

ILLUSTRATION BY JASIEK KRZYSZTOFIAK/NATURE

HOW COVID BROKE THE EVIDENCE PIPELINE

The pandemic stress-tested the way the world produces evidence – and revealed all the flaws. **By Helen Pearson**

It wasn't long into the pandemic before Simon Carley realized we had an evidence problem. It was early 2020, and COVID-19 infections were starting to lap at the shores of the United Kingdom, where Carley is an emergency-medicine doctor at hospitals in Manchester. Carley is also a specialist in evidence-based medicine – the transformative idea that physicians should decide how to treat people by referring to rigorous evidence, such as clinical trials.

As cases of COVID-19 climbed in February, Carley thought that clinicians were suddenly abandoning evidence and reaching for drugs just because they sounded biologically plausible. Early studies Carley saw being published often lacked control groups or enrolled too few people to draw firm conclusions. “We were starting to treat patients with these drugs initially just on what seemed like a good idea,” he says. He understood the desire to do whatever is possible for someone gravely ill, but he also knew how dangerous it is to assume a drug works when so many promising treatments prove to be ineffective – or even harmful – in trials. “The COVID-19 pandemic has arguably been one of the greatest challenges to evidence-based medicine since the term was coined in the last century,” Carley and his colleagues wrote of the problems they were seeing¹.

Other medical experts echo these concerns. With the pandemic now deep into its second year, it's clear the crisis has exposed major weaknesses in the production and use of research-based evidence – failures that have inevitably cost lives. Researchers have registered more than 2,900 clinical trials related to COVID-19, but the majority are too small or poorly designed to be of much use (see ‘Small samples’). Organizations worldwide have scrambled to synthesize the available evidence on drugs, masks and other key issues, but can't keep up with the outpouring of new research, and often repeat others' work. There's been “research waste at an unprecedented scale”, says Huseyin Naci, who studies health policy at the London School of Economics.

At the same time, shining examples of good practice have emerged: medical advances based on rigorous evidence have helped to chart a route out of the pandemic. The rapid trials of vaccines were spectacular successes, and well-run trials of possible treatments have shown, for instance, that some steroids help to fight COVID-19, but the drug hydroxychloroquine doesn't. Many physicians point to the United Kingdom's RECOVERY trial as exemplary in showing how quick action and simple protocols make it possible to conduct a large clinical trial in a crisis. And researchers have launched ‘living’ systematic reviews that are constantly updated as research emerges – essential in a fast-moving disease outbreak.

As the COVID-19 response turns from

a sprint to a marathon, researchers are taking stock and looking ahead. In October, global-health leaders will meet for three days to discuss what's been learnt from COVID-19 about supplying evidence in health emergencies. COVID-19 is a stress test that revealed the flaws in systems that produce evidence, says Elie Akl, an internal-medicine specialist and clinical epidemiologist at the American University of Beirut. “It would be shameful if we come out of this experience and not make the necessary change for the next crisis.”

The evidence revolution

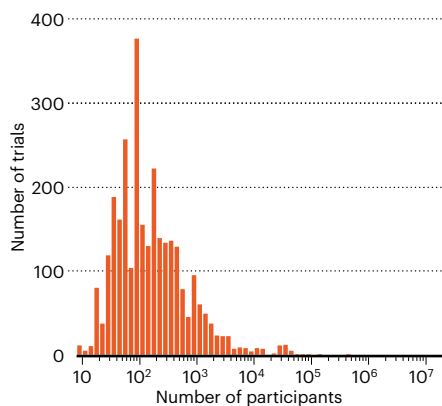
The idea that medicine should be based on research and evidence is a surprisingly recent development. Many doctors practising today weren't taught too much about clinical trials in medical school. It was standard to offer advice largely on the basis of opinion and experience, which, in practice, often meant following the advice of the most senior physician in the room. (Today, this is sometimes called evidence-based medicine.)

In 1969, a young physician called Iain Chalmers realized the lethal flaw in this approach when he worked in a Palestinian refugee camp in the Gaza Strip. Chalmers had been taught in medical school that young children with measles should not be treated with antibiotics unless it was certain that they had a secondary bacterial infection. He obediently withheld the drugs. But he found out later that what he'd been taught was wrong: six controlled clinical trials had shown that antibiotics given early to children with measles were effective at preventing serious bacterial infections. He knows that some children in his care died as a result, a tragedy that helped set him on a mission to put things right.

In the 1970s, Chalmers and a team set about systematically scouring the medical literature for controlled clinical trials relating to care in pregnancy and childbirth, a field in which the use of evidence was shockingly poor. A decade or so later, they published what they'd found in

SMALL SAMPLES

In one database of COVID-19 trials, 40% stated that they were enrolling fewer than 100 patients – a sample size that is generally too small to be useful.



a database and two thick books with hundreds of systematic reviews showing that many routine procedures – such as shaving the pubic hair of women in labour or restricting access to their newborn babies – were either useless or harmful. Other procedures, such as giving antenatal steroids for premature births, convincingly saved lives. It was a landmark study², and in 1993, Chalmers was central in founding the Cochrane collaboration, which aimed to follow this model and synthesize evidence across other medical specialties.

