposals on the UK Statistics Authority's Research Accreditation Panel (RAP),¹¹ and previously on the Administrative Data Research Network's (ADRN)^{12,13} approvals panel, has provided exposure to hundreds of projects requesting access to, and linkage of, de-identified government data.¹² Projects must meet strict criteria including the demonstration of public benefit. In the case of RAP, the legal gateway to access data is the Digital Economy Act 2017¹¹ which allows for accredited and approved researchers to access data for research and statistical purposes. As such researchers request access and linkage of data that have already been generated—either as a result of government providing a service (administrative data, including tax records) or conducting a survey (including the COVID-19 Infection Survey, the Longitudinal Survey and the Census). In the case of administrative data, it is often associated with the phrase Data = all, in the sense that the data are not a sample from a population but “are its entirety.”¹⁴ (Perhaps the closest analogy in drug development would be the integrated summaries of efficacy and safety that are prepared for drug registration that include all the individual data from all of the trials conducted for an experimental drug.) Although prospectively collected in a highly organised manner, the Data = all concept is somewhat similar for the UK Census. Coming from a background of drug development it has been interesting to compare and contrast the different approaches to statistical analysis planning. For both ADRN and RAP, a methods guidance document was developed by the respective approvals panels that aimed to bring in some of the good statistical practices seen elsewhere.³ These methods documents request that researchers specify the planned methods of analysis and articulate the outcome or dependent variable(s). In particular that they indicate the starting point for the modelling process—including the independent or explanatory variables that comprise the initial set to be considered—and the type of model to be employed. This aims to control p-hacking (that is, multiple uncontrolled attempts to find statistical significance) to a certain extent but is much weaker than the rigour seen in late-stage drug development. The researchers are asked to address how potential selection/causal bias will be addressed (e.g., by including a control group with information on how this control group will be created). Methodology references are requested for non-standard methodology. Importantly, the researchers are asked to explain how the methodological approach will answer the research question(s) and demonstrate public benefit.

Since a researcher has no control over the data-generating process, it is important that pre-specification requirements are proportionate. Indeed, it is not unusual for researchers to iterate and request other data sets or other variables once limitations emerge on closer inspection when data access is granted. Baldwin *et al.*¹⁵ have proposed solutions including declaring prior access to data and the use of synthetic and hold-out samples. In this respect, pre-specification has not advanced in the methodology guideline to stating the size of effect that would be regarded as meaningful. In general, the focus of these proposals is on statistical significance determining whether a finding is meaningful or not, which in general has causal limitations within a framework that lacks randomisation. That being said,

there are some proposals that contain quasi-experimental methods that are both creative and appealing and which are perhaps over-looked in drug development. We will return to these later in the article (section 5).

Overall, it appears that randomisation is one of the key drivers of pre-specification and the objectivity that results when it comes to analysis and interpretation—although regulation clearly has a key role in drug development. It follows that public health policy and some areas of social science would benefit from greater rigour at the planning stage.

Recommendation 1. *Public policy should have clearer guidance on what size of effect warrants the introduction of a policy, including benefit/risk assessment, and size of effect should be estimated within, where possible, an experimental framework.*

3 | META DATA, STANDARDISATION, AND VALIDATION

Bacon and Goldacre¹⁶ detail their frustrations in accessing health data in the UK, describing data sets that, without warning, change location or structure (without documentation), are simply impossible to locate, or require a CAPTCHA. These hinder the ability to automate downloads and lead to excessive manual intervention. There is also the issue of similar datasets with unexplained differences. Indeed, the same DataLab team based at Oxford University published one of the largest studies exploring factors associated with COVID-19-related hospital deaths using linked electronic health records from 17 million adult National Health Service (NHS) patients during the early months (1 February 2020 to 25 April 2020) of the pandemic.¹⁷

What Bacon and Goldacre highlight is the need for better meta data—the data about the data. This has been a hot topic during the SARS-CoV-2 pandemic, when reporting something as seemingly straightforward as the number of deaths by date and cause, proved anything but that. For such important data, it was remarkable how long it took for the meta data to reveal themselves and for the implications and limitations to be noted. For instance, it took until 12 August 2020 for Public Health England [subsequently replaced by the UK Health Security Agency] to modify its definition of deaths in relation to COVID-19 after it became apparent that deaths were originally being counted regardless of time since a positive PCR test.¹⁸ This was, of course, less of an issue in the early weeks and months of the pandemic, but as time progressed and PCR testing increased substantially, this definition began to pick up deaths that were clearly unrelated to much earlier positive test results. The new narrower measure became “A death in person with a laboratory-confirmed positive COVID-19 test and died within (equal to or less than) 28 days of the first positive specimen date.” (A lesser-known broader definition was also created using 60 days or after 60 days, only if COVID-19 is mentioned on the death certificate.) This new definition reduced the cumulative number of COVID-19 deaths by 13% (5377/42,072) overnight, with a noticeable impact (separation from previous measures) from late April 2020 onwards. Earlier there had also been confusion¹⁹ between date of death and date of registration of death, with the latter clearly lagging the former. The former, provided by Public Health England, excluded deaths outside of the NHS with the latter, produced by the Office for National Statistics (ONS), included all certified deaths (usually within 5 days of death).

