

# Choosing a simpler path to drug discovery

March 29 2017

---



A new, adaptive decision tree method lights a path toward reduced development costs. Credit: Kyoto University

Drug discovery is in essence the designing of compounds to interact with disease-related proteins. And in many recent development efforts, this process increasingly relies on "big data" and complex "deep learning", requiring the harnessing of supercomputing power. But what if this could be done much more simply, requiring less time and expense?

Now a team of scientists has done just that, developing a method using simple models and small data sets—but still achieving a high degree of predictive ability. The researchers from Kyoto University, MIT, and ETH Zurich reported their findings 6 March in the journal *Future Medicinal Chemistry*.

The study demonstrates that large amounts of data generated by testing compound activity on [protein](#) groups—known for roles in cancer and other physiological processes—could be reduced to a small fraction of the total, which could still accurately explain the full set. The subset required was less than a quarter in most cases, and in some, even less than 10%.

The authors examined 13 aspects of the new method to test its usefulness.

"We tried to intentionally break our system in multiple ways. Not only did it show resilience, but many of the analyses yielded views that supported each other," says corresponding author JB Brown of Kyoto University. "After the analyses and repeated testing for reproducibility, it became apparent to us that this could become a platform for molecule design."

The authors began with publicly available compound and protein activity data, and taught a computer program how to make decisions based on available information by using a collection of 'decision trees'. Hospital doctors, for example, use decision trees to arrive at diagnoses based on patient answers to general questions.

Brown and his team gave the program some basic experience, and then showed that it could be sufficiently predictive when working on additional cases.

"There is nothing wrong with acquiring huge amounts of data and having it available as a reference. However, the extra data's utility in predicting the relationship between drugs and proteins is questionable," explains Brown, emphasizing that the new method could lead to a reduction in drug development costs.

"Drug discovery can fall into a trap of trying tens or hundreds of thousands of compounds against proteins, with 1% or less success rates," continues Brown, emphasizing that the new technique can reduce the number of initial tests to a few thousand, from which point scientists can check just the most promising ones.

"Not only are the financial implications large, but decision trees let us ask and understand the key science question: why?," says Brown. The team is now evaluating the technology in practical applications for pharmaceutical, medical, and agricultural research.

**More information:** Daniel Reker et al, Active learning for computational chemogenomics, *Future Medicinal Chemistry* (2017).  
[DOI: 10.4155/fmc-2016-0197](https://doi.org/10.4155/fmc-2016-0197)

Provided by Kyoto University

Citation: Choosing a simpler path to drug discovery (2017, March 29) retrieved 23 April 2024 from <https://medicalxpress.com/news/2017-03-simpler-path-drug-discovery.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--