

# Artificial intelligence, drug repurposing and peer review

Can traditional computational analysis and machine learning help compensate for inadequate peer review of drug-repurposing papers in the context of an infodemic?

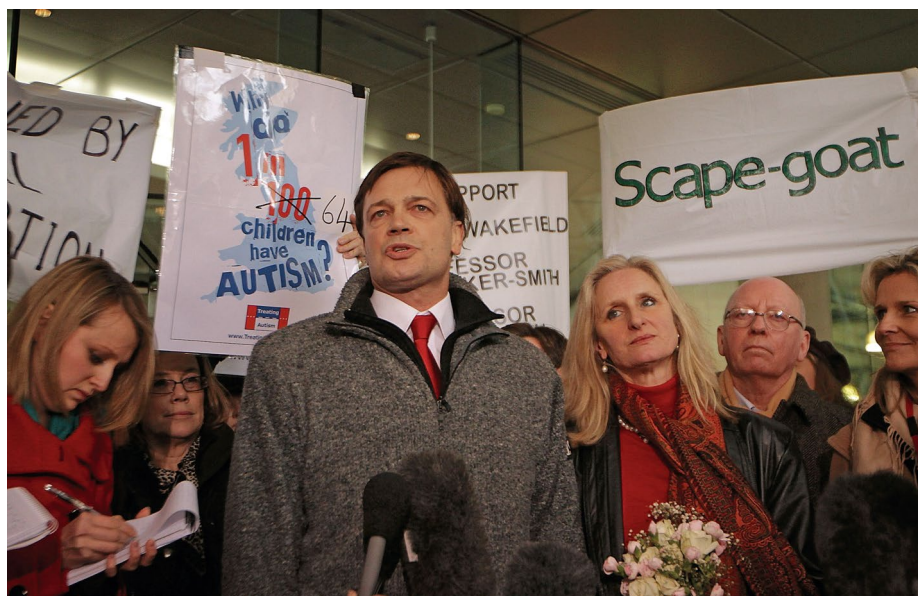
Jeremy M. Levin, Tudor I. Oprea, Sagie Davidovich, Thomas Clozel, John P. Overington, Quentin Vanhaelen, Charles R. Cantor, Evelyne Bischof and Alex Zhavoronkov

The COVID-19 pandemic has transformed the way scientific and clinical results are shared and disseminated. According to a recent analysis, an average of 367 COVID-19 papers are being published every week, with a median time from submission to acceptance of just 6 days (compared with 84 days for non-COVID-19 content)<sup>1</sup>. These unprecedented peer review turnaround times — and in some cases relaxed editorial standards — are justifiable in a context where new information may accelerate knowledge and solutions to the emerging global medico-socio-economic disaster, but they also risk the release of preliminary or flawed publications that can mislead research and development efforts, compromise clinical practice and misinform policy makers. What can be done to compensate for inadequate peer review in the context of a pandemic? Here, we propose a strategy whereby rigorous community and peer review is coupled to the use of artificial intelligence to prioritize research and therapeutic alternatives described in the literature, enabling the community to focus resources on treatments that have undergone appropriate and thorough clinical testing.

## When papers get it wrong

In 1998, Andrew Wakefield and his colleagues published a paper linking vaccinations and autism<sup>2</sup>. Twelve years after that publication — a misstep that turned hundreds of thousands of parents around the world against vaccinating their children because it implied a causal link between immunization and autism — *The Lancet* announced on 2 February 2010 that “several elements” of the Wakefield paper “are incorrect, contrary to the findings of an earlier investigation.”<sup>3</sup>

The risk of wide dissemination of less than rigorous clinical science is thus not a new problem; however, the COVID-19 pandemic has exacerbated the problem by precipitating the release of preliminary



A flawed paper from the group of Andrew Wakefield (pictured) linking vaccinations and autism<sup>2</sup> is archetypal of the types of risks inherent in wide dissemination of less-than-rigorous clinical science — a problem witnessed time and time again during the COVID-19 pandemic. Credit: Peter Macdiarmid / Staff / Getty Images News

research findings that may be subsequently revised in the light of new evidence — or in some case proven to be completely wrong. This is further compounded by the recent establishment of non-peer-reviewed preprint servers, which greatly lower the barrier to publishing and bypass critical peer review.

