Pacifying A Pandemic
With Numbers
A key reveal from this past year of pandemic response is how timely, accurate and accessible data can inform and guide health system decision-making. Good data, effectively applied, can transform clinical outcomes. The UK’s recent RECOVERY (Randomized Evaluation of COVID-19 Therapies) trial is a case in point: it found that a common steroid drug, dexamethasone, although not effective in all patients, still reduced mortality by one-third for the sickest patients, providing an evidence-based alternative to costly hospitalizations.

Yet the converse is also true: 18 months into the pandemic, little data exists on the impact of non-pharmaceutical interventions against COVID-19, especially among the most vulnerable populations classified by geography, gender, age, race or disability. And despite the fact that some 3,000 COVID-related trials are now in process, many are seen as too small, too rushed and statistically underpowered to render clinically useful judgments against a virus whose main feature is its variability – indeed, estimating the overall course of the pandemic has “largely failed,” according to a recent assessment by prominent researchers in the International Journal of Forecasting.

One positive note is that policymakers now have a better understanding of obstacles in existing infrastructure for the evaluation of large volume of data during a global health emergency. These include the low level of public trust in information sources generally; researchers’ differing perceptions of data quality, consistency and interpretation; lack of agreed standards on data governance and systems interoperability; and poor coordination among stakeholders in the public and private sectors. Combined, the effect is to limit rapid prioritization of initiatives that carry the greatest benefit to disease mitigation and prevention.

Fixing this sloppy architecture is long overdue, particularly as the sheer volume of health information grows: the US National Institute of Medicine (IOM) estimates that the average clinical trial today generates more than 3 million...
data points, yet due to absent or inconsistent rules on open access much of this data is in private hands and not available to researchers. Big commercial drug companies account for a substantial portion of the flow, but precisely what they should do with it all remains a gray area — such data could be considered an asset, but it is a contingent one, because monetizing data is a challenge when others see it as a public good. In addition to underlying issues around IP, this helps explains why the industry has often been hesitant about initiatives to consolidate and share the bounty from “big data” more widely, to the benefit of society overall.

“Are we a drug company or an information company?” It is a strategic question that the big pharma C-suite has been grappling with since the start of the genomic revolution two decades ago. Today, that question is top of mind, because making information both useful and visible might go far in filling gaps in the evidence armamentarium against this pandemic, as well as future ones. In short, at no time in history has the demand for shared, actionable data been higher. But what to do?

POWERHOUSE OF DATA: FROM MANY, ONE

A new collaborative partnership – the International COVID-19 Data Alliance (ICODA) – has come forward with a straight answer: in the institutional data space, it is time to move from segmentation to synthesis. Open access collaboration that connects siloed data assets with diverse human expertise, in pursuit of clinically sound answers to prevent viral spread, is the most powerful mechanism to defeat COVID-19.

Formally launched in July 2020, ICODA has a sterling pedigree, convened by Health Data Research UK (HDR/UK), an independent charitable research enterprise supported by 10 local funders, including the UK Health Department, which together manages one of the world’s richest repositories of scalable data on population health. The two other conveners are the Minderoo Foundation, based in Australia and one of the Asia region’s largest philanthropies, and the Bill & Melinda Gates Foundation, whose $50bn in assets puts it at the center of evidence-based decision-making in science, development and public health worldwide.

At the start of the pandemic in early 2020, the two philanthropies, along with Mastercard Inc’s Impact Fund, founded the COVID-19 Therapeutics Accelerator (CTA). Its mission, to study potential interventions to treat COVID-19, soon evolved to focus on building a trusted research platform – called the “workbench” – of curated, multi-dimensional data sets. Together, such data could answer key questions on the characteristics of the virus and impact of different therapeutic approaches on patient outcomes, including the most effective drug treatments and when to administer these during the course of care.

The workbench is now up and running as a key part of ICODA. The promise the workbench has shown in connecting a huge volume of disparate data sources is where ICODA meets the moment – to unite, improve and use data to inform clinical decision-making, particularly in areas where such knowledge is lacking.

