Roche buys into RET inhibitor showdown

Blueprint Medicines has [partnered with Roche](#) to sell the RET inhibitor pralsetinib, in a deal worth \$775 million in up-front payments and up to \$927 million in potential milestones. If the FDA approves pralsetinib later this year, the drug will compete with Lilly's first-in-class RET inhibitor Retevmo (selpercatinib).

RET, a proto-oncogene that encodes a transmembrane receptor tyrosine kinase, is mutated or aberrantly expressed in around 2% of cancers. This has made Ret an appealing target for precision therapeutics in oncology. Companies like Blueprint Medicines and Loxo Oncology developed their RET inhibitors initially for patients with RET-aberrant non-small-cell lung cancer, as well as other cancer indications with particular promise. Loxo was acquired for \$8 billion by Lilly in 2019 in part for access to Retevmo, and in May Lilly picked up an [FDA approval](#) for this small molecule. Blueprint completed its FDA filing for pralsetinib in March and is expecting a regulatory decision by the end of the November. In anticipation of facing off against deep-pocketed Lilly, Blueprint is now bringing Roche in on the future development and commercialization of the drug.

The \$1.7 billion biodollar deal — granting Roche exclusive license to commercialize the drug outside the United States, excluding greater China — will speed up Blueprint's ability to get its drug to patients, said the company's CEO Jeff Albers in a statement. The collaboration is transformative for the 12-year-old biotech, added Albers, because it enables the company to continue to further invest in its growing pipeline, which includes a systemic mastocytosis program. Blueprint and Roche, like Lilly, are evaluating further development opportunities for RET inhibitors, including [tissue-agnostic cancer indications](#) and combination strategies.

Published online: 4 September 2020
<https://doi.org/10.1038/s41587-020-0675-0>

Massive data initiatives and AI provide testbed for pandemic forecasting

Initiatives to gather massive epidemiological datasets aim to cut through national COVID-19 stats in a bid to understand the new coronavirus and aid public health policymakers.

Two population-scale COVID-19 studies in the UK, OpenSAFELY and the COVID-19 Symptom Study, are the first robust demonstrations of the power of big data to cut through the confusing mass of information and statistics generated by the pandemic to uncover biological signals. Although machine learning has the potential to deepen researchers' understanding of this new virus and how it affects its human host, most platforms capable of analyzing real-world data accurately are still immature given the lack of high-quality data required to develop such models. But even if many artificial intelligence (AI) initiatives may not have much of an impact on the present crisis, in the long run some could have a profound effect on future pandemic preparedness.

The OpenSAFELY team [analyzed the National Health Service \(NHS\) electronic health records](#) of over 17 million individuals — about 40% of England's adult population — to identify the main risk factors for COVID-19-related death. In addition to old age and the presence of underlying medical conditions, they found that Black and South Asian ethnicity were among the main risk factors for mortality. The [COVID-19 Symptom Study](#) collected self-reported

data from over 2.6 million people using a smartphone app developed by King's College London and London-based AI firm Zoe Global. By analyzing real-time reports from people with and without symptoms, the researchers established the combination of symptoms most likely to predict infection — including loss of taste and smell, which they were first to flag as a COVID-19 symptom.

Because of their scale, both studies have a statistical rigor that has been lacking in many other COVID-19 population studies. One of the most troubling uncertainties weighing on epidemiologists is the confusion about the actual number of cases. Behind the slick interfaces of the [many dashboards](#) developed to track the virus in real time lies a morass of messy, inconsistent data, which renders any cross-country comparisons invidious. "We've ended up with a jumble of metrics, some of which are useful, some of which are not useful," says Murray Aitken, executive director at the IQVIA Institute for Human Data Science. Patchy approaches to testing and uncertainties about both asymptomatic carriers and how long immunity persists in those who have recovered from the infection mean that epidemiologists' understanding of the pandemic remains incomplete. The

Machine learning can aid epidemiologists' understanding of the coronavirus pandemic, but not all comparisons will be useful if standards differ between countries. Pitinan Piyavatin / Alamy Stock Photo.

