

Lessons from COVID-19: How Data and Modelling Evolved During the Pandemic - audio summary

Google NotebookLM/WorldPop, 24 October 2025

Transcript

Speaker 1: Welcome to the Deep Dive. We streamline complex research so you can get straight to the essential insights. Now, before we jump in just a quick but important note, the voices you're hearing there are actually AI derived from source material that WorldPop uploaded, and this audio has been carefully edited, checked and validated by the experts at WorldPop.

So today our mission is to take a really insightful look at how disease modelling and use of data actually evolved during the COVID-19 pandemic. We're kind of tracing the story of the crisis through the eyes of the modellers themselves.

Speaker 2: And our sources for this Dive, they draw heavily from a brand new MOOD Horizon 2020 study - it just got published this week. It's led by Dr Esther van Kleef, who's a technical consultant for the World Health Organization, and co-authors are Dr Shengjie Lai and Professor Andy Tatem. It really offers an insightful look, as you said, at how disease modelling and data use changed throughout the pandemic.

Speaker 1: Let's unpack this then. So, we're looking specifically at researchers across the UK and key mainland European countries based on, was it 66 contributions? How exactly did the focus of their modelling work shift as the crisis just dragged on and on?

Speaker 2: Well, what the study really highlights is this clear evolution and priorities. It seems to move across 3 pretty distinct phases. So, in the early phase, let's say January to June 2020, the main goal was just sheer understanding. You know, modellers were focused on things like transmission parameters, trying to quantify the immediate burden of the disease. They're basically trying to figure out the size and the speed of the thing.

Speaker 1: That makes total sense. I mean, brand new pathogen emerges - you first need to understand its basic mechanics.

Speaker 2: Exactly. But then once we moved into the mid and late phases, so starting around summer 2020, the scope just brought and dramatically the focus shifted really from just understanding the disease to actively controlling it. So they move towards evaluating interventions specifically Non Pharmaceutical Interventions.

Speaker 1: Lockdowns, masks, that sort of thing.

Speaker 2: Precisely lockdowns, mask mandates and then later on they focused very heavily on vaccination strategies.

Speaker 1: And methodologically, when you look at how they modelled all this, you mentioned mathematical models were preferred for looking ahead.

Speaker 2: Yes descriptive statistics were used a lot too, but mathematical models really came into their own when the goal was anticipation trying to predict the pandemics' course and definitely for studying vaccine impact.

Speaker 1: OK, so the modellers knew the questions they needed to ask. But did they have the data, the fuel to run the engine, so to speak?

Speaker 2: Well, that brings us to the data story and this is maybe the most revealing part of the whole study. Traditional surveillance data, things like your daily case counts, incidents, reports - that stuff was generally available often in near real time.

Speaker 1: OK, standard stuff.

Speaker 2: But here's where it gets really interesting. More than half of all the modelling efforts reported that relevant crucial data was missing. It was the real time, non-traditional data that was persistently lacking. Things like real time social contact data. How are people actually mixing? Twenty three studies reported that missing.

Speaker 1: OK. That's huge for NPIs.

Speaker 2: Exactly. And attitude and behavioural surveys, you know, are people actually following the rules? How do they feel about interventions? Twelve studies missed that. And also serological data.

Speaker 1: Right. The blood test showing who'd already had it. Crucial for immunity estimates. So why was it missing? Was it funding, logistics? Or just not prioritised?

Speaker 2: The main reason reported by twenty seven studies was depressingly simple. It often just wasn't collected in the first place.

Speaker 1: It just wasn't part of the plan.

Speaker 2: Overwhelmingly, it seems it was a preparedness and protocol issue. The systems weren't set up before the crisis to capture that kind of dynamic behavioural data systematically.

Speaker 1: So, to evaluate if a lockdown works, you don't just need case counts. You need to know if people are actually staying home, reducing contacts. That's the behavioural gap you mentioned.

Speaker 2: And that failure to have systems ready for collecting social and behavioural metrics, that's a major takeaway here for preparedness.

Speaker 1: What's fascinating, then, and a bit surprising maybe, is who stepped in to fill some of these gaps? You mentioned private companies.

Speaker 2: Yeah, this was quite notable for-profit organisations think mobile phone companies, Google, Facebook. They actually ended up being relevant data providers for about a quarter of the studies.

Speaker 1: How did that work?

Speaker 2: Sometimes directly, sometimes through these 'Data for Good' initiatives that popped up.

Speaker 1: That reliance on the private sector, it shows adaptability, I guess, but raises big questions for the future, doesn't it?

Speaker 2: It certainly does. And speaking of adaptation and working together, the collaboration between the modellers and the public health authorities, the PHA's, was really frequent and

bidirectional. It shaped the work, it kind of matured over time too. Started more informal quick calls in the early phase, then it evolved into more formal participation, like sitting on high level advisory committees during the mid and late phases.

Speaker 1: OK, so this close collaboration, the modelling work itself. Did it actually translate into real world impact? Did policymakers listen?

Speaker 2: It seems they did. The impact looks pretty significant. About half the studies, 34 out of the 66 were cited at least once in an official policy document. And in total there were 121 citations across over 100 unique policy docs.

Speaker 1: That's substantial. Were certain types of studies cited more?

Speaker 2: Yes. Interestingly, the studies focused on anticipating the pandemics trajectory and those evaluating NPIs, they had the highest average citation rates. It really shows their direct usefulness for crisis decision making.

Speaker 1: So, boiling it all down, what's the key lesson for next time?

Speaker 2: The key implication seems twofold, really. First, for future preparedness, we desperately need to rethink our data collection systems. They need to be standardised, sustainable and go way beyond just basic case counts. We need systems with ready-to-deploy protocols specifically for capturing that social and behavioural data quickly and reliably.

Speaker 1: Right. Bake it in from the start.

Speaker 2: Exactly. And second, we need sustained investment in those cross-border collaboration networks before a crisis hits.

Speaker 1: OK, let me leave you our listener with a final thought and something provocative perhaps. If data from for-profit companies proves so critical this time around, how do we ensure ethical, sustainable long term access? That feels like a massive, complex question for future preparedness planners to wrestle with.

To read the full journal article that informed our discussion today, just follow the link below.