“A hard lesson from past viral outbreaks is that big data in the raw is never prescriptive – by itself, it’s a mountain of nothing,” ICODA’s scientific and strategy advisory committee co-chair, Martin Murphy, told In Vivo. “What’s really needed is smart data, broken down and refined through systematic learning analysis tools that allow collaborators to test different research hypotheses, the results of which can be applied in real time, directly to practitioners in the field.” Murphy says ICODA is unique in that it is not just a “digital library,” like other data-gathering exercises launched in recent years, but a “digital laboratory” as well, in which the value-added from analytics tools like the ICODA workbench is made freely available to qualified researchers, from rich and poor nations alike, for the singular aim of defeating the virus.

In pursuit of this goal, ICODA intends to build an institutional presence for the long haul. “What’s distinctive about this group is our data infrastructure, compiled in a user-friendly format where researchers who may not otherwise have access to those big “data lakes” can explore, evaluate and generate better hypotheses on what’s going on in different sub-populations due to the pandemic,” Peter Mesenbrink, executive director of biostatistics at Novartis, told In Vivo.
Mesenbrink, a member of ICODA’s Statistical Expert Group that provides methodological advice to researchers, was influential in designing the ICODA workbench for its current assignment to evaluate data from randomized controlled trials (RCTs) for COVID-19 treatments. ICODA allows researchers to bring their own data sets to a private workspace on the platform, to be evaluated alongside data provided by ICODA, in concert with additional regional, national and private-sector data bases. This is usually preceded by expert review of the research proposals to certify these are aligned with ICODA’s mandate to contribute practical solutions that drive policy and clinical practice.

The 26-year veteran of Novartis notes that RCT inputs to the workbench consist of summary, not patient-level, data so it is effectively “de-risked,” with regard to patient privacy concerns. It also has integrative processing features that enhance the capability to gather data across many RCTs involving multiple pharma company or other sponsors. “This allows us to conduct the kind of deeply networked meta-analyses that are highly useful in addressing issues relating to population sub-groups or hypotheses that could not be answered relying on a standalone study from a single sponsor,” Mesenbrink said. He notes that his company has signaled its support for the workbench by releasing one RCT study to date but faces a challenge in that many others conducted by Novartis have multiple co-sponsors. This makes obtaining consent from them on release of the data, even at a summary level, time consuming and arguably more difficult.

Overall, the methodological calculus driving the workbench means that research conclusions can be rendered with more focus on securing the right policies on the ground. In a pandemic, this means one thing: saving lives. “We’ve got a big collection of talent in place around a data framework that not only helps us now, in real time, but will be ready for the world to apply against the next public health emergency. That’s why I see ICODA as a precedent in the annals of big data,” said Mesenbrink.

**THE USUAL SUSPECTS: BUT THIS TIME ALL IN ONE PLACE**

ICODA also stands out for the experiential firepower of its leadership. Although it operates as a collective, three influential innovators provide the real gravitas behind the operation:

- **Andrew Morris**, professor of medicine at the University of Edinburgh, where his lab produced major scientific breakthroughs against diabetes, and more recently as the founding director of HDR/UK, which works in close partnership with the UK National Health Service (NHS) and is also convenor and chief talent resource for ICODA;

- **Steven Kern**, deputy director of quantitative sciences at the Seattle-based Bill & Melinda Gates Foundation with a doctorate in bioengineering from the University of Utah and early proponent of the FAIR Data Service Principles (Findable; Accessible; Interoperable; and Replicable) adopted for use by the global data community in 2016; and

- **Steve Burnell**, CEO of the Collaborative Against Cancer initiative (CAC) at the Australia-based Minderoo Foundation, who leads the group’s efforts in COVID-19 related projects.

Specifically, Morris acts as spokesperson for ICODA and has been instrumental in mobilizing UK Health Data Research’s requisition and curating capabilities to drive ICODA’s initial work stream. CTA has also provided funding as well as introducing ICODA to high-profile ventures like the Gates Foundation’s Grand Challenges initiative, which now includes a COVID-19 data science workstream managed by ICODA. According to Morris, “This is data at scale, globally, in useable formats, in a modern environment – all geared to making the most direct impact against the pandemic.”

