Scientists use machine learning to fast-track drug formulation development

January 10 2023

[Left to Right] Christine Allen and Alán Aspuru-Guzik from the University of Toronto are combining expertise in pharmaceutical sciences, AI and machine learning to develop new drug formulations faster. Credit: Steve Southon

Scientists at the University of Toronto have successfully tested the use of

machine learning models to guide the design of long-acting injectable drug formulations. The potential for machine learning algorithms to accelerate drug formulation could reduce the time and cost associated with drug development, making promising new medicines available faster.

The study was published today in *Nature Communications* and is one of the first to apply machine learning techniques to the design of polymeric long-acting injectable drug formulations.

The multidisciplinary research is led by Christine Allen from the University of Toronto's department of pharmaceutical sciences and Alán Aspuru-Guzik, from the departments of chemistry and computer science. Both researchers are also members of the Acceleration Consortium, a global initiative that uses artificial intelligence and automation to accelerate the discovery of materials and molecules needed for a sustainable future.

"This study takes a critical step towards data-driven drug formulation development with an emphasis on long-acting injectables," said Christine Allen, professor in pharmaceutical sciences at the Leslie Dan Faculty of Pharmacy, University of Toronto. "We've seen how machine learning has enabled incredible leap-step advances in the discovery of new molecules that have the potential to become medicines. We are now working to apply the same techniques to help us design better drug formulations and, ultimately, better medicines."

Considered one of the most promising therapeutic strategies for the treatment of chronic diseases, long-acting injectables (LAI) are a class of advanced drug delivery systems that are designed to release their cargo over extended periods of time to achieve a prolonged therapeutic effect. This approach can help patients better adhere to their medication regimen, reduce side effects, and increase efficacy when injected close

to the site of action in the body. However, achieving the optimal amount of drug release over the desired period of time requires the development and characterization of a wide array of formulation candidates through extensive and time-consuming experiments. This trial-and-error approach has created a significant bottleneck in LAI development compared to more conventional types of drug formulation.

"AI is transforming the way we do science. It helps accelerate discovery and optimization. This is a perfect example of a 'Before AI' and an 'After AI' moment and shows how drug delivery can be impacted by this multidisciplinary research," said Alán Aspuru-Guzik, professor in chemistry and computer science, University of Toronto who also holds the CIFAR Artificial Intelligence Research Chair at the Vector Institute in Toronto.

To investigate whether machine learning tools could accurately predict the rate of drug release, the research team trained and evaluated a series of eleven different models, including multiple linear regression (MLR), random forest (RF), light gradient boosting machine (lightGBM), and neural networks (NN). The data set used to train the selected panel of machine learning models was constructed from previously published studies by the authors and other research groups.

"Once we had the data set, we split it into two subsets: one used for training the models and one for testing. We then asked the models to predict the results of the test set and directly compared with previous experimental data. We found that the tree-based models, and specifically lightGBM, delivered the most accurate predictions," said Pauric Bannigan, research associate with the Allen research group at the Leslie Dan Faculty of Pharmacy, University of Toronto.

As a next step, the team worked to apply these predictions and illustrate how machine learning models might be used to inform the design of new

LAIs, the team used advanced analytical techniques to extract design criteria from the lightGBM [model](). This allowed the design of a new LAI formulation for a drug currently used to treat ovarian cancer. "Once you have a trained model, you can then work to interpret what the machine has learned and use that to develop design criteria for new systems," said Bannigan. Once prepared, the drug release rate was tested and further validated the predictions made by the lightGBM model. "Sure enough, the formulation had the slow-release rate that we were looking for. This was significant because in the past it might have taken us several iterations to get to a release profile that looked like this, with machine learning we got there in one," he said.

The results of the current study are encouraging and signal the potential for machine learning to reduce reliance on trial-and-error testing slowing the pace of development for long-acting injectables. However, the study's authors identify that the lack of available open-source data sets in pharmaceutical sciences represents a significant challenge to future progress. "When we began this project, we were surprised by the lack of data reported across numerous studies using polymeric microparticles," said Allen. "This meant the studies and the work that went into them couldn't be leveraged to develop the machine learning models we need to propel advances in this space," said Allen. "There is a real need to create robust databases in pharmaceutical sciences that are [open access]() and available for all so that we can work together to advance the field," she said.

To promote the move toward the accessible databases needed to support the integration of machine learning into pharmaceutical sciences more broadly, Allen and the research team have made their [datasets]() and [code]() and available on the open-source platform Zenodo.

"For this study our goal was to lower the barrier of entry to applying machine learning in pharmaceutical sciences," said Bannigan. "We've

made our data sets fully available so others can hopefully build on this work. We want this to be the start of something and not the end of the story for machine learning in drug formulation."

More information: Machine learning models to accelerate the design of polymeric long-acting injectables, *Nature Communications* (2023). DOI: 10.1038/s41467-022-35343-w

Provided by University of Toronto

Citation: Scientists use machine learning to fast-track drug formulation development (2023, January 10) retrieved 25 April 2024 from https://phys.org/news/2023-01-scientists-machine-fast-track-drug.html