Another example was hospitalisation and the important distinction between hospitalised for COVID-19 versus hospitalised with COVID-19. The Oxford Centre for Evidence-Based Medicine highlighted the issue with public briefings, noting the 13 April change to “with confirmed COVID-19” and calling for clearer definitions.²⁰ (It is also interesting to read from the NHS England website, the number of submissions containing erroneous data from various hospital trusts. Data may = all, but that clearly does not mean all = correct and current.)

For statisticians working in drug development, such ambiguity is an anathema. The operating procedures, work instructions and templates that describe “the why” and “the how” are all requirements of working in a regulated industry. These direct team members to document what they do and when they do it. Regulatory auditors will trace data points from datasets to statistical outputs via statistical analysis plans, table shells, derived data specifications and in some cases programs themselves. Version control and change management underpin quality and reproducibility of results. Standardisation is at the heart of data processing and analysis with the Clinical Data Interchange Standards Consortium (CDISC)²¹ providing a suite of standards supporting the clinical and non-clinical end-to-end processes. These CDISC standards are required by the FDA (and the Japanese PMDA, to a lesser extent) but the global nature of drug development means that they have become the *de facto* standard. CDISC provides naming conventions for datasets and variables, dataset structure, variable lengths and formats, data derivation rules and has extended into therapeutic area standards with disease-specific meta data. Data flows from the Study Data Tabulation Model (SDTM), for organising and formatting data, to the Analysis Data Model (ADaM), that defines data and meta data standards. There

are also export standards, so that the regulators can easily import data and undertake independent analyses of the data. The Clinical Data Acquisition Standards Harmonisation (CDASH) establishes a standard way to collect data across studies and sponsors. Importantly CDISC standards are interoperable with standards of reference such as MedDRA and HL7.

Now consider a related issue. Neil Ferguson set some alarm bells ringing with a series of tweets (@neil_ferguson) on 22nd March 2020 stating: “I’m conscious that lots of people would like to see and run the pandemic simulation code we are using to model control measures against COVID-19. To explain the background - I wrote the code (thousands of lines of undocumented C) 13+ years ago to model flu pandemics...” and later “I am happy to say that Microsoft and Github are working with Imperial and MRC to document, refactor and extend the code to allow others to use without the multiple days training it would currently require (and which we don’t have time to give)...”

Statisticians and programmers working in drug development understand the need to control the risks associated with having a single-point of failure with independent checking at the heart of the process. Of course, there is no indication that Ferguson’s code contained errors, but that is not the point. In drug development, derived datasets and outputs are often independently programmed from detailed specifications and compared electronically to systematically identify and investigate differences. One small coding error can have profound implications for aggregated data—much more so than for a data transcription error at an investigator site that impacts just one data record. Such processes and procedures have been developed over many years as a result of guidance that began with ICH *Good Clinical Practice*²² that led to more specific guidance such as ICH E9 and also the US computer systems validation guideline 21CFR11.²³ Such processes and procedures do not guarantee error-free statistical modelling, but they allow others to navigate the data sets generated and perform independent checks, and they also allow corrections to be made in a timely and ordered manner. Importantly there is an audit trail that provides re-assurance and builds trust. Similar standards should apply elsewhere when computer programs have the potential to have public policy impact.

Now one might argue that the Ferguson code was directed at simulation rather than data analysis *per se*, and that even in drug development, simulation code is frequently developed by individual experts rather than procedurally with independent checks. That is to some extent true although there are key differences. Validated simulation software does exist for clinical trial design while the outputs from such simulations are mostly directed at assessing the operating characteristics of a trial design, that will subsequently generate prospective data to inform decision making. In cases where the simulation itself plays an important part in regulatory approval, then the usual rigorous standards apply. Smith and Marshall²⁴ discuss the importance of simulation plans to define the data generating process, the methods and the decision criteria in relation to clinical drug development.

Recommendation 2. *The UKSA, MHRA and UKHSA should jointly develop data standards for health related data including the storage and documentation of data sets using these standards to facilitate the timely provision of evidence to support public policy.*

Recommendation 3. *Modelling (including simulation) used to support the implementation of NPIs should be subject to standards similar to that expected for the implementation of PIs, in terms of documentation, independent checking and validation.*

4 | MASTER PROTOCOLS

The UK’s RECOVERY trial,²⁵ led by the University of Oxford, is one of the most important and impactful clinical trials of recent years, yet it was Academia that stepped up to the challenge of initially re-purposing drugs to treat COVID-19 and evaluating them in an experimental setting. It must have been an organisational feat to set it up so quickly (March 2020, within 6 weeks of funding). In many ways the RECOVERY Trial could be described as Back to the Future. On the one hand, it had characteristics that indicated a return to simpler times with its short protocol and short eCRF. It is a randomised, parallel group study, with some interim analyses but it is not double-blind and collects minimal data. On the other hand, it is a platform design with a master protocol that has multiple treatments (with concurrent controls) and it can add promising treatments, including newly developed drugs or drop ineffective ones. Its biggest breakthrough was to demonstrate the effectiveness of dexamethasone²⁶ in June 2020 (reducing 28-day mortality among those who were receiving either invasive mechanical ventilation or oxygen alone at randomisation). Subsequently in February 2021, it demonstrated the effectiveness of tocilizumab,²⁷ a re-purposed intravenous drug used to treat