The other key funders of CTA – and thus indirectly ICODA, as well – include the Chan Zuckerberg Initiative; Mastercard Inc.; the Wellcome Trust; and Microsoft’s Artificial Intelligence (AI) for Health program, which is providing ICODA with resources to apply AI to advance data analytics using the ICODA workbench and other evaluative tools. With initial funding from these foundations in launching the CTA, and with the human capital on loan through the HDR/UK, ICODA now has the budget to fulfill its ambitious plans, including activities that will extend beyond the course of this pandemic.
In addition to these principal backers, ICODA has assembled a diverse and globally representative set of partners to provide both strategic support and data assets and technology expertise. Some 20 organizations have joined since ICODA held its first formal assembly meeting in November 2020. The list includes the developer of the ICODA workbench, Aridhia Informatics PLC, which creates data platforms for the analysis of biomedical data, as well as other commercial and non-profit data repositories and analytics vendors; publicly-funded bodies like the African Academy of Sciences and the UK National Institute for Health and Care Excellence (NICE); and one global pharmaceutical company, Novartis AG.

“Rarely in the global research space do we find so many key participants aligned under a single alliance like we now have with ICODA,” said Ben Kramer, vice president of US medical affairs for the immunology, ophthalmology, respiratory and infectious disease portfolios at Genentech Inc. “That’s important in addressing a virus as complex as COVID-19, the effects of which are going to impact every aspect of global health status for the indefinite future.”

PRINCIPLES, MISSION AND GOALS
ICODA functions as a collaborator rather than controller – the aim is to assemble high-quality data, make it useful to policymakers and then make it available for researchers who need it. Nevertheless, the founders have created a finely tuned operation that adheres to strict governance guidelines in pursuit of its goals.

At its inaugural meeting in November 2020, ICODA endorsed a set of 10 operating principles, covering, among other things, commitments on global equity and inclusion; transparency and accountability, through a “best practices” approach to technology, privacy and public outreach; collaboration, through a sharing work culture; innovation, especially in methods and practices that yield verifiable progress against disease; and making data more robust, as an aid to reproducibility in research.

In addition, ICODA is notable in endorsing both the FAIR Data Services Principles as well as the more recent “Five Safes” focused on data integrity. The latter cover protections like vetting researchers and projects on trustworthiness to “serve the public good” and prevent the exploitation of confidential data for private gain. Perhaps the most important guarantee covers safe outputs from research to prevent inadvertent identification of individual patients.

In pursuit of its mission to bring forward better data, trusted by all, the ICODA workflow is organized to secure maximum accountability and oversight. The centerpiece is the aforementioned annual Alliance Forum, which in November 2020 brought together virtually the

ICODA’S STRATEGIC PARTNERS

African Academy of Sciences
Aridhia Inc.
BREATHE Health Data Research Hub
CAIAC (Canadian Artificial Intelligence Association)
Certara Inc.
Cytel Inc.
DNA/Stack
Generation Scotland
Genomics England
Global Alliance for Genomic Health
Global Partnership for Sustainable Development Data
Health Data Research Network Canada
Infectious Diseases Data Laboratory
ISARIC (International Severe Acute Respiratory and Emerging Infection Consortium)
MMS (Medical Management Services) Inc.
National Institute for Health and Care Excellence (NICE)
Novartis AB
PA Consulting PLC
Preva Group LLC
Research Data Alliance
SAIL (Secure Anonymized Information Linkage Databank
UK Secure e-Research Platform (UKSeRP)
Shivom PLC

Source: ICODA
ICODA leadership and some 400 members of the data research community from 37 countries. A key outcome was the launch of the Grand Challenges ICODA Data Initiative, funded by the Bill and Melinda Gates Foundation, to provide grants for COVID-19 research studies in four categories: epidemiology of the disease and transmission of the virus; vaccines and prophylaxis; infections and prevention; and therapeutics and clinical care management. Steve Kern described it to In Vivo as “a community-building effort to recognize the contributions from all scientists around the globe.”

Selections will be based on the tangible public health impact from the research; robustness and relevance of the data sets intended for review; involvement of patients and public representatives in study design and execution; and whether the proposed research can address particular challenges of COVID-19 in low- and middle-income countries. Grants will cover six to 12 months of active work. All told, 418 project applications have been received and that list has been winnowed down to 25 grants of up to $100,000, with the winners being announced in July 2021.