Given the lack of effective vaccines and validated therapeutic choices against SARS-CoV-2, biomedical scientists are racing to publish suggestions for rapidly deployable therapeutic options. Although novel chemical and biologic entities are being evaluated as potential therapeutics for SARS-CoV-2 infection, the repositioning and off-label use of existing agents approved for unrelated conditions is widely advocated as a therapeutic approach<sup>4</sup> against COVID-19 as it offers more rapid, actionable

interventions against the virus. Off-label use allowed several groups to report potential efficacy of some agents (for example, umifenovir<sup>5</sup> and remdesivir<sup>6,7</sup>) in clinical studies. Table 1 summarizes some of the repositioning proposals undergoing clinical evaluation, many of which are investigator initiated.

## The hydroxychloroquine saga

A case in point is the use of hydroxychloroquine (HCQ) and azithromycin in COVID-19. The first evidence for potential came from an in vitro study of a combination of remdesivir and chloroquine (CQ) — a drug structurally related to HCQ — in which growth inhibition of SARS-CoV-2 was reported<sup>4</sup>. The original paper was submitted to

**Table 1 | Drugs repositioned for COVID-19 based on information summarized by ASHP**

Type	Drug	Mode of action and approved use	Clinical experience in COVID-19
Off label	Oseltamivir	Neuraminidase inhibitor, effective against influenza infection	No evidence
Off label	Zanamivir	Neuraminidase inhibitor, effective against influenza infection	No evidence
Repositioning	Remdesivir (not an approved drug)	RNA-dependent RNA polymerase inhibitor, broad spectrum antiviral	Multiple clinical trials ongoing; approved under specific EUA guidelines by US FDA (since 1 May 2020) and available for compassionate use; remdesivir shortens time to recovery but does not improve survival in COVID-19 <sup>29</sup>
Repositioning	Umifenovir	Fusion inhibitor targeting hemagglutinin with in vitro activity against SARS-CoV-2, approved in China and Russia for influenza infection	Not FDA approved; no efficacy against COVID-19 established
Repositioning	Anakinra	Interleukin-1 receptor antagonist, for rheumatoid arthritis	A cohort study shows that anakinra improves survival in severe COVID-19 when compared with standard of care <sup>30</sup>
Repositioning	Tocilizumab	Interleukin-6 blocker, for rheumatoid arthritis	Used in China to treat severe COVID-19 in patients with high IL-6 levels <sup>31</sup> ; COVACTA (NCT04320615), a randomized, placebo controlled trial, did not meet its primary endpoint of improved clinical status in patients with COVID-19-associated pneumonia
Repositioning	Azithromycin	Macrolide antibiotic that blocks translation by binding to the 50S ribosomal subunit; shows antiviral activity in vitro	Effective as adjunct therapy in respiratory infections; a small study claimed effectiveness when combined with HCQ
Repositioning	HCQ	Antimalarial endosomal acidification fusion inhibitor, used in autoimmune diseases (lupus, rheumatoid arthritis)	Small study claimed effectiveness when combined with azithromycin; temporary approval under EUA guidelines by FDA between 28 March and 15 June; not approved in Europe
Repositioning	CQ	Antimalarial endosomal acidification fusion inhibitor, used in autoimmune diseases (lupus, rheumatoid arthritis)	Temporary approval under EUA guidelines by FDA between 28 March and 15 June; not approved in Europe
Repositioning	Dexamethasone	Glucocorticoid receptor agonist; immunosuppressive agent	Following a trial in the UK, dexamethasone is recommended for mechanically ventilated patients with COVID-19 or those who require supplemental oxygen
Repositioning	Baricitinib	Janus kinase (JAK1 and JAK2) inhibitor, used in rheumatoid arthritis	Contraindicated in viral infections and respiratory diseases; no evidence yet for clinical efficacy <sup>32</sup> in COVID-19

ASHP, American Society of Health-System Pharmacists.