rheumatoid arthritis, while in June 2021, it demonstrated the effectiveness of the experimental monoclonal antibody combination casirivimab/imdevimab²⁸ in patients hospitalised with severe COVID-19, who had not mounted their own natural antibody response. Equally importantly, it confirmed in June 2020 that the anti-malarial, Hydroxychloroquine²⁹ (HCQ), was ineffective. HCQ had been the subject of much debate³⁰ with observational data from early in the pandemic pointing to its effectiveness, initially from China, but also extending to other regions and countries including the US and Europe. The delay in studying HCQ within an experimental framework inevitably became a diversion that wasted resources at a key time. (Data provenance also came into question in relation to HCQ. The Surgisphere study³¹ published (and then retracted) in the *Lancet* showed HCQ to be ineffective and confirmation bias may have led to the lack of critical review. It led the *Lancet* to change its review process with a greater focus on provenance and proof of data sharing agreements—something other do well, as described earlier in section 2.) In this respect, RECOVERY certainly strengthened the case for randomisation by highlighting the unreliable nature and limitations of observational data, particularly in relation to causal modelling.

So, should RECOVERY be viewed as the future of drug development? Firstly, consider simplification. Research from Tufts³² shows that from 2001–05 to 2011–15, the distinct number of procedures (blood draws, biopsies, scans, rating scale completion etc.) in a typical Phase III study increased by 59%. The number of eligibility criteria per study increased by 61% and the number of scientific endpoints used to judge success increased by 86% while the number of planned patient visits increased by 25%. It is not hard to imagine how this increased burden on patients and investigators has acted as a disincentive, such that finding the right sites with the right patients continues to be the number one challenge facing drug development. The Tufts data also show that 54% of investigators conducted just one trial and then opted out of further work. In contrast the RECOVERY trial was set-up quickly with an emphasis on collecting just key data and this point should not be lost on the pharmaceutical industry. Drug development project teams need to be held much more accountable for, and challenged more to justify, each eligibility criterion and procedure that they propose to include in a protocol, with each one truly earning its place.

Secondly, consider collaboration and coordination.³³ Drug development, although a massive team effort across Sponsors, Investigators, Patients, Regulators and frequently Contract Research Organisations (CRO), it is usually conducted separately from competing drug development programs. There are, of course, commercial reasons for this including pricing and reimbursement, and the desire to be first, or one of the first to market. However, it is inefficient in terms of use of control groups and also in terms of learning—understanding early how an experimental treatment compares with others in development. Consider two experimental treatments sharing a concurrent control arm (with allocation 1:1:1). This requires 25% fewer patients overall, compared to two separate studies (each with allocation 1:1), since only a third of patients now get control (often placebo alongside standard-of-care) versus a half of patients. In challenging research areas such as Alzheimer's disease, where successes have been few and far between, surely a master protocol approach would have provided a more efficient use of resources (across sponsors, investigators, and patients) over many years. COVID-19 vaccine development has meant that more than 50,000 participants will have received control while the comparative efficacy of leading vaccines has been open to debate in the absence of head-to-head data. Given the record speed observed with vaccine development, it would be somewhat churlish to be critical here, but if the Regulators had insisted upon a master protocol, then the evidence base would now be much richer in terms of comparative efficacy and safety. As new variants of SARS-CoV-2 emerge and placebo-controlled trials become increasingly hard and unethical to implement, then surely a master protocol becomes the most efficient way of generating the evidence required to optimise vaccination. There is of course the issue of concurrent control for unbiased treatment comparisons, but this can be overcome by stratified analyses that account for the timing (that is, implementation) of randomisation amendments.

Thirdly, consider integration—embedding clinical trials as a treatment option. RECOVERY has more than 45,000 participants and was essentially embedded into the clinical practice at participating sites, subject to the usual patient informed consent. This is an appealing model that aims to make clinical research, including the benefits, available to a wider and more diverse group of participants. It points to a healthcare model that is constantly learning and gathering knowledge around comparative effectiveness, optimisation and safety—in an experimental setting that is underpinned by a causal model.

In relation to SARS-CoV-2, master protocols would create an efficient framework to develop new and modified vaccines, including in a non-inferiority setting, particularly against new variants and would facilitate the need to move towards vaccine optimisation. Indeed, with regard to optimisation, the UK Vaccine Task Force and the National Institute for Health Research funded the Com-COV study³⁴ that enrolled participants in February 2021 to compare heterologous (different vaccines) with homologous (same vaccine) two dose schedules of the Pfizer/BioNTech and AstraZeneca/

Oxford vaccines in terms of immunological response. In effect a 2 (first dose) \times 2 (second dose) \times 2 (28 day vs. 84 day dosing interval) factorial design. Such studies are designed to begin to answer important questions around vaccine optimisation in terms of mixing vaccines and dosing intervals. A second study³⁵ (Com-COV2) was subsequently undertaken that incorporated randomised 2nd doses of either Moderna, Novavax or the homologous vaccine using a stratified design (non-randomised first dose Pfizer/BioNTech or AZ/Oxford).

