Open Science in the Pharmaceutical Industry: Analysis and Recommendations

Jared Ratto
University of Exeter
Department of Computer Science

Abstract

The problems of expensive drug discovery costs, high drug prices, and public distrust are well documented in the pharmaceutical industry. Open science, or the practice of openly disclosing research methods and data, has been proposed as a solution for these ills. M4K Pharma is studied in this paper as an example of a company that has incorporated open science principles into its business model. The impact of the potential growth of similar open science businesses is explored and recommendations are presented for the role of open science in the pharmaceutical industry.

Introduction

Pharmaceutical companies in developed countries have made remarkable achievements in developing breakthrough, therapeutic drugs. However, their drug discovery process has significant flaws. The innovative drugs produced by this process are generally very expensive, it is not uncommon for these to be priced at thousands of dollars per treatment plan (UCL Institute for Innovation and Public Purpose, 2018: 16-17). The drug discovery process is also fraught with inefficiencies and is marked by a trend in which the rate of new drugs per dollar of research and development (R&D) spent has steadily decreased, a phenomenon known as Eroom' law (Scannell et al., 2012). Each of these flaws at least partially stem from the industry's practice of withholding various proprietary data that could otherwise be opened to assist other organizations in drug discovery.

In this paper, I will explain how these closed data practices have led to these problems and will argue that the solution lies in making more of the drug discovery data available to the public and the research community. To do this, I will first elaborate on the closed data practices that are common in the pharmaceutical industry and I will introduce open science as an alternative to these. I will then introduce the company M4K Pharma as an example of how open science can be implemented in in the industry and address some of the limitations of the business model. I will then recommend how the industry can adopt open science and thereby solve the challenges of efficiency and affordability.

In this paper, I will use the term pharma to refer to pharmaceutical companies as a whole and as shorthand for the word pharmaceutical. This convention is also followed in academic papers on the subject.

Pharmaceutical Data Governance

Closed Data Governance

The practice of pharma to keep data proprietary can be more formally referred to as closed data governance. The term data governance refers to the policies and practices regarding how data can be stored, accessed, and shared. Under closed data governance, data are treated as private, commoditized goods that are obtained via investment and whose access is generally restricted from outsiders. If data are shared or moved outside of the company, it is often done so for a price or other compensation. The data are viewed as valuable, intellectual property assets that give the company a form of competitive advantage (Beaulieu and Leonelli, 2021: 116-122).

In a way it is understandable for pharma to treat its data as guarded assets. The traditional process of discovering and developing drugs requires large financial investments, the Tufts organization estimated that the cost of bringing a new therapeutic to market using a newly approved molecule is $2.6B (DiMasi, Grabowski and Hansen, 2016). Furthermore, all drug R&D projects involve a great deal of risk and nearly 90% of candidate drugs that enter Phase I clinical trials fail (Bountra, Lee and Lezaun, 2017: 7). Thus, for-profit businesses are incentivized to protect their intellectual property and reap as much benefit as possible from it.

It is important to note that not all of pharma data is strictly considered to be under closed data governance. For example, the summaries of all clinical trials conducted in the United States are required to uploaded to clinicaltrials.gov (The Royal Society, 2012: 43). Additionally, pharma companies are major contributors of biomedical research. The research is published in medical journals, which then are accessible at a price according to the journals' paywalls (OpenPharma, no date). However, there are several other data sources in the drug discovery process that are shared less openly, such as more detailed clinical trial datasets, pre-clinical trial data, post-marketing data (Joana Osório, personal communication), as well as the processes behind designing the drug (UCL Institute for Innovation and Public Purpose, 2018: 16). More detailed clinical trial datasets are a particularly significant data source because they have the potential to reveal deeper insights than those conveyed in the summaries alone and are less susceptible to bias. Selective bias to only include results that lead to greater likelihood of publication is still a concern of regulators in the pharmaceutical industry (The Royal Society, 2012: 43). A study in 2016 showed that unpublished documents were more likely to report adverse effects of treatments than published clinical trial results (Golder et al., 2016). The closed nature of these several datasets therefore leads to the conclusion that pharma companies still overall operate within closed data governance.

Multiple works have referred to the role of patents in pharma's secrecy with data (Balasegaram et al., 2017: 2; Morgan, Roberts and Edwards, 2018: 3). While these cited criticisms do not explicitly detail why patents lead to data secrecy, my reasoning is that patents encourage secrecy both prior to and after patent approval. To receive a patent in the United States, one must establish that the invention was never publicly disclosed prior to applying for it (Walker, no date). After the patent is approved, the inventor is granted monopoly power over the invention and has little incentive to share information, perhaps to discourage competitors from creating different products that address the same need.