*Cell Research* on 25 January, accepted on 28 January after just three days in peer review, and published on 4 February. This publication was followed by a letter to the editor in *BioScience Trends*<sup>8</sup>.

In early March, a medical group from IHU-Méditerranée Infection, Marseille, France used HCQ, a related antimalarial drug used off-label for autoimmune diseases such as systemic lupus erythematosus, to treat patients infected by SARS-CoV-2. The results of this open-label, non-randomized clinical trial were submitted to *International Journal of Antimicrobial Agents*<sup>9</sup> on 16 March, accepted on 17 March and published on 20 March. This report suggested that HCQ, in combination with azithromycin, successfully clears SARS-CoV-2 infections<sup>9</sup>. Upon close examination, however, this paper reported data from only 14 patients (in a HCQ monotherapy treatment arm)

and 6 patients (undergoing combination treatment with azithromycin to prevent bacterial superinfection) out of a total of only 26 patients. Despite the fact that the authors correctly mentioned the small sample size and very short follow-up time in their study, they nevertheless recommended that both drugs could be used as a curative and preventative therapy for COVID-19.

Several other trials followed up on this report, together with many publications in the lay press. As yet, the outcomes from these trials are inconclusive<sup>10</sup>. The original study in March 2020 led, however, to an enormous level of activity and focus on HCQ, with serious ramifications across the industrial, medical, political and societal landscape. Although rigorous, evidence-based results rooted in more than a century of experience in drug safety (the Pure Food and Drug Act of 1906)

are lacking, massive efforts have focused on HCQ as a COVID-19 therapy, with global implications. Major multinational companies responding to requests from government leaders ramped up manufacturing of CQ and HCQ. Other countries have hoarded the products, while regulatory authorities are rushing through emergency approvals of the drugs without data on effective dosage or safety protocols. None of the related studies explicitly referred to existing posology and pharmacokinetic properties of the drugs (for example, dosage, half-life and clearance) that are essential for application guidance and approval in original indications.

The clinical community did not wait for the results of more conclusive trials. And while the initial endorsement by the US Food and Drug Administration (FDA) was revoked in fewer than three months (see

**Box 1 | AI/ML in drug repurposing for COVID-19**

According to Mount Sinai School of Medicine's COVID-19 Drug and Gene Set Library<sup>33</sup>, at least 81 studies have focused on computational drug repurposing for COVID-19, compared with 27 studies based on experimental screening data. Nine months into the pandemic, there remains a paucity of confirmatory data on any of the repurposed drugs. For example, the antiviral activity of remdesivir, one of the most prominently discussed drugs repurposed for SARS-CoV-2, has been independently shown in just six independent studies; for ten other repurposed drugs, only one further confirmatory study has been published since the initial report<sup>33</sup>. This illustrates the uncertainties regarding reproducibility and rigorous evidence with respect to SARS-CoV-2 and drug repurposing facing the scientific community. Another application of AI/ML aims to identify

patterns of viral spread and disease course likelihood in different patient groups.