Master protocols for drug development would require a cultural shift and perhaps a commercial risk-share type of model but would be more efficient and expose fewer patients overall to placebo, whilst making better use of their data. In many respects, it is an ethical imperative. Even if viewed as too challenging commercially for initial marketing authorisation purposes, there should be scope for master protocols to be used subsequently to study both broader populations and special populations. For example, master protocols for paediatric populations, where studies are often undertaken later according to a paediatric investigation plan (PIP).

Recommendation 4. *Regulatory agencies and funding bodies should encourage the greater use of master protocols to promote the judicious use of trial participants randomised to control and to limit their number. Comparative efficacy and safety data would also enhance decision making at various levels.*

5 | QUASI-EXPERIMENTATION AND LESSER USED DESIGNS

When reviewing prospective research projects directed at accessing and linking de-identified government data, predominately in the area of social science, it is interesting to read about the various approaches proposed. Some are novel and take advantage of public policy implementation that can be regarded as having a random component. One particular area is regression discontinuity designs.

5.1 | Regression discontinuity designs

Regression discontinuity designs are relatively common in social science but almost unheard of in drug development. As described by Gelman et al.³⁶ such designs can be thought of as a randomised block design with complete confounding of the blocking variable with treatment group. Consider an example. Vaccines to prevent COVID-19 disease are rolled out within a country by age, during a specific time window. In fact, in the UK, roll-out began on 8 December 2020 in Coventry when Margaret Keenan, aged 90 years old, became the first person in the world, not just the UK, to receive a vaccine as part of a mass vaccination program.³⁷ Now in this respect vaccination (treatment) is, within this first time period, confounded with age (block), with the priority being those aged over 80 years, front-line health care workers and nursing home workers. A comparison of those vaccinated with those unvaccinated with regard to outcome (typically one of infection, hospitalisation or death) is therefore also a confounded comparison of age, leading to an inherently biased comparison. As such there is no possibility of obtaining either balance within a block (age, say) or overlap (of treatment within block). The causal trick that is introduced in a RDD is to consider those people on either side of the age cut-off. The argument is as follows. The elderly population around 80 years will be very similar in the sense that someone aged 79 years and 364 days, will not be inherently different to someone aged 80 years and 1 day. The fact that one person was eligible and another ineligible could be regarded as fairly random given that birth varies within a gestation window. So, if the analysis is limited to a narrow range around the threshold, then participants may be regarded as similar. Indeed, Bermingham et al.³⁸ undertook such an analysis comparing the age group 80–84 years with the group 75–79 years, finding a 71% reduced risk of death with vaccination. Gelman provides some examples from the field of education and, for instance, the Bank of England used a discontinuity design to assess the impact of COVID-19 lockdowns on business activity,³⁹ while in Japan it was used to assess the impact of COVID-19 school closures.⁴⁰

RDDs are undoubtedly inefficient compared to RCTs, and also biased if the model assumptions do not hold.⁴¹ In this respect the method is not a substitute for randomisation—rather these designs have the potential to tease out causal relationships (information) in high volume observational settings that would otherwise remain uncovered.

One area where RDD could be informative in drug development is with regard to patient eligibility. In large trials with high screen failure rates, where screened patients are excluded on the basis of a strict cut-off for a baseline continuous variable (and not re-screened), there may be potential to follow ineligible patients with respect to a key outcome.

5.2 | Stepped-wedge cluster randomisation designs

Stepped-wedge designs, in the simple sense, are actually very common in drug development, since the key feature of the design is to delay treatment commencement for trial participants in a series of steps. In this respect, a typical phase III trial that randomises patients to experimental treatment or placebo (in a blinded fashion) for a fixed period, after which all patients enter a second (or extension) period, where all receive experimental treatment is a stepped-wedge design, even if seldom analysed as one. Such an approach is commonly used to expand the safety database to gather more data on the experimental treatment—both on more patients and for a longer duration. (The primary efficacy analysis using data from the first period only.) The design also encourages participation since those randomised to placebo in the first period are essentially guaranteed experimental treatment later—and important feature particularly in rare diseases where few, if any, other treatment options are available. (Note, those assigned to placebo would typically receive standard of care alongside placebo.) However, stepped-wedge cluster randomisation designs were devised to tackle other issues—namely logistics in service delivery type interventions. As described by Hemming et al.⁴² “The design involves random and sequential crossover of clusters from control to intervention until all clusters are exposed.” Initially no cluster is exposed to the intervention, and then at regular intervals (i.e., steps) a cluster or groups of clusters receive the intervention until all clusters have received the intervention. The roles of logistical and resource constraints are key in the pragmatic use of such designs and early examples come from developing countries where the availability of health care professionals (HCP), funding and geographic spread of the population means that an intervention has to be rolled out over a prolonged period of time. However, we have seen that such logistical challenges are not always limited to developing countries and the SARS-CoV-2 vaccine roll-out again provides an example of where such a design could have been used to generate both effectiveness and safety data within an experimental (that is, causal) setting. Care homes would have provided perhaps the most obvious experimental unit here, where HCPs attended and vaccinated the willing *en masse*, over a period of time.

Stepped wedge cluster designs outperform standard cluster designs when intra-cluster correlation is larger or the clusters are large,³⁹ although temporal trend which is a confounding factor has to be taken into account.