Table 1 | Selected AI and data science initiatives for tracking COVID-19

Participants	Project	Description	Status
Adaptive Biotechnologies, Microsoft	TCR-Antigen Map	Large-scale mapping of T cell responses to disease, combined with the development of machine learning models that can diagnose specific disease states	Ongoing initiative extended to COVID-19
C3.ai, Microsoft, University of Illinois, University of California, Berkeley, Carnegie Mellon University, Massachusetts Institute of Technology, Princeton University, University of Chicago, Stanford University, National Center for Supercomputing Applications, Lawrence Berkeley National Laboratory	C3.ai Digital Transformation Institute	Research consortium to advance AI and digital technologies to mitigate COVID-19 and future pandemics	First 26 research projects have received \$5.4 million funding, access to high-performance computing resources and C3.ai's Data Lake
IQVIA	IQVIA CARE Project	Web-based opt-in registry for US and UK residents to study symptom onset, disease progression and treatment outcomes	First 20,000 participants recruited
MIT Center for Collective Intelligence, Community Biotechnology Initiative at MIT Media Lab, MilliporeSigma	Pandemic Response CoLab	Online research collaboration for identifying COVID-19-related problems and proposing solutions	Early-stage initiative
Reagan-Udall Foundation for the Food and Drug Administration, Friends of Cancer Research and over 50 partners	COVID-19 Diagnostics Evidence Accelerator	Collaborative project to share and analyze real-world evidence on COVID-19 molecular diagnostics and serology tests, symptoms, test performance, and surveillance trends	Weekly meetings began on 28 May
Sophia Genetics, Paragon Genomics	Hospital-level multimodal analysis of COVID-19 patient data	Combined AI-powered analysis of genomic, lung radiomic and clinical data to predict likely outcomes and inform patient triage in COVID-19	Initial proof of concept established in lung cancer; COVID-19 study planned
University of Oxford, London School of Hygiene and Tropical Medicine, TPP (Leeds, UK), NHS	OpenSAFELY	Secure health analytics platform housing electronic health records for 40% of England's adult population	Published analysis of factors associated with 10,926 COVID-19-linked deaths among 17,278,392 patients
xCures	BEAT19 study	Population study of the natural history of COVID-19 symptoms in confirmed and suspected cases	Analysis of data from the first 4,000 participants will be published shortly
Zoe Global, King's College London, NHS, Massachusetts General Hospital, Lund University (Sweden)	COVID-19 Symptom Study	Analysis of population-level COVID-19 symptom data to detect COVID-19 symptoms and estimate national prevalence and incidence rates	Data collection app downloaded by over four million users in the UK, US and Sweden, who upload regular reports

Sources: organization websites

uncertainty is compounded by divergent reporting standards across and within countries and by the constant revision of official case data. “I can’t emphasize how chaotic the situation is,” says Aditya Prakash, associate professor in Georgia Institute of Technology’s school of computational science and engineering.

Traditional statistical analysis and machine learning can help to clean up messy or incomplete data. Zoe and Intellegens are among several firms (Table 1) repurposing machine learning capabilities developed in other disciplines for pandemic forecasting. Intellegens, which has been focused on a range of industries, including materials science and drug discovery, received funding from Innovate UK to apply its deep learning algorithm Alchemite to build a predictive

tool for governments and healthcare providers. The aim is to improve the algorithm’s forecasting accuracy so analysts can assess the likely impacts of different policy interventions.

Zoe was set up, not for COVID-19, but for weight loss and dietary health. Its initial focus was to apply machine learning to analyze individuals’ metabolic responses to food. Its COVID-19 work has demonstrated that it is feasible to gather high-quality data directly from users. “This is real citizen science,” says Zoe CEO Jonathan Wolf. “You can collect data from millions of people today,” he says. In addition to the symptom study, which was led by Zoe’s cofounder and long-term collaboration partner Tim Spector, of King’s College London, it has also recruited 800,000 volunteers to participate in

a clinical study to explore whether machine learning could serve as a digital diagnostic for COVID-19. Obtaining a data signal is difficult at present, however, because of the relatively low levels of coronavirus in the United Kingdom.

