Machine Learning can Help Predict New Uses for Existing Drugs

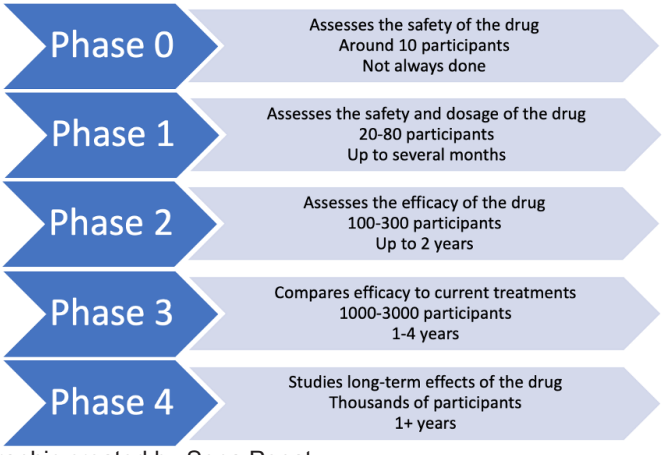
Sona Popat

In an age of antibiotic resistance and viruses becoming more transmissible, new drugs need to be developed faster than ever. But drug development is a long process: The pipeline from a potential drug candidate to an approved and usable drug can take up to several decades, and diseases are evolving too fast for drug development to keep up. One way scientists aim to make drug development a more efficient process is by repurposing drugs: using currently available drugs to treat different diseases - diseases other than the ones they were first developed to treat.

Drug repurposing is useful for rapidly responding to new outbreaks of disease as well as for rare diseases, where there may not be enough people to conduct large-scale clinical trials. To make drug repurposing more efficient, machine learning techniques can be used to identify drug repurposing candidates. A new study at The Ohio State University, led by Dr Ping Zhang, has developed a machine learning framework that identifies drug repurposing candidates more efficiently than ever before.

Why does it take so long to develop new drugs?

Typically, drugs are designed to stop disease progression by targeting components of cellular mechanisms. For example, tumours arise due to uncontrolled cell growth and replication, so components of the signalling telling cells to grow and replicate may be of interest to target and inhibit tumour growth. From initial studies into drug targets to potential drug candidates being studied in cell cultures and animal models, it is already a lengthy process before drugs even reach clinical trials in humans. There are several phases of clinical trials for drugs to pass through before they can be approved:



Graphic created by Sona Popat.

In Phase 0, scientists use small numbers of healthy volunteers and low doses of the treatment to confirm it is not harmful, and Phase 1 progressively scales the dose of the treatment upwards to identify the maximum dose of the drug humans can withstand before side-effects are seen. Phase 2 tests for side effects and aims to get a rough idea of whether the treatment is effective. By Phase 3, trials are usually randomised and double-blind: patients are randomly put into groups to receive either a new drug, a control, or a placebo, without the patients or scientists knowing which individual is in which group.

This process ensures no bias colours the outcomes of the trial due to the placebo effect or the way that scientists interpret the results. This phase is much larger-scale, involving thousands of patients across a range of demographics so the conclusions are more reliable and generalisable across groups. If the new treatment is safer and more effective at treating the disease than current treatments, it could go on to be approved for market distribution. Further trials, known as Phase 4, may investigate long term benefits and side-effects, even after the drug has been approved. Due to the lengthy processes and, depending on the rarity of the disease, the difficulty in obtaining enough volunteers for these trials, it can take an extremely long amount of time for a drug to be approved.

Enter drug repurposing

Drug repurposing has already been used to treat a wide range of diseases. For example, Viagra was initially developed to treat hypertension but is currently used for erectile dysfunction, and anti-inflammatory drugs made to treat arthritis are now being used for COVID-19 patients. The main benefit of drug repurposing is the shortening of the clinical trial process. As there is no need to repeat the initial safety phases, these existing drugs can be immediately tested for efficacy in treating the new disease, making the process safer, faster, and more cost-effective. Candidates for drug repurposing are usually identified using the results of previous research, genomics, or by considering drug action mechanisms that could be effective and common to multiple diseases, but the development of artificial intelligence and machine learning has sped up this process dramatically.