Recommendation 5. *Once a drug or vaccine receives marketing authorisation, greater thought needs to be given to additional research that can be efficiently conducted and how this can be embedded within an experimental framework, including using quasi-experimental methods.*

6 | POST-MARKETING PHARMACOVIGILANCE

The SARS-CoV-2 vaccine development programs have not only put vaccine efficacy under the spotlight but also safety. Although the phase III vaccine programs were sized on the basis of clinical efficacy (symptomatic infection) they provide substantial safety databases to compare each SARS-CoV-2 vaccine to their control (in most cases, placebo). For instance, a 40,000 participant trial, with a 1:1 randomisation, would generate safety data for circa 20,000 participants on the experimental vaccine. Adopting the WHO⁴³ classification for adverse drug reactions (very common event $\geq 1/10$; common event $\geq 1/100$ and $<1/10$, uncommon event $\geq 1/1000$ and $<1/100$; rare event $\geq 1/10,000$ and $<1/1000$; and, very rare event $<1/10,000$), it is straightforward to determine that with $n = 20,000$ participants receiving an experimental vaccine, then the probability (p) is 0.86 (86%) of observing at least one event of a specific type, with a true incidence of $r = 1/10,000$ (the threshold between rare and very rare event), where $p = 1 - (1 - r)^n$. Such a calculation takes no account of multiplicity, and the type of event that could be recorded is, in effect, unbounded. However, it is convenient for illustrative purposes. Vaccines tend to record few serious adverse events related to treatment and typical adverse events are recorded around the time of dosing (or within a day or two). Expected adverse events include mild injection sites reactions resulting from intra-muscular delivery—but since vaccines are also given to large populations, very rare but related, serious adverse events can still lead to a substantial number of occurrences (1000 events in 10 million exposed). Indeed, despite the large SARS-CoV-2 vaccine programs, potential drug reactions did emerge beginning with two NHS workers with anaphylactoid reactions in the UK who received the Pfizer/BioNTech vaccine in December 2020. Interestingly the Pfizer/BioNTech protocol⁴⁴ used to support MHRA approval excluded those with a “History of severe adverse reaction associated with a vaccine and/or severe allergic reaction (e.g. anaphylaxis) to any component of the study intervention(s)” and the resulting UK patient leaflet stated that the vaccine should not be given to individuals who are allergic to the active substance or any of the other active ingredients. As a result of these occurrences the

MHRA moved quickly to strengthen their guidance. Subsequent attention turned to potential “blood clotting” events, specifically sinus vein thrombosis and the AstraZeneca/Oxford vaccine in March 2021,⁴⁵ with further investigation leading to age-related modification of the vaccine roll-out in the UK and elsewhere. This was followed by cases of myocarditis reported in April 2021 and the Pfizer/BioNTech vaccine, and thromboembolic events and the Janssen vaccine.

Post-marketing signal detection requires investigations that go both broad and deep. It adopts a broad approach by comparing events reported for an experimental drug (captured from various diverse sources) with those recorded in reference databases (e.g., the FDA’s Sentinel database)⁴⁶ to estimate excess events of a particular type. It adopts a deep approach by considering individual events in detail, evaluating temporal relationships (to dosing), underlying conditions and medications to treat those conditions, age related hepatic and renal functioning, and so on. It can also revert to the clinical trials (previous and ongoing) and re-examine the data for specific signs of an event that have been subsequently observed post regulatory approval. The fundamental challenge with signal detection however is the unreliable reporting (and under-reporting) of data from various sources (the numerator), establishing the denominator and establishing a casual framework from observational data that disentangles the various potential confounding effects. The IMI WEB-RADR⁴⁷ initiative investigated whether it was possible to find rare events—and find these earlier than in other systems—to alleviate the potential under-reporting of events. This initiative evaluated social media as a source—notably Twitter (4.2 million tweets) and Facebook—in what could be described as a very broad approach. The conclusion⁴⁴ was that both “had very low value in the given context” and it was more strongly stated that such an approach “has the potential to negatively impact Signal Detection systems.” The challenge here is clear, to identify the drug in question from a social media post, to identify an event, to assign the event to the drug, to exclude events where the drug is being given to treat the event, and to remove duplicates.

That the use of high-volume data collected for a different purpose becomes a distraction in terms of signal detection should not come as a surprise since it is well established⁴⁸ that “without taking data quality into account, population inferences with Big Data are subject to a Big Data Paradox: the more the data, the surer we fool ourselves.” Furthermore, without the control of probabilistic sampling, the estimation error relative to the benchmark, $1/\sqrt{n}$, where n is the size of the dataset, increases with \sqrt{N} , where N is the population size.⁴⁸ To illustrate the point using an example from political science, Meng⁴⁸ notes that self-reporting data accumulated on 1% of US eligible voters (roughly 2,300,000 voters) would have had the same mean squared error as a simple random sample of 400 voters, when estimating the proportion of voters selecting Donald Trump in the 2016 presidential election.

