Accepted Manuscript

Title: Can we accelerate medicinal chemistry by augmenting the chemist with Big Data and artificial intelligence?

Authors: Edward J. Griffen, Alexander G. Dossetter, Andrew G. Leach, Shane Montague

PII: S1359-6446(17)30578-0
DOI: https://doi.org/10.1016/j.drudis.2018.03.011
Reference: DRUDIS 2210

To appear in:

Please cite this article as: Griffen, Edward J., Dossetter, Alexander G., Leach, Andrew G., Montague, Shane, Can we accelerate medicinal chemistry by augmenting the chemist with Big Data and artificial intelligence?.Drug Discovery Today https://doi.org/10.1016/j.drudis.2018.03.011

Can we accelerate medicinal chemistry by augmenting the chemist with Big Data and artificial intelligence?

Edward J. Griffen[1], Alexander G. Dossetter[1], Andrew G. Leach[2], and Shane Montague[1]

[1]MedChemica Ltd, Biohub, Alderley Park, Macclesfield, SK10 4TG, UK
[2]Liverpool John Moores University, School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK
Corresponding author: Griffen, E.J. (ed.griffen@medchemica.com)
Keywords: Medicinal chemistry; lead optimization ; AI; big data; MMPA.
Teaser: AI comes to lead optimization: medicinal chemistry in all disease areas can be accelerated by exploiting our pre-competitive knowledge in an unbiased way.

Highlights:

- Systematizing medicinal chemistry knowledge can make it evidence based and identify valuable exceptions to established medicinal chemistry dogma
- There is a huge amount of unique medicinal chemistry knowledge
- Evidence based medicinal chemistry can counter individual and team biases
- IP sensitive medicinal chemistry knowledge can be exchanged securely by using context dependent matched molecular pair analysis
- The corporate ecosystem of drug discovery is the last barrier to knowledge sharing

It is both the best of times and the worst of times to be a medicinal chemist. Massive amounts of data combined with machine-learning and/or artificial intelligence (AI) tools to analyze it can increase our capabilities. However, drug discovery faces severe economic pressure and a high level of societal need set against challenging targets. Here, we show how improving medicinal chemistry by better curating and exchanging knowledge can contribute to improving drug hunting in all disease areas. Although securing intellectual property (IP) is a critical task for medicinal chemists, it impedes the sharing of generic medicinal chemistry knowledge. Recent developments enable the sharing of knowledge both within and between organizations while securing IP. We also explore the effects of the structure of the corporate ecosystem within drug discovery on knowledge sharing.

Introduction

Here, we how it is possible to accelerate drug discovery by analyzing, systematizing, and sharing medicinal chemistry knowledge in an unbiased and coherent way. This is at the core of AI approaches based on supervised learning. We review the well-known decrease in productivity of drug discovery and highlight the role of medicinal chemistry in addressing this issue. We discuss how the human aspects of working within a drug discovery team impact the practice of medicinal chemistry and how using more evidence-driven approaches can counter-balance natural human cognitive biases. The central part of this article discusses how the application of modern approaches to systematizing knowledge in an unbiased way can both extract new knowledge and circumvent the confidentiality issues created by the need to generate (IP. Finally, we discuss the corporate challenges and benefits to global drug discovery in sharing medicinal chemistry knowledge broadly between large pharmaceutical companies and more widely with the academic, not-for-profit, and biotech sectors.

The central challenge for chemists in drug discovery

Drug discovery is facing severe economic stress against a background of increasing societal need. The output of global drug discovery has held surprisingly constant, with a median of 16 new molecular entities (NMEs) launched per year between 1950 and 2014. Taking a straightforward definition of productivity in drug discovery as the number of new chemical entities (NCEs) brought to market divided by cost, longitudinal analysis shows an average annual increase in cost of 8% per annum, christened 'Erooms Law' [1]. This is a drop in productivity that few budgets in any industry can tolerate. Using 2016 prices: in 1950, US$1 billion would deliver over 30 drugs to market, whereas, today, it would