

This article appeared in Drug Discov. Today 20(9) pp. 1093-1103 (2015) and explores the benefits of a more intuitive and flexible approach to viewing and interacting with drug discovery data. We illustrate how this can help to quickly identify high quality compounds and strategies for further compound optimisation.



Abstract

Drug discovery scientists often consider compounds and data in terms of groups, such as chemical series or clusters, and relationships, representing similarity or structural transformations, which help to navigate the complex process of compound selection and optimisation. This is often supported by cheminformatics algorithms that analyse complex compound data and extract relevant patterns, for example clustering and matched molecular pair analysis. However, the software that supports drug discovery chemistry almost always presents these data as spreadsheets or form views; essentially long lists that make it hard to find relevant patterns or even conveniently compare related compounds. In this paper we review methods that are commonly used to extract information from chemistry data and the ways in which these data are typically viewed. We then introduce a new framework that breaks free from the restrictions of chemical spreadsheets to work with drug discovery data in the way that scientists think about them. We also illustrate how this approach can be used to view and interact with the output of algorithms to quickly and intuitively identify key structure-activity relationships with which to guide further optimisation.

You can download a copy of this article as a [PDF](#).