On the other side of the world, meanwhile, a group of doctors led by David Sackett working at McMaster University in Hamilton, Canada, had been developing a new way of teaching medicine, in which students were trained to critically appraise the medical literature to inform their decisions. In 1991, the term evidence-based medicine was coined, and it was later defined³ as the “conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients”.

Today, it's common for doctors to use evidence, alongside their clinical expertise and a patient's preferences, to work out what to do. A bedrock of evidence is built from systematic reviews, in which researchers follow standard methods to analyse all relevant, rigorous evidence to answer a question. These reviews often include meta-analyses – the statistical combining of data from multiple studies, such as clinical trials. Cochrane and other groups published more than 24,000 systematic reviews in 2019 alone.

Organizations in areas ranging from education to conservation also create evidence syntheses, and policymakers find them an invaluable tool. When faced with a slew of conflicting studies, an evidence synthesis “has the power to identify important conclusions about what works that would never be possible from assessing the underlying trials in isolation”, says Karla Soares-Weiser, editor-in-chief of the Cochrane Library and acting chief executive of Cochrane, based in Tel Aviv, Israel.

The rise of evidence syntheses has been “an invisible and gentle revolution”, says Jeremy Grimshaw, a senior scientist and implementation researcher at the Ottawa Hospital Research Institute. Once you see the compelling logic of assessing an entire landscape of science in this way, “it's very hard to do anything else”, he says.

At least, that is, until COVID-19 hit.

Tumult of trials

Carley compares the time before and after COVID-19 to a choice of meals. Before the pandemic, physicians wanted their evidence like a gourmet plate from a Michelin-starred restaurant: of exceptional quality, beautifully presented and with the provenance of all the ingredients – the clinical trials – perfectly clear. But after COVID-19 hit, standards

slipped. It was, he says, as if doctors were staggering home from a club after ten pints of lager and would swallow any old evidence from the dodgy burger van on the street. “They didn’t know where it came from or what the ingredients were, they weren’t entirely sure whether it was meat or vegetarian, they would just eat anything,” he says. “And it just felt like you’ve gone from one to the other overnight.”

Kari Tikkinen, a urologist at the University of Helsinki who had run clinical trials in the past, was equally shocked early last year to talk to physicians who were so confident that untested therapies such as hydroxychloroquine were effective that they questioned the need to test them in clinical trials. It was “hype-based medicine”, he says – fuelled by former US president Donald Trump, who announced last May that he had started taking the drug himself. “It very quickly got ahead of us, where people were prescribing any variety of crazy choices for COVID,” says Reed Siemieniuk, a doctor and methodologist at McMaster University.

Many doctors and researchers did race to launch clinical trials – but most were too small to produce statistically meaningful results, says Tikkinen, who leads the Finnish arm of SOLIDARITY, an international clinical trial of COVID-19 treatments coordinated by the World Health Organization (WHO). Hydroxychloroquine was the most-tested drug according to a database of 2,900 COVID-19 clinical trials called COVID-NMA: it was tested in 250 studies involving nearly 89,000 people (see ‘Too many trials?’). Many are still under way, despite convincing evidence that the drug doesn’t help: the RECOVERY trial concluded that hydroxychloroquine should not be recommended to treat COVID-19 in June last year.

Researchers have known for well over a decade that colossal amounts of medical research are wasted because of poorly designed trials and a failure to assess what research has been done before⁴. A basic calculation at the start of a COVID-19 trial, Tikkinen says, would have shown the large number of participants necessary to produce a meaningful result. “There was no coordination,” he says.

Instead, hospitals should have joined up, as was done in a handful of mega-trials. SOLIDARITY has enrolled nearly 12,000 people with COVID-19 in more than 30 countries. And many researchers look with awe at the RECOVERY trial, which the United Kingdom launched rapidly in March 2020, in part because it was kept simple – a short consent procedure and one outcome measure: death within 28 days of being randomly assigned to a treatment or control group. The trial has now enrolled nearly 40,000 people at 180 sites and its results showing that the steroid dexamethasone reduced death rates changed standard practice almost overnight.

One clear take-home lesson, researchers say, is that countries need more large-scale

national and international clinical-trial protocols sitting on the shelf, ready to fire up quickly when a pandemic strikes. “We will learn a lot of lessons from this, and I think RECOVERY has set the standard,” Tikkinen says.