Another outcome from the Alliance Forum was the hosting of a symposium in March this year examining barriers to the dissemination of research on COVID-19. Some 600 researchers from around the world attended Equitably Harnessing the Power of Health Data. One key data point was a participant survey that identified data governance issues as the single biggest hurdle to sharing of data, followed by the absence of financial incentives and poor data and IT infrastructure. There is also the traditional academic concern that work has to undergo a time-consuming hurdle of peer review publication before its underlying data can be shared. To advance the process, the symposium recommended that more research grants should be made conditional on a commitment to prompt and fair data sharing by recipients.

INTEGRATIVE APPROACH TO DECISIONS AND LEADERSHIP
Day-to-day management of ICODA is the responsibility of a seven-member executive management team that provides strategy direction and is accountable to a board composed of representatives from key data partners and funders. There are also five separate committees that provide additional guidance to management and staff.

The scientific and strategic advisory committee, co-chaired by Martin Murphy, founding CEO of the US CEO Roundtable on Cancer and its lead initiative, Project Data Sphere, on cancer clinical trial data sharing; and Agnes Binagwaho, a Rwandan pediatrician who served as her country’s Minister of Health and is now vice-chancellor at Rwanda’s University of Global Health Equity and adjunct professor at Harvard Medical School. The committee’s focus is seeking impactful research opportunities to fight COVID-19, ensuring the scientific rigor of the entire enterprise and addressing different ways to remove barriers to data sharing.

The ethics advisory committee, chaired by professor Effy Vayena, who heads the bioethics department at the Swiss Institute of Technology (ETHZ), teaches at the University of Zurich and has served in numerous multilateral institutions, including the OECD and the WHO as well as the Wellcome Trust and Commonwealth Fund. It monitors adherence to ICODA’s data governance regime, to ensure ICODA is meeting commitments to transparency, privacy and collaboration in the use of data for COVID-19 research.

The statistics experts group is chaired by Jonas Haggstrom, vice-president at Cytel Inc., the largest industry provider of statistical software and analytics used in drug clinical trials. The group oversees the application of the ICODA workbench, other analytical tools and methods and quality standards to ensure research results are fit for purpose in delivering useful evidence to policymakers worldwide. In addition to several academic members, the group includes representatives from Novartis AG, Pfizer Inc., AstraZeneca PLC, a number of academic institutions and a former top official of the US FDA.

The patient and public expert group, launched at an October 2020 meeting convened by Bella Starling, co-director of Public Programs for the Wellcome Trust, promotes patient and public participation in ICODA’s work. It manages outreach to patients, providers, health advocacy groups and other parts of government and civil society to ensure ICODA’s work translates to lay audiences.
and gets recognition at the grassroots of clinical care. Its work is just getting underway, but it is expected to serve as the external voice of ICODA with key external constituencies, especially in low- and middle-income countries.

Finally, the core funders committee exists to represent the key donors to ICODA and to guarantee it has the resources to fulfill its mission, particularly as it relates to research needs in low- and middle-income countries. It also looks at financing needs for ICODA in the context of addressing future public health emergencies – a key goal of the organization, one contingent on it making a difference in the battle against COVID-19.

**DATA AS AN ACTIVE VERB: PURSUING PROJECTS TO DRIVE POLICY**

With the organizational structure in place, ICODA’s attention is now on implementing a work program designed to make data “meet the moment.” Attention is focused on what ICODA calls “Driver Projects” which tackle big outstanding research questions, identify new models of managing data from RCTs and other sources, and get practical evidence that informs COVID-19 policy and practice out to the field, all in six months or less. “With all this data expertise now focused on a common goal against COVID, we realized it was the right time to plan something new,” said Névine Zariffa, former senior vice president of biometrics and information sciences for drug maker AstraZeneca and currently a member of the ICODA executive leadership team. She is leading work on the first Driver Project.

“Over the years, most of us have been involved in various approaches to data-sharing, none of which have delivered what the world needs during a pandemic. We now know what the pitfalls are and what we have to do: forget the data silos and focus on something that is accessible and global and that recognizes the importance of building viable data repositories in low- and middle-income countries, not just the affluent OECD bloc. Most important, we must strive for immediacy and impact. We can’t be satisfied simply around the number of research papers published, but rather what we can do with this information to fight disease and help health care decision-makers.”