IQVIA and xCures have also been running community-based population studies, albeit at a smaller scale. IQVIA has so far recruited about 20,000 volunteers for its web-based CARE registry study and is following participants longitudinally. “The main focus is on progression of symptoms and severity,” says Nancy Dreyer, CSO at IQVIA’s Real World Solutions arm. “We have very little information in the epidemic about symptoms and symptom severity outside of the hospital setting.” xCures’ BEAT19 study has recruited about 4,000 volunteers

FDA okays second CAR-T for Kite

Gilead subsidiary Kite Pharma has [gained an accelerated approval](#) for Tecartus (brexucabtagene autoleucel), a CD19-directed autologous T cell therapy for relapsed or refractory mantle cell lymphoma. With the US Food and Drug Administration's go-ahead, Tecartus becomes the third FDA-approved CAR-T cell immunotherapy and the first CAR-T therapy for mantle cell lymphoma, an aggressive, hard-to-treat form of non-Hodgkin lymphoma. The chimeric antigen receptor (CAR) protein in Tecartus consists of a mouse single-chain variable fragment specific to human CD19 linked to signaling domains from CD28 and CD3 ζ . The approval, issued on 24 July, is based on data from the ongoing [Zuma-2 trial](#), in patients who had previously received up to five treatments, including chemotherapy, a CD20 monoclonal antibody and a Bruton tyrosine kinase inhibitor. Of 60 patients whose treatment could be evaluated for efficacy, [87% responded to Tecartus](#), including 62% who cleared their cancer completely.

Kite's other CAR-T therapy is Yescarta (axicabtagene ciloleucel), FDA approved in October 2017 for third-line treatment of relapsed or refractory diffuse large B-cell lymphoma. Tecartus and Yescarta express identical CARs. But Tecartus undergoes an extra T cell enrichment step needed for use in B cell malignancies in which there are lymphoblasts in patients' blood. This manufacturing step separates T cells from CD19-expressing tumor cells among the cells collected from patients by leukapheresis to prevent these tumor cells driving expansion and exhaustion of CD19-directed CAR-T cells.

[Kite was acquired by Gilead](#) for \$11.9 billion shortly before Yescarta's approval. But the buyout isn't yet living up to its lofty price tag: in the first half of this year, Yescarta pulled in just \$296 million in sales. To boost patient numbers, and hence sales, Kite is eyeing approvals in other indications. A phase 3 trial is testing Yescarta as a second-line therapy for diffuse large B cell lymphoma, and a phase 2 trial is underway as a first-line therapy in the same indication. Tecartus is in phase 2 trials for acute lymphoblastic leukemia and chronic lymphocytic leukemia.

Published online: 4 September 2020
<https://doi.org/10.1038/s41587-020-0676-z>

in the United States and is now tapping into Brazilian data through relationships with several medical centers there.

Machine learning is a relatively new addition to epidemiology forecasting. It is not a replacement for classical mechanistic modeling, but it can integrate unstructured and informal data from multiple data streams and can uncover hidden patterns within diverse datasets. In contrast, mechanistic models draw on more limited datasets but incorporate an understanding of the transmission dynamics of the disease outbreak. As Inga Holmdahl and Caroline Buckee, of the Harvard T.H. Chan School of Public Health, note in a [recent Perspective](#), the two approaches address different questions and have different limitations. Machine-learning forecasts are more suited to short-term predictions, which can help prioritize the allocation of healthcare resources, for example, whereas mechanistic models are used to examine long-term trends and the impacts of policy measures, such as encouraging social distancing or mandating face masks. The distinction between the two approaches is becoming blurred, Prakash says, by hybrid approaches that include elements of both.

The annual [FluSight challenge](#) run by the US Centers for Disease Control on influenza forecasting, in which Prakash participates, provides an important testbed for assessing new approaches to disease forecasting. Prakash is also one of several dozen researchers working on a large-scale [computational epidemiology initiative](#) funded by the US National Science Foundation, which aims to build sophisticated network models, operating at multiple scales and on multiple data layers, to develop insights into the control of epidemics and pandemics. The [C3.ai Digital Transformation Institute](#), an AI research consortium founded in March by the AI software firm C3.ai, Microsoft and several US academic institutions, is also developing modeling and AI-based tools to mitigate pandemics. Its first research awards are focused on a broad swathe of topics that intersect with COVID-19, including social issues, such as housing precariousness and the social determinants of health, as well as technical problems, such as mathematical modeling and computational biology.

For example, Vince Poor, professor of electrical engineering at Princeton University, and collaborators in Princeton, Carnegie Mellon University and the University of Pennsylvania are applying network engineering concepts to model the epidemiology of COVID-19. A key element of their approach is to incorporate a more nuanced description of R_0 , which is the average number of new infections

expected to arise in a naïve population from one infected individual. By modifying the 'susceptible, infected and recovered' (SIR) model, they aim to develop a more accurate picture of the spread of the virus. "Instead of applying a uniform R_0 across the entire population, the idea is to apply a probability distribution to the transmissibility of each individual in the population," he says.