Carley says that in February, he treated a man with COVID-19 who desperately wanted to receive monoclonal antibodies, but the only route to do so at Carley’s hospital was by enrolling him in RECOVERY. The randomization protocol assigned him to receive standard care, rather than the therapy. “Which was tough – I still think it’s the right thing to do,” says Carley, who adds that the man did OK. The RECOVERY trial announced⁵ in February that the monoclonal antibody tocilizumab cut the

“You have this huge amount of inappropriate and probably wasteful duplication of effort.”

risk of death in people hospitalized with severe COVID-19; testing of another antibody cocktail is still under way.

The rise of reviews

The pandemic is “evidence on steroids”, says Gabriel Rada, who directs the evidence-based health-care programme at the Pontifical Catholic University of Chile in Santiago. Research on the disease has been produced at a phenomenal rate. And that created a knock-on problem for researchers who try to make sense of it.

The number of evidence syntheses concerning COVID-19 went through the roof, as governments, local authorities and professional bodies flocked to commission them. “We’ve never seen this level of demand from

decision makers saying ‘help, tell us what’s going on,’” Grimshaw says. Rada runs a giant database of systematic reviews in health called Epistemonikos (a Greek term meaning ‘what is worth knowing’). It now contains nearly 9,000 systematic reviews and other evidence syntheses related to COVID-19. But ironically, just like the primary research they are synthesizing, many of the syntheses themselves are of poor quality or repetitive. Earlier this year, Rada found 30 systematic reviews for convalescent plasma, based on only 11 clinical trials, and none of the reviews had included all the trials. He counted more than 100 on hydroxychloroquine, all out of date. “You have this huge amount of inappropriate and probably wasteful duplication of effort,” Grimshaw says. “There’s a fundamental noise-to-signal problem.”

One possible solution lies in PROSPERO, a database started in 2011 in which researchers can register their planned systematic reviews. Lesley Stewart, who oversees it at the Centre for Reviews and Dissemination at the University of York, UK, says that more than 4,000 reviews on COVID-19 topics have been registered so far, and the PROSPERO team appeals to researchers to check the database before embarking on a review, to see whether similar work already exists. She’d like to see better ways to identify the most important questions in health policy and treatment and make sure that researchers generating and synthesizing evidence are tackling those.

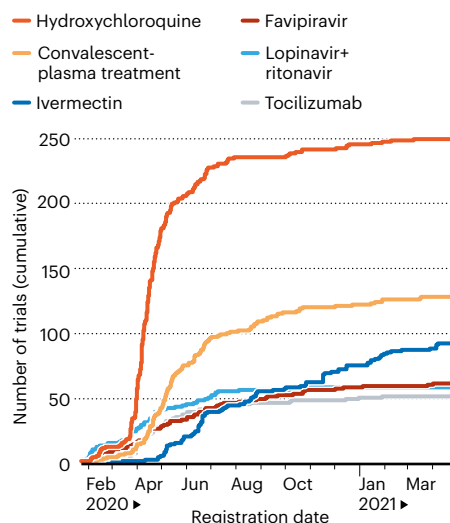
Researchers already knew that evidence syntheses took too long to produce and fell quickly out of date, and the pandemic threw those problems into sharp relief. Cochrane’s median time to produce a review is more than two years and, although it commits to updating them, that isn’t nimble enough when new research is flooding out. So, during the pandemic, Cochrane cut the time of some reviews to three to six months.

Systematic reviews are slow to produce in part because academics have to work hard even to identify relevant clinical trials in publication databases: the studies are not clearly tagged and researchers who do trials rarely talk to those collecting them into reviews. Julian Elliott, who directs Australia’s COVID-19 Clinical Evidence Taskforce, based at Cochrane Australia, Monash University in Melbourne, says it’s as if one group creates a precious artefact – its clinical-trial paper – and then tosses it into the desert, leaving the reviewers to come along like archaeologists with picks and brushes to try to unearth it in the dust. “It sounds completely insane, doesn’t it?” he says.

Rada is trying to help. During the pandemic, he has compiled one of the largest repositories of COVID-19 research in the world, containing more than 410,000 articles by early May. The team uses automated and manual methods to trawl literature databases for research and then classify and tag it, for example as a

TOO MANY TRIALS?

Studies assessing drugs against COVID-19 included 250 trials of hydroxychloroquine – a duplication that researchers say represents wasted effort.





Some US protesters against COVID-19 restrictions promoted drugs not backed by evidence.

randomized controlled trial. The goal is for the database, called COVID-19 Living Overview of Evidence (L-OVE), to be the raw material for evidence syntheses, saving everyone a monumental amount of work.

Drawing on this and other sources, a handful of groups including Cochrane have been developing living systematic reviews. Siemieniuk had produced such reviews before and helped to convene a group to build one on COVID-19 therapies. The international team, now about 50–60 people, combs the literature daily for clinical trials that could change practice and distills findings into a living guideline that doctors can quickly refer to at a patient's bedside and which is used by the WHO. "It's a very good concept," says Janita Chau, a specialist in evidence-based nursing at the Chinese University of Hong Kong and co-chair of a network of Cochrane centres in China. Chau says it's important to compile the evidence now rather than see interest in it fade away with the disease itself, as she saw during the SARS outbreak in 2003.