Hence the first Driver Project charts a different path. It is looking holistically at the efficacy and safety of the full range of treatments for COVID-19 using RCT data sets contributed by ICODA’s industry and philanthropic partners, including the COVID-19 R&D Alliance, which encompasses all the major pharmaceutical players committed to open data-sharing worldwide. The aim is to provide good evidence of what works best from the standpoints of policy and clinical practice.

The Statistical Expert Group and members of the COVID-19 R&D Alliance are contributing to the preparation of a “data dictionary” intended to harmonize trial endpoints, definitions, classifications and identifiers so that side-by-side reviews can be made. The same was done in trying to structure the etiology of the virus, such as disease severity and characteristics of biomarkers and assays like d-dimer and c-reactive protein. The intent is that in creating a common basis for compiling, curating and digitizing RCT data, researchers will be able to deploy summary, population-level insights that are actionable by policymakers. Certara Inc., a leading trial software vendor and ICODA data partner, has also provided summaries, in an analyzable format, of approximately 500 COVID-related studies that have so far been reported in the public domain.

“What we’ve achieved so far is a data compilation with statistical tools that lowers the threshold for access and interpretation in a way that will expand the evidence base around COVID-19 and allow governments and health providers to get ahead of the pandemic,” said Zariffa. Many more studies are planned to be submitted but the majority received are still not in the fully harmonized data format that makes them most useful to researchers. “This process is proving to be slower than anticipated,” Zariffa told *In Vivo*. 
Nevertheless, the plan is to grow that number significantly in the months ahead, now focusing on the drug classes where researchers still have questions about the efficacy of the drug in question. When the work is completed, the Driver Project will follow up by seeking educational and communications opportunities to spread the message among researchers on the best interventions available.

The second Driver Project is focused on how the pandemic has impacted vulnerable populations. Specifically, the International Perinatal Outcomes in the Pandemic (iPOP) study is investigating the impact of pandemic mitigation measures – such as population lockdowns – on the rate of pre-term births worldwide. Premature birth is a leading cause of death, for infants and mothers alike, in many countries but especially for those in low- and middle-income countries, where access to primary care has been sharply curtailed by the pandemic-induced strain on health services. Andrew Morris told In Vivo, “I especially like this Project as it includes participation of over 80 countries, right from the start; the group has thus been productive in a very short time. The collaborative and team science approach here resonates with the distinctive way that ICODA works. I am hopeful the word will spread to others.”

With support from a UK data bank called SAIL (Secure Anonymized Information Linkage), a principal source for UK population, health and social care data intended solely for research, ICODA has been able to demonstrate federated access and analysis of more than 30 international iPOP COVID-19 data sets hosted by SAIL. Ultimately, more than 100 researchers on the iPOP team from 80 countries are seeking to deposit their data into SAIL to evaluate how maternal workloads, hygiene practices, air pollution and other externalities are impacting pre-term birth rates across countries. That data can then be requested by other researchers from within ICODA. Since the requests go to the iPOP team, Kern says ICODA is actually facilitating large and productive collaborations that otherwise would not exist. Armed with the data, iPOP will make recommendations to shape perinatal care practices worldwide, not just for the duration of this pandemic but beyond. Adds Zariffa, “The project is representative of ICODA’s approach: tackle a big, relevant health question; seed it with good accessible and trustworthy data from diverse sources; and get to the core of an issue very quickly; with results that the world community can build on, not just in a health emergency but for the long-term state of health itself.” More Driver Projects are in the planning stage. Most of these will be linked to the grants being awarded through the Grand Challenges COVID-19 Data Initiative program.

FINDING MEANING IN NUMBERS
ICODA faces some challenges as it moves forward to execute on its promises. One is the effort to keep growing the large repositories of data required to deliver fresh policy and practice insights as the pandemic continues to evolve – in a dizzying number of directions. There are complex statistical methodological issues that require finessing, such as retrospective pooling of data sets, which reflects a basic reality that trials are rarely, if ever, designed in unison, around a single template. “In a global health emergency, it’s important to focus on the end result of saving lives and not get lost in insisting on 100% comparability with a study sponsor’s interpretation of how data is defined,” says Novartis’ Mesenbrink. “What we need is to ensure the data itself is accurately compiled and directed toward answering a solid underlying hypothesis.”