According to the Milken Institute, at least 20 efforts to repurpose drug candidates are underway, some using experimental approaches and others using AI/ML or a combination. Few peer-reviewed reports of AI/ML in drug repurposing have been published, with baricitinib the best example to date<sup>32</sup>. Various AI/ML models have been explored for prioritizing pharmacological targets for further investigations<sup>34</sup>. A dedicated AI/ML suite, REDIAL-2020 (<http://drugcentral.org/Redial>), uses six independent models related to viral entry, live virus infectivity and viral replication to estimate anti-SARS-CoV-2 activities from molecular structure. Similarly, the inClinico platform developed by one of us (A.Z.) incorporates multiomics data and clinical data into models to predict clinical trial outcome.

below), many self-medicating patients and physicians prescribing HCQ ignored or criticized the subsequent study results<sup>11,12</sup>. As evidence of lack of efficacy of HCQ has continued to be published, proponents of the treatment put forward further factors, such as the lack of zinc salts in combination with HCQ, as an explanation for the negative clinical results.

**Lessons learned**

The one-day peer-review process that led to the publication of the *International Journal of Antimicrobial Agents*<sup>9</sup> paper is highly irregular. Had a normal peer-review process occurred, it would have likely ensured that the publication that triggered this event would have had to reference and discuss a prior failed clinical study on HCQ by Chen et al.<sup>13</sup> published on 6 March in the *Journal of Zhejiang University*. In that controlled, open-label study of 30 patients with COVID-19, 13 of the patients who were on the drug tested negative for the virus after seven days, compared with 14 of the people on placebo. One of the patients on HCQ went on to develop severe illness, and the median recovery time was similar in the two groups. The study, albeit on only a small cohort of patients, concluded that HCQ is no more effective than the standard of care.

The article published by the group from IHU-Méditerranée Infection presents contradictory data based on a small sample of patients. It makes ambitious claims on the

basis of weak evidence that appears to have undergone minimal peer review, but was published in a respectable peer-reviewed journal. On 19 March, after this article was submitted but before it was published, the White House press secretary announced that HCQ had shown encouraging early results against COVID-19. This was followed by a blizzard of Twitter- and television-based communications claiming a “100% cure” of COVID-19 via treatment with HCQ, azithromycin and zinc sulfate, with no need for hospitalization or ventilators. A preprint posted on 10 April (which remains to be published in a peer-reviewed journal) of a Wuhan, China-based randomized clinical trial showed shorter remission time and body temperature recovery time upon administration of HCQ in 62 patients with mild SARS-CoV-2 infection<sup>14</sup>, whereas a peer-reviewed study in 150 Chinese patients reported no difference in viral conversion rate<sup>15</sup>.

On 28 March, the FDA (but not the European Union) granted Emergency Use Authorization (EUA) for both CQ and HCQ for certain hospitalized patients with COVID-19. The FDA revoked the EUA on 15 June, after a series of addendum warnings about HCQ side effects, especially the potential for fatal arrhythmia due to QTc prolongation, as well as other cardiac events. Despite more evidence gathered over the past months, enforcing the lack of benefit of HCQ in prophylactic

settings, people continue to advocate prophylactic self-medication<sup>16</sup>. Neither the study published by the group from the IHU-Méditerranée Infection nor any of the subsequent studies embracing HCQ as an effective treatment for COVID-19 has turned out to be reproducible. And yet, despite their scientific weakness, they became the object of political attention, causing a global shortage of these drugs that affected patients with legitimate therapeutic needs receiving the drugs for autoimmune disease<sup>17</sup>.

The two initial HCQ reports bypassed existing scientific checks and balances, highlighting the importance of sound peer review, which is particularly important when disclosing novel therapies during a global crisis. Ensuring that scientific publications undergo rigorous peer review, even in times of emergency, is paramount<sup>18</sup>. Moreover, in the case of HCQ, the central tenet of medical practice — *primum non nocere* (first, do no harm) — was compromised because (i) drug shortages meant that patients with a legitimate medical need for the therapy (for example, for rheumatoid arthritis) could not procure it, and (ii) those patients who did receive the HCQ/azithromycin combination gained no benefit, but were exposed to the therapy's increased risk of cardiovascular mortality<sup>14,19,20</sup>.