At a completely different scale, genomics represents another domain that may be amenable to analysis with machine learning. In response to the COVID-19 pandemic, Adaptive Biotechnologies and Microsoft have extended an existing alliance to map T cell receptor (TCR) sequences to specific disease states. Adaptive is making TCR sequence data from de-identified geographically and ethnically diverse COVID-19-infected people available to vaccine and drug developers, to enable them to assess T cell responses in clinical trials. At the same time, the partners are developing a COVID-19 diagnostic powered by machine learning, which works by identifying all possible TCRs in a blood sample capable of binding a SARS-CoV-2 antigen. "Our big thesis is [that] the immune system is really, really good at detecting disease," says Harlan Robins, cofounder and CSO at Adaptive. Because TCRs only recognize antigen fragments presented by major histocompatibility molecules, the 'rules' for understanding TCR-antigen recognition are constrained. "It's a hard problem, but it's tractable," Robins says. The company aims to ship the test during the fourth quarter. "The diagnostic is complete. Now we have to validate it in a way the FDA agrees to."

AI-driven diagnostics based on lung imaging have received substantial attention, although some critics argue that the field is still at a nascent stage. "It's so easy to create machine learning models," says Roger Noble, founder and CEO of Oxford-based Zegami, which combines development of machine learning tools with image management and analysis software. "The difficulty is in making sure that they work correctly, that they're unbiased, that they're fair, and they work against the real-world data that they're seeing, not just the training and validation data that's been collected maybe in a more sterile environment."

Sophia Genetics, a Lausanne, Switzerland-based firm that cut its teeth in AI-powered analysis of genomic data for cancer, is convinced that a multimodal approach to COVID-19 will generate useful insights. This involves layering lung imaging data on top of other clinical and host and viral genomic data to predict the course of disease in patients. "You can probably only make predictions if you have multimodal

data,” says Philippe Menu, chief medical officer at Sophia Genetics.

In other domains, such as [facial recognition](#), [computer vision](#) and [police profiling](#), AI has been widely criticized because of a range of biases that result in racially skewed outcomes. In healthcare, those who would introduce AI in ways that may alter clinical practice have a high bar to clear in terms of equity as well as safety and efficacy. The [string of retractions associated with Surgisphere](#), the notorious healthcare analytics firm that claimed to have built a global database of electronic health records from hundreds of hospitals, epitomizes the fast-and-loose culture of data publishing that has thrived during the COVID-19 pandemic.

At the same time, clinging to the status quo may not represent the best response either. The OpenSAFELY project, which is jointly led by Ben Goldacre, director of the Evidence-Based Medicine DataLab at the University of Oxford, and Liam Smeeth,

professor of clinical epidemiology at the London School of Hygiene and Tropical Medicine, demonstrates the potential of unlocking the enormous data resources of the UK's NHS to better understand and improve health outcomes. Goldacre decries the “insufficient focus on the implementation of data science in healthcare.” The present crisis has [lowered the administrative and cultural barriers](#) that in more normal times would have either prevented or delayed the implementation of such a study. Crucially, its design and implementation have been exemplary, in stark contrast to the discredited studies based on Surgisphere's data. A handful of AI-based diagnostics in several indication areas have already gained approval, but the [regulatory framework](#) for assessing the utility of AI-based technologies in healthcare is still evolving. Whichever way these emerging technologies are assessed, the usual clinical parameters of safety and efficacy will still apply.

Of course, no smart technologies or sophisticated models can serve any useful purpose if political leaders and their supporters continue to flout basic public health guidance. The extraordinary trajectory of the pandemic in the United States and several other countries led by populist politicians who pay little heed to scientific advice is a stark illustration of the [price ordinary people pay](#) when the public health arena is transformed into a political battleground. The chaos that is a feature of the COVID-19 pandemic in some countries is as much a function of bad politics as it is of inconsistent reporting. No technology can alleviate that fundamental problem. □

Cormac Sheridan
Dublin, Ireland

Published online: 4 September 2020
<https://doi.org/10.1038/s41587-020-0671-4>

Around the world in a month

Credit: Map: © iStockphoto; Flags: pop_jop / DigitalVision Vectors / Getty

Published online: 4 September 2020
<https://doi.org/10.1038/s41587-020-0669-y>