Machine learning to identify drug repurposing candidates

Machine learning (or deep learning) frameworks are algorithms that improve automatically through experience, where the algorithm can receive a "training" dataset to learn from, and then apply that knowledge to different datasets. Ohio State University scientists created a deep learning framework that uses real-world data from patients to identify drug repurposing candidates and their estimated effects to treat a particular disease.

This is the first framework of its kind to use real-world data instead of just preclinical data. The framework receives insurance claims data (e.g. information on assigned treatments, disease outcomes, patient demographics) as inputs, and then emulates a randomised clinical trial for each drug. By comparing pre-defined measures of disease status before and after treatment, the model identifies drugs that have positive disease outcomes (i.e. an improvement in the predefined symptoms and disease status of interest) as potential candidates to be repurposed.

But does this framework actually work?

To test their framework, scientists inputted a dataset of over 1.1 million coronary artery disease patients with the aim of identifying potential drug repurposing candidates for coronary artery disease treatment. The pre-defined measures of disease outcomes selected to assess whether drugs would be effective at treating coronary artery disease included a lower risk of heart failure and stroke. Of the nine top hits the framework produced, three were drugs that are currently used to treat coronary artery disease, confirming that the framework could successfully identify effective drugs. The other six drugs identified have not been used to treat coronary artery disease before, so existing literature was used to learn about their mechanism of action in the body and see if they are plausible drugs for treating coronary artery disease. Existing research supports the idea that these six drugs are likely to be effective at treating coronary artery disease: They act on areas common to coronary artery disease and heart failures, such as blood pressure and vasculature. Of course, before any of these drugs could be administered to patients to treat coronary artery disease, they would require clinical trials to confirm their efficacy, optimum dosage, and safety.

What's next? Coronary artery disease and beyond

When compared to three other machine learning frameworks that have been developed to identify drug repurposing candidates in the past, the new model created by Dr Zhang's group at Ohio State consistently outperformed the other models. Despite the success of testing the model, Dr Zhang has emphasised that the outcome of this single case study is less important than the development of the framework itself.

"My motivation is applying this, along with other experts, to find drugs for diseases without any current treatment," Zhang said. "This is very flexible, and we can adjust case-by-case. The general model could be applied to any disease if you can define the disease outcome."

Not only can drug repurposing speed up the drug development process to allow rapid responses to new outbreaks of disease, but it is also an exciting prospect to identify therapeutic options for diseases that currently do not have any treatments. For example, for rare diseases, there may not be enough data or patients available for the long drug development and clinical trials process. So, while identifying the six potential drug repurposing candidates for coronary artery disease is promising, it is more exciting to consider the future uses of this model for diseases that need new treatments.

But this is still only the beginning: Machine learning frameworks like this can be expanded to include more demographic information that may be important for disease outcomes, such as race and medical history. This is important as drugs have historically been designed based on data where ethnic minorities are underrepresented, so it is essential that technology is used to close this gap instead of reinforcing it. The rapid development of technology and machine learning such as this has the potential to revolutionise the drug development process in future.

REFERENCES

COVID-19 rapid evidence summary: tocilizumab for COVID-19. NICE. Updated February 2021. Available at: https://www.nice.org.uk/advice/es33/chapter/Product-overview [Accessed 26 February 2021].

Did You Know That Viagra Was Originally Developed For The Treatment Of Hypertension And Angina Pectoris? MediHelp. March 2020. Available at: https://medihelp.life/did-you-know-that-viagra-was-originally-developed-for-the-treatment-of-hypertension-and-angina-pectoris/ [Accessed 16 February 2021].

Four Challenges to Successful Rare Disease Drug Development. ICON. November 2019.

Liu, R., Wei, L., Zhang, P. A deep learning framework for drug repurposing via emulating clinical trials on real-world patient data. Nature. January 2021. Available at: https://www.nature.com/articles/s42256-020-00276-w [Accessed 12 January 2021].

Phases of Clinical Trials. Cancer Research UK. Updated February 2019. Available at: https://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/what-clinical-trials-are/phases-of-clinical-trials [Accessed 20 February 2021].