AN AUDIENCE WITH…

RECOVERY 1 year on: a rare success in the COVID-19 clinical trial landscape

Peter Horby, co-lead on the RECOVERY platform trial, discusses the origins, lessons learned and future plans for one of the most informative studies of COVID-19 therapeutics with **Asher Mullard**.

Credit: David Fisher, Fisher Studios, Oxford

On 23 March, the UK National Health Service's RECOVERY trial marked its first anniversary. Initially set up to test four possible treatment options in a few thousand hospitalized COVID-19 patients, the platform trial is still going strong. It has now enrolled nearly 40,000 patients — just under 10% of the UK's hospitalized COVID-19 patient population — into 13 treatment arms.

In a landscape marked by poorly designed and underpowered trials, RECOVERY stands out as a rare bright spot. Its results have demonstrated the benefits of the steroid dexamethasone and the IL-6-targeting antibody tocilizumab, while showing the lack of efficacy for azithromycin, colchicine, convalescent plasma, hydroxychloroquine and lopinavir–ritonavir. Ongoing arms are testing other agents, including an antibody cocktail and the JAK inhibitor baricitinib.

Peter Horby, co-lead of the trial and Professor of Emerging Infectious Diseases and Global Health at the University of Oxford, discussed the origins of the trial, the lessons learned and the future prospects for large, simple platform trials.

How did RECOVERY come together?

I've been working on epidemic infections for quite some time, and we've been trying to improve the speed with which we can start clinical studies for some years. If you go back to the 2009 influenza pandemic, it was clear that there was a big failure to do any good, meaningful RCTs. The next big challenge was the Ebola outbreak, where we managed to get trials started within a few months. But, still, these were started towards the end of the epidemic, and in that West Africa outbreak we really didn't get any good answers.

When this outbreak started, I was working with a Chinese colleague and his team to set up trials in Wuhan itself right away. We used some pre-prepared protocols that we had developed for the MERS coronavirus with colleagues in Saudi Arabia, and adapted those quickly. Within 20 days of the first announcement of the outbreak, we managed to enrol the first patient into a trial of [the HIV drugs] lopinavir and ritonavir. Not long afterwards, we started a second trial with [Gilead's antiviral] remdesivir in Wuhan.

But epidemics are a bit unpredictable, and very aggressive public health control measures in China actually meant that case numbers plummeted in Wuhan. We didn't reach the target sample size for either of those studies.

I had applied for additional funding to continue those trials, to turn these into a platform trial in China with Chinese colleagues. The funders rang me up and said "Yes, we'll fund you. But the cases are now in Europe. You can do the trial in the UK".

We had a sort of design and a pot of money. And that was the point that [University of Oxford's] Martin Landray, the other co-chief investigator of RECOVERY, and I got in contact. He'd had a discussion with The Wellcome Trust's Director Jeremy Farrar [on a bus], and Jeremy knew I'd been funded to do this research. This brought together two traditions within Oxford — emerging infections, which I've been working on for a long time, and large-scale pragmatic cardiovascular trials. It has been a wonderful marriage.

How did the protocol change with this union?

We had been in discussion with the World Health Organization about working with them on their Solidarity platform trial, but Martin and I looked at the protocol and decided it was too complicated. Martin picked up the protocols for the ISIS (International studies of Infarct Survival) trials, mega-trials from the 1980s looking at the effect of various drugs on survival following heart attacks. Martin got those protocols out of a drawer, and we decided that we should do something very simple like those.

In the UK, we have the advantage of data linkage. We knew that we could keep the follow-up very simple, and ascertain outcomes like death, particularly, but also intensive care admission and ventilation from existing data streams. That gave us the ability to simplify things.

The context was very clear: we were going to have a really big outbreak. There was a very real chance that the NHS would be completely overwhelmed. Only a simple trial would work in those very stressed circumstances. If people don't have time, they are not going to enrol patients into very complicated trials. As we didn't have much time, we also needed to get this thing started very quickly. And we realized that we were going to need a large number of patients, because there was not going to be a miracle drug. Severe viral respiratory infections are generally pretty difficult to treat. So we needed to have a trial that would detect a modest benefit. All of that pushed us to do a very simple trial.

Did you expect this trial to get so big?

No. We estimated that we should look for treatment effects of about a 20% reduction in mortality, and to see that on a background of around 20% mortality we would need about 2,000 patients versus 2,000 patients. So we always anticipated we were going to need thousands of patients. If I remember correctly, I think the original approval was for 12,000 patients. And we've had to keep pushing that up, and up, and up.

When we started we had four drugs and a control group, and we weren't sure we would finish all those.

We hadn't expected how things would turn out.

How do you decide which drugs to test?