Mesenbrink also references IP concerns that have surfaced in the industry’s RCT data sharing discussions, such as how researcher explorations of ICODA’s summary-level data sets could compromise the guardrails around patient-level data available on the same trials through other RCT data sharing initiatives, such as Transcelerate in the US. More disclosures of the treatment effect of particular therapies could result, jeopardizing a patent claim.

On a positive note, scientific advisory committee co-chair Martin Murphy adds that AI-enabled technology will help resolve these problems by enhancing data integration capabilities. “One of the unexpected benefits from the pandemic is the progress in using AI to make data fit-for-purpose in addressing a vastly more sophisticated research agenda, in all areas of science, not just health. The era of interdisciplinary insight is upon us, and it’s up to us to take advantage of it, especially in areas of the world that have not benefited from these new technologies.”
Nevertheless, negotiating all those data sharing agreements can be time consuming. “We've been asking all the biopharmaceutical majors to provide us with their summary trial and treatment data on COVID-19, in a prescribed format that allows researchers a richer and more augmented experience in reviewing it,” Zariffa said.

So far, ICODA has signed 10 such agreements with multinational pharma companies and another with a global COVID-19 researcher consortium focusing on the evaluation of several drug classes in a study called REMAP-CAP (Randomized Embedded Multifactorial Adaptive Platform for Community-Acquired Pneumonia). ICODA’s biggest cooperation deal to date is now being negotiated with Oxford University's Infectious Diseases Data Observatory (IDDO), which has extensive data links to the public health research and epidemiological communities of low- and middle-income countries. More such multilateral agreements are expected to come forward in the next few months.

ICODA would also like to work with some other key data sources, including the US National Institutes of Health (NIH). NIH represents a challenge due to the organization’s specialized public mandate and approach to managing the trial data it obtains for its own data registries. The fact is that a successful and trustworthy data ecosystem needs to combine data from numerous providers to create value as a decision-making tool. Building a collaborative network that motivates parties with similar – but sometimes competing – interests will require a sustained effort between ICODA and private philanthropies, government and industry.

GETTING BIG PHARMA’S BUY-IN

To address these issues head on, Zariffa and others at ICODA with backgrounds in industry are working to persuade biopharma companies to take a more active role in the group’s work. “The biopharma industry has made an invaluable investment in addressing COVID-19. What I'd like to see is them going that final mile where every RCT data set is sent to ICODA. And I'd like the industry CEOs to address those internal silos that can slow things down – starting with sharing that summary-level data with us right away. Data scientists, researchers, health providers and policy-makers need it now. In fact, your own teams need other group’s data to provide context to your findings and help design the next sets of trials.”

Zariffa also emphasizes that ICODA is very explicit about its partnering proposition. “Trying to solve every research question with data is a recipe for failure, so we intend to be selective in focusing on data that hits the sweet spot in fighting the pandemic and its aftermath.”

Overall, ICODA represents an idea planted precisely at the right time to make a difference. If anything, COVID-19 has shown that, despite an unprecedented – and costly – acceleration in the pace of vaccine development, mobilization of data to make sense of the contagion itself lags far behind. As the World Bank notes in its 2021 World Development Report, Data for Better Lives, data is an essential, relatively cheap but still missing part of an improved global infrastructure for health – one that can spell the difference between life and death for millions. ICODA represents a bit of “radical transparency” in pursuit of that goal, so why not give it a try? Especially when one considers that the harsh year of COVID-19 just passed could just be one of many to come.

Published online 30 June 2021

SIGNED DATA-SHARING AGREEMENTS

Amgen Inc.
AstraZeneca PLC
Gilead Sciences Inc.
GlaxoSmithKline PLC
Novartis AB
Pfizer Inc.
Roche/Genentech Inc.
Sanofi SA
Takeda Pharmaceuticals Co.
UCB SA/NV
REMAP-CAP Study Team (Randomized Embedded Multifactorial Adaptive Platform for Community-Acquired Pneumonia)

Source: ICODA

Published online 30 June 2021