**What to do about misinformation?**

It is critical we learn from the HCQ experience as we look for immediate medical solutions to this crisis. The COVID-19 pandemic has resulted in an unprecedented number of other therapeutic proposals, from peer-reviewed and preprint publications to blog posts, tweets, television and other communication channels. No matter how rational or sound, this many proposals cannot be systematically evaluated and prioritized by any single group, institution or regulator. There is a clear need to understand the ever-growing stream of data, information and knowledge being published in order to collate, process and structure it in real time.

In this respect, dictionary-based text mining<sup>21</sup>, coupled with specialized artificial intelligence (AI) or machine learning (ML) originally called statistical pattern recognition, such as BioBERT, a bidirectional biomedical language representation model<sup>22</sup>, can help achieve that potential. Rigorous, evidence-based peer review coupled with open data and computer-aided technologies offer a way out of the dilemma and provide the opportunity to make reasoned scientific breakthroughs in a crisis (Box 1). Although preprint services allow rapid dissemination

**Table 2 | Examples of data science and AI/ML techniques for drug repositioning for COVID-19**

Approach	Target and network analysis	Compound-based virtual screening	Drug-specific filters
Data science	COVID-19 literature text mining for target enrichment	Ligand-based similarity	Reverse L1000 genome signatures
	SARS-CoV-2-human protein interactome models		Drug exposure, compartment and tissue accumulation, toxicity and efficacy
AI/ML	Model development based on proteomics data	Analysis of antiviral activity from in vitro drug screen	Drug-drug interactions and drug synergy
		Molecular docking (SARS-CoV-2 and human proteins)	

of new findings, they are not peer reviewed and can easily mislead non-experts and encourage sensationalist and erroneous public coverage. At present, there are several explainable AI/ML systems designed to predict the outcomes of clinical trials on the basis of multiomics data, chemistry and textual information in prospective validation studies. These systems may be used to caution the regulators and encourage more careful review in cases where predictions disagree with the published results. Comprehensive end-to-end AI-powered drug discovery systems such as *pharma.ai* (<http://insilico.com/platform/>) integrating the preclinical and clinical datasets and providing target hypotheses; small molecule screening; and grant, publication, patent and clinical trial data analysis may be used for both repurposing and prediction of clinical trial outcomes.

Clinical drug development requires distinct areas of expertise. Unbiased clinical data, compiled in real time, ought to be made accessible without restrictions, and intellectual property rights should be waived for this and future pandemics. Although coordinated efforts are ongoing, this pandemic offers a unique opportunity to lay the foundation for synchronized global workflows that will ensure data veracity, provide an unbiased and multi-viewpoint assessment of therapeutic alternatives, and allow efficient allocation of computational, human and experimental resources. This must occur in the context of allowing peer review, fact checking and incorporation of relevant domain expertise.

Traditionally, such activities would take place in laboratories and be followed by human clinical trials. Given the almost

complete shutdown of animal research facilities and need for focused rational clinical trials, it is worthwhile to explore how many of these activities can be supplemented or even replaced by the capabilities that are now available through in silico technologies, such as ML, systems biology and computer-aided drug repurposing. These technologies have matured in recent years and are ready to become an integral part of the global workflow, to prioritize novel drug targets and new chemical entities as well as to evaluate off-label or drug repositioning proposals<sup>22</sup>. Such workflows — based, for example on Drug Repositioning Evidence Level (DREL)<sup>23,24</sup> — could be used to evaluate drug repositioning candidates. By integrating multiple layers of data, information and knowledge and processing the massive stream of repositioning proposals, validated machine-intelligence-based methods could serve, in the near future, as a decision support system for policymakers, healthcare providers and society at large<sup>25</sup>. If properly resourced and implemented, such a synchronized workflow could assist in assembling disparate evidence and hypotheses into actionable healthcare solutions to tackle the current and future inevitable pandemics.