Consider, for instance, the estimates of SARS-CoV-2 infection prevalence that are obtained from PCR testing versus the ONS’s COVID-19 Infection survey⁴⁹ that is based on a random sample of households, and which has longitudinal elements built in (repeated cross-sectional with nested serial sampling of a subset of participants). Although limited to estimating community infection (and antibody) levels, the survey’s de-identified data have provided the most reliable estimates of infection prevalence in England (initially targeting 11,000 households from 20,000 invited to participate in April 2020) and then in the UK (from October 2020, with the aim of collecting 150,000 swab test results from individuals at least every fortnight). Might a post-marketing pharmacovigilance signal detection system benefit from a similar approach with the aim of collecting more detailed data from those receiving a drug or vaccine that has received post marketing authorisation? This could be linked to specific PV focussed centres—establishing a network of Electronic Health Record (EHR) enabled sites that closely monitor the safety of patients that receive a newly authorised drug. Patients could be selected at random, with consent sought to participate, within such a framework. In both cases, the numerator, denominator and quality of information are more clearly defined. Indeed, the FDA’s Sentinel⁴⁶ safety surveillance system primarily accesses electronic healthcare and administrative claims data and is based on a distributed data network and a common data model while OHDSI⁵⁰ (pronounced Odyssey) is an established international network of researchers and observational health databases (centrally coordinated at Columbia University) that undertakes large-scale analytics of health data—including medical product safety surveillance.

Other experimental methods could also be considered in specific circumstances. Farrington and Whittaker⁵¹ describe the case series model that uses data from only cases to estimate the relative incidence of events. The model applies self-matching and therefore matches for “all age-independent confounders that act multiplicatively on the baseline incidence.” The method is directed at the potential association between exposure and events, and utilises the timing of both to evaluate causality—in effect, adopting the deep approach referred to above that requires the temporal detail to be captured. Farrington and Whittaker generalise the case series model, developing a semiparametric approach that allows covariate effects to be modelled.

In terms of the SARS-CoV-2 vaccines, embedding an experimental framework into the initial roll-out, could have provided valuable data. For instance, a stepped-wedge cluster design (as described in section 5.2) could have been used

for care homes and could have provided an opportunity to estimate both efficacy and safety in the elderly. It could also have been used for 2nd dose or booster dose comparisons in a similar manner—particularly when logistics and supply challenges were present. It could also have been used for children, with clusters being schools.

Recommendation 6. *The current broad and deep approach to pharmacovigilance signal detection should be augmented, for a defined period after product authorisation by a more systematic approach to signal detection that is based on representative sampling and, where possible, an experimental or quasi-experimental framework.*

7 | DISCUSSION

This article has presented an eclectic mix of topics were the SARS-CoV-2 pandemic could become a catalyst for the cross-over of statistical thinking and practices from one area of research to another. Indeed, there are signs that this cross-over of thinking has begun.

In relation to NPIs, WHO⁵² has stated that “Member States can improve the transparency of the decision-making process by establishing criteria for evaluating NPIs. Predefined criteria on the efficacy and socioeconomic costs of NPIs will facilitate multisectoral deliberations of the measures and assist public health officials when the data and evidence on these dimensions are incomplete.” In this respect, the statement can be viewed as the need to assess both risk (including potential wider adverse impact) and benefit, noting that the SARS-CoV-2 vaccines were subject to a strict safety evaluation and risk/benefit assessment by the regulatory agencies prior to marketing authorisation.

In terms of embedding an experimental framework with clinical practice, post marketing authorisation, the Centres for Medicare and Medicaid Services (CMS)⁵³ in the US have proposed that they will cover FDA approved monoclonal antibodies for the treatment of patients with Alzheimer’s disease, where the approval is based on amyloid reduction, only in CMS approved RCTs or in National Institute of Health (NIH) supported trials. The trials must be aimed at demonstrating “statistically significant and clinically meaningful” impact on the decline in cognition and function. Such trials can be extended to include longitudinal follow-up, but the direction of travel is clear in a disease where drug approval has been controversial.

When randomisation is impractical, observational studies, whether in health research or the social sciences, should be clearer at the outset as to what magnitude of effect would constitute public benefit, even in cases where the data that are generated are for purposes other than research, for example, as part of service delivery. While in drug development, there is also potential to make greater use of experimental and sampling frameworks following marketing authorisation of a treatment or vaccine. The experimental framework includes stepped-wedge cluster randomisations designs, master protocols where simple clinical trials are embedded into clinical practice and quasi-experimental techniques such a regression discontinuity designs. Representative sampling could be undertaken to provide a richer and higher quality data to evaluate safety.

Overall, the SARS-CoV-2 pandemic has seen research conducted at great speed with undoubted success in the areas of COVID-19 prevention and treatment. The daily reporting of test results, hospitalisations and deaths has also been an exemplar of real-time big data curation. However, what is striking is how the tried and tested methods of randomisation and random sampling, with a few modern twists along the way, have delivered the most reliable data on which to base public health policy. An embedded experimental framework, even if quasi-experimental, should become a more routine feature for NPIs in healthcare but also for of the post-marketing evaluation of vaccine and treatments.

ACKNOWLEDGMENTS

We thank the reviewers and Associate Editor for their helpful comments.

CONFLICT OF INTEREST

The author is employed by ICON which is a contract research organisation. ICON provides pharmaceutical services to the pharmaceutical and biotechnology industries. ICON conducts clinical trials on behalf of Sponsors, including COVID-19 vaccine trials. The author is also a member of the UK Statistical Authority’s Research Accreditation Panel. The views expressed in the article are attributable to the author alone.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed.