Isabelle Boutron, an epidemiologist at the University of Paris and director of Cochrane France, is co-leading another extensive living evidence synthesis, the COVID-NMA initiative, which is mapping where registered trials are taking place, evaluating their quality, synthesizing results and making the data openly available in real time. Ideally, she says, researchers planning trials would talk to evidence-synthesis specialists in advance to ensure that they are measuring the types of outcome that can be usefully combined with others in reviews. "We're really trying to link the different communities," she says.

Grimshaw, Elliott and others would like to see living reviews expanded. That's one focus of COVID-END (COVID-19 Evidence Network for support Decision-making), a network of

organizations including Cochrane and the WHO that came together in days in April 2020 to better coordinate COVID-19 evidence syntheses and direct people to the best available evidence. The group is now working out its longer-term strategy, which includes a priority list for living evidence syntheses.

As the world moves into a recovery phase, Grimshaw, who co-leads COVID-END, argues that it will be served best by a global library of a few hundred living systematic reviews that address issues ranging from vaccine roll-out to recovery from school closures. "I think there's a strong argument that you'll get more bang for the buck if, in selected areas, you invest in living reviews," he says.

Mosaic of evidence

Even when rigorous clinical trials are too slow or difficult to run, the pandemic served as a reminder that it's still possible to recommend what to do. In the United Kingdom, Trish Greenhalgh, a health researcher and doctor at the University of Oxford, expressed frustration at those who wanted bullet-proof evidence from randomized controlled trials before recommending the widespread use of face masks, even though there was a wealth of other evidence that masks could be effective and, unlike an experimental drug, that they posed little potential harm. (The United Kingdom mandated face masks on public transport in June 2020, long after some other countries.) "I think that was just a blast of common sense," says David Tovey, co-editor in chief of the *Journal of Clinical Epidemiology* and an adviser to COVID-END, based in London. "People have focused too much on randomized trials as being the one source of truth."

The issue is familiar in public health, says David Ogilvie, who works in the field at the

MRC Epidemiology Unit at the University of Cambridge, UK. In the standard paradigm of evidence-based medicine, researchers collect evidence on a therapy from randomized controlled trials until it gets a green or red light. But in many situations, such trials are unethical, impractical or unfeasible: it's impossible to do a randomized controlled trial to test whether a new urban motorway improves people's health by siphoning traffic out of town, for example. Often, researchers have to pragmatically assess a range of different evidence – surveys, natural experiments, observational studies and trials – and mosaic them together to give a picture of whether something is worthwhile. "You have to get on and do what we can with the best available evidence, then continue to evaluate what we're doing," says Ogilvie.

However well scientists synthesize and package evidence, there's of course no guarantee that it will be listened to or used. The pandemic has shown how hard it can be to change the minds of ideologically driven politicians and hardened vaccine sceptics or to beat back disinformation on Twitter. "We're definitely fighting against big forces," says Per Olav Vandvik, who heads the MAGIC Evidence Ecosystem Foundation in Oslo, which supports the use of trustworthy evidence.

Leaders in the field will pick up these debates in October during the virtual meeting organized by Cochrane, COVID-END and the WHO, to discuss what has been learnt about evidence supply and demand during the pandemic – and where to go next. One key issue, Soares-Weiser says, is ensuring that evidence addresses issues faced by low- and middle-income countries, as well as richer ones, and that access to evidence is equitable, too. "I really believe that we will come out of this crisis stronger," she says.

Carley, meanwhile, is still treating people with COVID-19 in Manchester, and sometimes still seeing new treatments recommended before they've been tested in trials. The last year has been exhausting and awful, he says, "seeing young, fit, healthy people coming in with quite horrific chest X-rays and not do terribly well".

At the same time, there's a thrill in seeing the enormous difference that evidence – that science – can make. "When results come out and you see that dexamethasone is going to save literally hundreds of thousands of lives worldwide," he says, "you think – 'that's amazing'".

Helen Pearson is *Nature's* chief magazine editor.

1. Carley, S. et al. *Emerg. Med. J.* **37**, 572–575 (2020).
2. Chalmers, I., Enkin, M. & Keirse, M. J. N. C. *Effective Care in Pregnancy and Childbirth* (Oxford Univ. Press, 1989).
3. Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B. & Richardson, W. S. *Br. Med. J.* **312**, 71 (1996).
4. Chalmers, I. & Glasziou, P. *Lancet* **374**, 86–89 (2009).
5. RECOVERY Collaborative Group. *Lancet* **397**, 1637–1645 (2021).