In practice, the deployment of AI/ML methods requires a comprehensive understanding of their advantages and weaknesses. AI/ML is powerful for identifying relevant patterns within large set of nonlinear data without the need for manual feature engineering as systems can learn implicit rules from the data provided. While the amount of data needed to train such algorithms might be an issue, the ability

of AI/ML to make sense of large amount of data is an advantage in many circumstances. For instance, the Smith–Waterman algorithm and Pfam are standard methods for the prediction of protein functions, but they are not fast enough to handle large number of protein sequences. AI/ML offers alternatives to address both issues. For instance, DeepFam<sup>26</sup> is an alignment-free method extracting functional information from sequences without requiring multiple sequence alignments. When compared to state-of-the-art methods, DeepFam performed better in terms of accuracy and runtime for predicting protein functions. In this context, the emergence of AI/ML approaches is incremental and can be built on classical sequence similarity and genome analysis with tools such as BLAST.

ML is already starting to be used to identify biological targets for therapeutic intervention in heterogeneous disease<sup>27</sup> and find suitable drug candidates that bind those targets (Table 2). Similarly, if empirical clinical observations in a paper are used to propose a drug as a potential treatment approach, ML could and should be used to rapidly simulate efficacy and side effects in (preferably stratified) populations. A synchronized workflow using ML methods could be based on resources available for analyzing targets, drugs and related potential side effects, such as the Side Effect Resource (SIDER), which combines data on drugs, targets and side effects recorded during clinical trials, and the FDA Adverse Event Reporting System (FAERS), which gives access to adverse event reports and medication error reports previously submitted to FDA.

There are still several unknowns about the biology and mode of action of SARS-CoV-2. However, information about the sequence of the viral genome, discovery of receptors used by the virus to infect cells and knowledge of structure of the virus allow the identification of potential targets for direct-acting antivirals. As data have emerged on population-scale pathology, it has become clear that an overactive host immune response is a clear driver of more serious disease. Naturally, the data gathering, analytics strategies and focus for therapy discovery have responded to these data.

Furthermore, ML algorithms can accelerate the design of clinical trials by automatically identifying suitable subjects, ensuring the correct distribution to groups of study participants and providing an early warning system for a clinical trial that is not producing meaningful results. Computational drug repurposing accelerates the drug development process and reduces the associated costs. To identify the right

repurposing candidates, it is important to identify known molecular targets, to predict novel molecular targets for known drugs, and to consider dosing, pharmacokinetic and safety-related parameters. With its ability to analyze millions of examples of drug and patient data to generate hypotheses and then provide evidence supporting or challenging them, ML can be used to identify new indications for known drugs and to combine existing drugs in ways that give them therapeutic powers that each lacks in isolation.

Within healthcare and drug discovery, AI/ML should be implemented as an adjunct to human workflows rather than as an alternative. Integration of AI/ML offers powerful options, but can only be successful within multidisciplinary teams that can ensure AI/ML solutions are adapted to each particular situation. From this viewpoint, human expertise and final decision-making will remain essential in drug discovery and development, as well as in clinical practice. There are currently few examples of large-scale integration of AI/ML technologies in drug discovery or clinical practice. In drug discovery, where the timeframe for a drug to undergo preclinical testing and clinical trials is especially long, more time is needed to assess their real impact.

AI/ML algorithms already deployed within the drug development pipeline have greatly improved. For instance, the synthetic tractability that was a weakness of the first AI/ML de novo design methods can now be evaluated using synthetic accessibility scores. When properly designed in collaboration with medicinal chemist experts, platforms for de novo design can prioritize synthetically tractable molecular structures with the desired biological activity. Moreover, state-of-the-art AI-based methods for de novo design can generate molecular structures using restricted information. Binding-site amino acid environment and cocrystallized fragment, for instance, provide the pocket and ligand features needed to perform either ligand-based or pocket-based generation. Nevertheless, challenges encountered when developing AI/ML solutions for de novo molecular generation or for medical imaging prognosis<sup>28</sup> also demonstrate that there is a need to develop and improve reporting standards and metrics, as well as best practices for data sharing and a requirement for algorithm availability, which should be adapted to the strict requirements and expectations of medical sciences and healthcare.