With the first four drugs that were chosen, we wanted drugs that were available in the cupboard. There was a lot of discussion

> There's been a huge amount of wasted resource, on multiple underpowered or poorly designed trials

about repurposed drugs, because we could use them very quickly and they would have a well-understood safety profile. These were [the HIV drugs] lopinavir and ritonavir; azithromycin, a safe antibiotic but also an immunomodulator that is used in inflammatory lung disease; dexamethasone, because there was controversy about whether steroids should be used in lung infections; and hydroxychloroquine, because we thought everyone's starting to use it and we ought to find out if it really works or not.

After that, we were just inundated with emails of things to test, from the quite sensible to the, frankly, absurd. It quickly became clear that we couldn't manage that triage process. The UK government set up an independent committee called the COVID-19 Therapeutics Advisory Panel to look at all of these things. Anyone can submit a drug, and the committee screens these and does very good due diligence. Then they make recommendations, which we are free to accept or reject. So far, we've accepted pretty much all of their recommendations.

Why didn't you test remdesivir?

We tried to get it into the trial, but we could not get access to it.

Another drug that caused a bit of controversy is [Synairgen's] inhaled interferon. This was in the original protocol, but was dropped because the company and the investigators who were developing it decided that RECOVERY wasn't the appropriate vehicle. They started their own phase II trial, which subsequently showed a potential benefit. But it was a small trial. They've moved on to a larger trial.

To date, two out of seven completed treatment arms have shown efficacy. What's your take on this hit rate?

I guess that's not bad, based on normal drug discovery pipelines. Obviously, we want every one of them to work, and it's always disappointing when they don't. But taking a step back, it's probably not surprising given that they're mostly repurposed drugs to date, with probably a low probability of efficacy.

The failure that I found so far most disappointing is convalescent plasma. There were observational data that looked as if it might be beneficial. When we didn't see an overall benefit, that was pretty disappointing.

If you could start the trial again, would you make any design changes?

There are some routine data that we started to collect later — such as C-reactive protein levels and oxygen saturation measurements — that we could have collected from the start. The other change is with regard to factorial design [in which treatment arms are tested both alone and in combination with one another]. We did this later, but we perhaps could have done this from the outset.

Are there signals RECOVERY can't assess, because of its relatively simple setup?

I think we took the right approach, which was to be all inclusive and big enough so that we could do subgroup analyses. But we get asked a lot about activity in subgroups of subgroups, and things like that. You can take two views on this. One view is that if a drug is saving lives, does it matter? The other view is that we've only very crudely stratified patients by disease severity, but what about if a patient does or doesn't have a certain biomarker?

I think we made the right trade off. What you don't want — which is what you have got in some trials — is an uncertain result with a great deal of granularity. What you want is a certain result, and you can try and resolve the granularity later if the result is positive. For example, dexamethasone offers a clear benefit. Now, there are ongoing questions about the dose, the patient subgroups that benefit most and the duration of treatment. But these can be ironed out in subsequent trials.

What do you make of the rest of COVID-19 clinical trial landscape?

It's very clear that there's been a huge amount of wasted resource, on multiple underpowered or poorly designed trials. That's really disappointing, but I suppose not surprising because we're in a pandemic and everybody wants to try and do something.

There's also been a huge amount of noise generated from those underpowered trials, as well as from unreliable observational data. The noise around hydroxychloroquine, vitamin D and [the antiparasitic drug] ivermectin distracts from just doing big trials and getting convincing answers.

I have also been a bit surprised that some of the trials sponsored by pharmaceutical companies appear to be quite underpowered and small, given that there's no shortage of patients. It's hard to see the rationale for doing trials in fewer than 1,000 patients for a pharmaceutical company that's trying to get regulatory approval for a drug. I'm not sure what the barriers are to bigger trials, and I think that's an important point that needs some reflection.

What's next for RECOVERY?

There have been a lot of learnings from RECOVERY. Rather than run this trial in an academic research center, we took this trial to the frontline. It was done across over 175 hospitals in the UK, by frontline nurses. We sort of democratized the research process. We've had very positive feedback from that. I think that's a model for the future. We also leveraged data linkage.

We have linkage to the hospitalization data, ICU admissions data, prescribing data and national fatality statistics. Almost 100% of the primary and secondary outcomes data have come through national data streams. This has been enormously powerful, and shows that clinical trials don't necessarily need to be intensive form-filling exercises.

The next thing for us is to export the model to other contexts. This is one of the reasons that we recently expanded internationally [into hospitals in Indonesia and Nepal]. We want to see how we can export the model to places that haven't got a national health service, national research infrastructure or electronic data linkage. They could still benefit enormously from simplified trials that focus on what really matters.

But other contexts also means not just in emergency and in infectious disease settings.

I also hope that RECOVERY — as well as the REMAP-CAP trial — have made some pharmaceutical companies rethink clinical trials. Initially, companies would say to us "we're not very interested in your trial. It's too simple, too big, and it won't meet regulatory requirements". But they come back to us once their very complicated, underpowered trials have given them an equivocal result. I think this will make drug companies think again about whether they want to spend tens of millions on a trial in 1,000 patients, or work with a national platform trial that will cost them a lot less and will give them a more definitive answer.