ORCID

Andrew D. Garrett  <https://orcid.org/0000-0002-0932-8024>

REFERENCES

1. Working Party on Statistical Issues in First-in-Man Studies. Statistical issues in first-in-man studies. *J R Statist Soc A*. 2007;170:517-579.
2. Development of licensure of vaccines to prevent COVID-19: Guidance for industry. Food and Drug Administration. Center for Biological Evaluation and Research, June. 2020.
3. ICH topic E9. Statistical principles for clinical trials. 5 February. 1998.
4. Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med*. 2020;383:2603-2615. doi:10.1056/NEJMoa2034577
5. Voysey M, Clemens SAC, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials on in Brazil, South Africa, and the UK. *Lancet*. 2021;397:99-111. doi:10.1016/S0140-6736(20)32661-1
6. Sadoff J, Gray G, Vandebosch A, Cárdenas V, Shukarev G, Grinsztejn B, Goepfert PA, Truyers C, Fennema H, Spiessens B, Offergeld K, Scheper G, Taylor KL, Robb ML, Treanor J, Barouch DH, Stoddard J, Ryser MF, Marovich MA, Neuzil KM, Corey L, Cauwenberghs N, Tanner T, Hardt K, Ruiz-Guiñazú J, le Gars M, Schuitemaker H, van Hoof J, Struyf F, Douoguih M, ENSEMBLE Study Group Safety and efficacy of single-dose Ad26.COV2.S vaccine against Covid-19. *N Engl J Med* 2021; 384:2187-2202. DOI:10.1056/NEJMoa2101544, 384.
7. Ferguson N, Laydon D, Nedjati-Gilani G, et al. Report 9: impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College London. 16 March, 2020. DOI: 10.25561/77482
8. Liu IT, Prasad V and Darrow JJ. Evidence for community cloth face masking to limit the spread of SARS-CoV-2: a critical review. Cato Working Paper. Cato Institute. No. 64; 2021.
9. Bundgaard H, Bundgaard JS, Raaschou-Pedersen DET, et al. Effectiveness of adding a mask recommendation to other public health measures in prevent SARS-CoV-2 infection in Danish mask wearers. A randomized controlled trial. *Ann Intern Med*. 2021;174(3): 335-343.
10. Abaluck J, Kwong LH, Styczynski A, et al. Impact of community masking on COVID-19: a cluster-randomized trial in Bangladesh. *Science*. 2021;375:eabi9069. doi:10.1126/science.abi9069
11. Public Registers. Accredited researchers and accredited projects. UK Statistics Authority. <https://uksa.statisticsauthority.gov.uk/digitaleconomyact-research-statistics/better-useofdata-for-research-information-for-researchers/list-of-accredited-researchers-and-research-projects-under-the-research-strand-of-the-digital-economy-act/>
12. Administrative Data Research UK. ESCR. <https://www.adruk.org>
13. List of projects. ADR UK. <https://www.adruk.org/our-work/browse-all-projects/>
14. Hand DJ. Statistical challenges of administrative and transaction data. *J R Statist Soc A*. 2018;181:555-605.
15. Baldwin JR, Pingault J-P, Schoeler T, et al. Protecting against researcher bias in secondary data analysis: challenges and potential solutions. *Eur J Epidemiol*. 2022;37:1-10. doi:10.1007/s10654-021-00839-0
16. Bacon S, Goldacre B. Barriers to working with National Health Service England's open data. *J Med Internet Res*. 2020;22:1-14. doi:10.1126/science.abi9069
17. Williamson EJ, Walker AJ, Bhaskaran K, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature*. 2020;584:430-436. doi:10.1038/s41586-020-2521-4
18. Public Health England. Technical summary: Public Health England data series on deaths in people with COVID-19 death. 12 August 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/916035/RA_Technical_Summary_-_PHE_Data_Series_COVID_19_Deaths_20200812.pdf
19. The Centre for Evidence-Based Medicine. Oxford University. Reconciling COVID-19 death data in the UK. 8 April 2020. <https://www.cebm.net/covid-19/reconciling-covid-19-death-data-in-the-uk/>
20. The Centre for Evidence-Based Medicine. Oxford. Are COVID-19 patients in hospital or admitted to hospital. 13 April 2020. <https://www.cebm.net/covid-19/are-covid-19-patients-in-hospital-or-admitted-to-hospital/>
21. Clinical Data Interchange Standards Consortium. <https://www.cdisc.org>
22. ICH E6(R2). Guideline for good clinical practice. 1 December 2016.
23. Guidance for Industry. Part 11, Electronic Records; Electronic Signatures—Scope and Application. FDA; 2003
24. Smith M, Marshall A. Importance of protocols for simulation studies in clinical drug development. *Stat Methods Med Res*. 2011;20(6): 613-622. doi:10.1177/0962280210378949
25. <https://www.ukri.org/our-work/tackling-the-impact-of-covid-19/vaccines-and-treatments/recovery-trial-identifies-covid-19-treatments/>
26. The RECOVERY Collaborative Group. Dexamethasone in hospitalised patients with Covid-19. *N Engl J Med*. 2021;384:693-704.
27. The RECOVERY Collaborative Group. Tocilizumab in patients admitted to hospital with Covid-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet*. 2021;397:1637-1645. doi:10.1016/S0140-6736(21)00676-0