## Conclusions

The COVID-19 pandemic has highlighted the need for new tools to complement existing peer-review mechanisms for ensuring the veracity and robustness of the biomedical information published to guide clinical practice and shape public health policy. It has also shown that the research community is capable of generating large amounts of heterogeneous data in a short period of time — information on a scale and speed that can confound human interpretation. In this context, we believe that AI/ML has a critical role in bolstering data to supplement peer-reviewed papers. We call for the rapid development and prospective validation of comprehensive, explainable AI/ML systems that use preclinical and clinical data, capable of not only rapidly predicting clinical trial outcomes, but also highlighting possible flaws in published work and the features contributing to the increased probability of failure. □

*Editorial note: This article has been peer reviewed.*

Jeremy M. Levin<sup>1,2</sup>, Tudor I. Oprea<sup>3,4,5,6</sup>, Sagie Davidovich<sup>7</sup>, Thomas Clozel<sup>8</sup>, John P. Overington<sup>9</sup>, Quentin Vanhaelen<sup>10</sup>, Charles R. Cantor<sup>11</sup>, Evelyne Bischof<sup>10,12,13,14</sup> and Alex Zhavoronkov<sup>10,14</sup>

<sup>1</sup>Ovid Therapeutics, Inc, New York, NY, USA.

<sup>2</sup>Biotechnology Innovation Organization (BIO),

Washington, DC, USA. <sup>3</sup>Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM, USA. <sup>4</sup>Autophagy

Inflammation and Metabolism Center of Biomedical Research Excellence, University of New Mexico Health Sciences Center, Albuquerque, NM, USA.

<sup>5</sup>Department of Rheumatology and Inflammation Research, Institute of Medicine, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden. <sup>6</sup>Novo Nordisk Foundation Center for

Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. <sup>7</sup>SparkBeyond, Netanya, Israel. <sup>8</sup>Owkin,

Paris, France. <sup>9</sup>Medicines Discovery Catapult, Cheshire, UK. <sup>10</sup>Insilico Medicine, Hong Kong Science and Technology Park, Hong Kong, China. <sup>11</sup>Retrotape,

Inc, Los Altos, CA, USA. <sup>12</sup>Shanghai University of Medicine and Health Sciences, College of Clinical

Medicine, Shanghai, China. <sup>13</sup>Department of Advanced Biomedical Sciences, Federico II University

of Naples, Naples, Italy. <sup>14</sup>These authors contributed equally: Evelyne Bischof, Alex Zhavoronkov.

✉e-mail: bischofevelyne@gmail.com;

alex@insilico.com

Published online: 14 September 2020  
<https://doi.org/10.1038/s41587-020-0686-x>

## References

- Palayew, A. et al. *Nat. Hum. Behav.* **4**, 666–669 (2020).
- Bhatt, R. *Lancet* **351**, 1357 (1998).
- Eggertson, L. *CMAJ* **182**, E199–E200 (2010).
- Wang, M. et al. *Cell Res.* **30**, 269–271 (2020).
- Huang, D. et al. *J. Med. Virol.* <https://doi.org/10.1002/jmv.26256> (2020).
- Touret, F. et al. *Sci. Rep.* **10**, 13093 (2020).
- Mehta, P. et al. *Lancet* **395**, 1033–1034 (2020).
- Gao, J., Tian, Z. & Yang, X. *Biosci. Trends* **14**, 72–73 (2020).
- Gautret, P. et al. *Int. J. Antimicrob. Agents* **56**, 105949 (2020).
- Yao, X. et al. *Clin. Infect. Dis.* **71**, 732–739 (2020).
- Cavalcanti, A. B. et al. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2019014> (2020).
- Tilangi, P., Desai, D., Khan, A. & Soneja, M. *Lancet Infect. Dis.* [https://doi.org/10.1016/S1473-3099\(20\)30430-8](https://doi.org/10.1016/S1473-3099(20)30430-8) (2020).
- Chen, J. et al. *Zhejiang Da Xue Xue Bao Yi Xue Ban* **49**, 215–219 (2020).
- Chen, Z. et al. Preprint at *medRxiv* <https://doi.org/10.1101/2020.03.22.20040758> (2020).
- Tang, W. et al. *Br. Med. J.* **369**, 1849 (2020).
- Cohen, M. S. N. *Engl. J. Med.* **383**, 585–586 (2020).
- Molina, J. M. et al. *Med. Mal. Infect.* **50**, 384 (2020).
- Touret, F. & de Lamballerie, X. *Antiviral Res.* **177**, 104762 (2020).
- Lane, J. C. E. et al. Preprint at *medRxiv* <https://doi.org/10.1101/2020.04.08.20054551> (2020).
- Borba, M. G. S. et al. *JAMA Netw. Open* **3**, e208857 (2020).
- Cook, H. V. & Jensen, L. J. *Methods Mol. Biol.* **1939**, 73–89 (2019).
- Lee, J. et al. *Bioinformatics* **36**, 1234–1240 (2020).
- Aliper, A. et al. *Mol. Pharm.* **13**, 2524–2530 (2016).
- Oprea, T. I. & Overington, J. P. *Assay Drug Dev. Technol.* **13**, 299–306 (2015).
- McCall, B. *Lancet Digit. Health* **2**, e166–e167 (2020).
- Seo, S., Oh, M., Park, Y. & Kim, S. *Bioinformatics* **34**, i254–i262 (2018).
- Zeng, X. et al. *Chem. Sci.* **11**, 1775–1797 (2020).
- Nagendran, M. et al. *Br. Med. J.* **368**, m689 (2020).
- Beigel, J. H. et al. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2007764> (2020).
- Huet, T. et al. *Lancet Rheumatol.* **2**, e393–e400 (2020).
- Guaraldi, G. et al. *Lancet Rheumatol.* **2**, e474–e484 (2020).
- Titanji, B. K. et al. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciaa879> (2020).
- Kuleshov, M. V. et al. *Patterns (N Y)* <https://doi.org/10.1016/j.patter.2020.100090> (2020).
- Zeng, X. et al. *Proteome Res.* <https://doi.org/10.1021/acs.jproteome.0c00316> (2020).

## Acknowledgements

T.I.O. was supported by the US National Institutes of Health, U24 CA224370. E.B. was supported by Krebsliga Schweiz, BIL KFS 4261-08-2017.

## Competing interests

J.M.L. works for Ovid Therapeutics, a public biotechnology company developing medicines for rare neurological diseases, and is the chairman of the Biotechnology Industry Organization. A.Z. and Q.V. are affiliated with Insilico Medicine, a company developing artificial intelligence solutions for target discovery, small molecule chemistry and prediction of clinical trial outcomes. A.Z. is the CEO of Deep Longevity, an artificial intelligence company. S.D. is CEO of SparkBeyond, an artificial intelligence company. T.C. is CEO of Owkin, an artificial intelligence company specializing in clinical research. T.I.O. has received honoraria from or consulted for Abbott, AstraZeneca, Chiron, Genentech, Infinity Pharmaceuticals, Merz Pharmaceuticals, Merck Darmstadt, Mitsubishi Tanabe, Novartis, Ono Pharmaceuticals, Pfizer, Roche, Sanofi, Wyeth and Insilico Medicine. J.P.O. was previously employed by Pfizer, Inpharmatica, EMBL-EBI and BenevolentAI, and has received honoraria or consulting fees from Insitro, Sanofi and Boehringer Ingelheim. C.R.C. is a founder and a director of Retrotape and an advisor to Insilico Medicine. E.B. is an advisor to Insilico Medicine.