28. The RECOVERY Collaborative Group. Casirivimab and imdevimab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. medRxiv. doi: [10.1101/2021.06.15.21258542](https://doi.org/10.1101/2021.06.15.21258542)
29. The RECOVERY Collaborative Group. Hydroxychloroquine in hospitalised patients with Covid-19. *N Engl J Med.* 2020;383:2030-2040. doi:[10.1056/NEJMoa2022926](https://doi.org/10.1056/NEJMoa2022926)
30. Saag MS. Misguided use of hydroxychloroquine for COVID-19. The infusion of politics into science. *JAMA.* 2020;324(21):2161-2162. doi: [10.1001/jama.2020.22389](https://doi.org/10.1001/jama.2020.22389)
31. Mehra MR, Ruschitzka F, Patel AN. Retraction-hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet.* 2020;395:P1820. doi:[10.1016/S0140-6736\(20\)31324-6](https://doi.org/10.1016/S0140-6736(20)31324-6)
32. *Impact Report: Analysis and insight into critical drug development issues.* Tufts centre for the study of drug development. Tufts University. 2018; 20(4):1-4.
33. Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. *N Engl J Med.* 2017;377:62-70.
34. Liu X, Shaw RH, Stuart ASV. Safety and immunogenicity of heterologous versus homologous prime-boos schedules with an adenoviral vectored and mRNA COVID-19 vaccine (com-COV) a single-blind, randomised non-inferiority trial. *Lancet.* 2021;398:856-869.
35. Study Protocol Com-COV2. <https://comcovstudy.org.uk/study-protocol>
36. Gelman A, Hill J, Vehtari A. *Regression and other stories.* Cambridge University Press; 2021.
37. Baraniuk C. Covid-19: how the UK vaccine rollout delivered success so far. *BMJ.* 2021;372:n421.d
38. Bermingham C, Morgan J, Ayoubkhani D, et al. Estimating the effectiveness of first dose pf COVID-19 vaccine against mortality in England: a quasi-experimental study. medRxiv; 2021. doi: [10.1101/2021.07.12.21260385](https://doi.org/10.1101/2021.07.12.21260385)
39. Hurley J, Walker D. Staff working paper no. 943. Did the Covid-19 lockdown reduce business activity Evidence from UK SMEs. Bank of England. November 2021. <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2021/did-the-covid-19-local-lockdowns-reduce-business-activity-evidence-from-uk-smes.pdf?la=en&hash=E8E116D61061BB71B018643E983FD306D8553C80>
40. Takaku R, Yokoyama I. What the COVID-19 school closure left in its wake: evidence from a regression discontinuity analysis in Japan. *J Public Econ.* 2021;195:1-10. doi:[10.1016/j.jpubeco.2020.104364](https://doi.org/10.1016/j.jpubeco.2020.104364)
41. Senn S. A brief note regarding group matching in medical research. February 2020. http://www.senns.uk/Stats_Notes/Matching.pdf.
42. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis and reporting. *BMJ.* 2015;350:h391.
43. WHO. Definitions. https://www.who.int/medicines/areas/quality_safety/safety_efficacy/trainingcourses/definitions.pdf
44. Pfizer protocol C4591001. https://cdn.pfizer.com/pfizercom/2020-11/C4591001_Clinical_Protocol_Nov2020.pdf
45. Medicines and Healthcare products Regulatory Agency. 07 April 2022. <https://www.gov.uk/government/news/uk-regulator-confirms-that-people-should-continue-to-receive-the-covid-19-vaccine-astrazeneca>
46. www.sentinelinitiative.org/
47. Steklenborg V, Ellenius J, Maskell S, et al. Recommendations for the use of social media in pharmacovigilance: lessons from IMI WEB-RADR. *Drug Saf.* 2019;42:1393-1407.
48. Meng X-L. Statistical paradises and paradoxes in big data (I): law of large populations, big data paradox, and the 2016 US presidential election. *Ann Appl Stat.* 2018;12(2):685-726.
49. COVID-19 Infection Survey: methods and further information. ONS. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/covid19infectionsurveyspilotmethodsandfurtherinformation>
50. Observational Health Data Sciences and Informatics. <https://ohdsi.org/>
51. Farrington CP, Whitaker HJ. Semiparametric analysis of case series data (with discussion). *J R Statist Soc C.* 2006;55(5):1-28.
52. WHO. Calibrating long-term non-pharmaceutical interventions for COVID-19. Principles and facilitation tools. 30 July 2021. <https://apps.who.int/iris/bitstream/handle/10665/332099/WPR-DSE-2020-018-eng.pdf?sequence=5&isAllowed=y>
53. The Centers for Medicare and Medicaid (CMS). Monoclonal antibodies directed against amyloid for the treatment of Alzheimer's disease. 2022. CAG-00460N. <https://www.cms.gov/medicare-coverage-database/view/ncacal-decision-memo.aspx?proposed=Y&NCAId=305>

How to cite this article: Garrett AD. The cross-over of statistical thinking and practices: A pandemic catalyst. *Pharmaceutical Statistics.* 2022;21(4):778-789. doi:[10.1002/pst.2221](https://doi.org/10.1002/pst.2221)