The prevalence of closed data practices encouraged by patents has contributed to various problems in the drug discovery process. This secrecy leads to inefficiency in R&D in the industry as a whole and can often lead to duplicated work (Morgan, Roberts and Edwards, 2018: 3; UCL Institute for Innovation and Public Purpose, 2018: 16). Clearly, research organizations cannot learn from the mistakes and processes of similar studies if the results are closed off. This siloed nature leads companies' research to be more expensive than it otherwise could be if more knowledge was shared. This explanation appears to be verified by the steady increase of inflation-adjusted costs per R&D project (Scannell et al., 2012). The resulting high cost of research, along with the aforementioned high risk of projects,

incentivizes corporations to charge high prices for innovative drugs to ensure profits and returns for shareholders. Furthermore, patents enable companies to charge high prices without threat of competition. These prices have implications for the sustainability of healthcare systems because they are priced at the upper limits of what institutions can bear (UCL Institute for Innovation and Public Purpose, 2018: 7). Finally, another symptom of closed data practices is lack of public trust, and this is evidenced by a study revealing that more than two-thirds of the public do not trust the research produced by the pharmaceutical industry (Ipsos Mori, 2016; Phillips and Williams, 2023).

Open Data Governance

A radically different approach to data is presented in open data governance. This practice, known as open science when applied in research fields, treats data as public goods that can be freely accessed and shared by any who would benefit from the data. An example of how this principle has been applied in science is the knowledge commons, which is a designation given to scientific data sets that are labelled as key information for the public good and is therefore freely accessible to all (Beaulieu and Leonelli, 2021: 122-124). Proponents of open science argue that its practice maximizes the public benefit from scientific research and reduces transactional costs of collaboration. It enables self-correction of research processes and permits the underlying logic to be scrutinized, thereby supporting or invalidating what is claimed (International Council for Science et al., 2015: 3). By being more transparent, it also strengthens the dialogue between the public and the research community, sometimes even summarizing its findings in plain language for the lay reader (International Council for Science et al., 2015: 3; Joana Osório, personal communication). In order for these benefits to be realized, there must be sufficient infrastructure, such as online repositories, for knowledge to be shared efficiently and there must also be sufficient contributions from community members so that value is added.

Despite the benefits of open science, there are legitimate concerns surrounding its implementation. Even among advocates for open data, there is consensus that there are limits on which data should be shared. For example, openly publishing patient-identifying data of research studies would be inappropriate because it violates the patients' rights to privacy. Simply removing the names of patients would not be sufficient, because such datasets can often still be used to identify individual patients, especially those with rare conditions, when combined with other data sources (Joana Osório, personal communication). There is also reasonable concern that open data governance can reduce incentives for innovators, such as those offered by patents, and thereby lead to less innovation. Various solutions have been debated to address this issue and some of them involve rewards from regulatory bodies that still allow the information to be freely shared (Morgan, Roberts and Edwards, 2018: 8).

M4K Pharma

Case Study

M4K Pharma, founded in 2017, is an example of a company that incorporates open science principles in its drug discovery model. It is designing a therapeutic for treating diffuse intrinsic pontine glioma (DIPG), a rare and fatal pediatric brain cancer that has no chemotherapies that provide benefit (M4K Pharma, no date). It plans to produce a drug to treat this condition with affordable prices that will be set in licensing agreements with manufacturers. Rare diseases and pediatric conditions are both areas that receive low amounts of R&D because they are markets where it is difficult to earn a profit. In these, the high costs

and risks of R&D often outweigh potential profits. M4K Pharma believes that their business model is well suited for this disease because open science research has the potential to cost much less than traditional pharma's R&D. M4K Pharma relies on both financial and in-kind contributions (i.e., noncash donations in the form of equipment, data, or expertise) to keep costs low. According to a whitepaper published by the company (Morgan, Roberts and Edwards, 2018), they have received generous donations from universities, research institutes, and the private sector. The company believes that its commitment to open science and affordable pricing will position it well to receive funding from public and philanthropic sources.

M4K Pharma is committed to openly sharing its findings from its preclinical and clinical trial stages. The company livestreams and uploads to YouTube their monthly team meetings, where they discuss their research progress and even include plans for their chemical compounds. When a drug candidate is developed, they will share the protocols and analysis plans of its clinical trials to elicit feedback from the open source community and will release analyzable datasets and metadata of the results. The company plans to ensure that personally-identifiable information is removed from clinical trial data before publishing.

M4K Pharma has made it clear that none of its research activity will become patented. Instead, the company will seek to protect its intellectual property by obtaining regulatory exclusivities. These are similar to patents in that they grant protection from generic competition, but they would better foster the company's open science practices. The two main forms of these exclusivities are for drugs with new chemical entities and for drugs that treat rare diseases (known as orphan drugs). The protection of these exclusivities is significantly shorter than what is granted by patents, new chemical entities exclusivities yield 5 years of protection and orphan drugs yield 7 years whereas patents last 20 years. However, the fact that the exclusivities are nearly costless to obtain, easier to gain approval for, and that the exclusivity only begins once the company starts marketing the drug makes these exclusivities compare favorably with patents in some respects (Morgan, Roberts and Edwards, 2018: 4).

Commentary: Role of Open Source

In one of the peer reviews of M4K Pharma's whitepaper, Bernard Munos poses a question: "which is role of open-source in the drug R&D ecosystem. Is it a replacement, an alternative, or a partner to pharma?" (Morgan, Roberts and Edwards, 2018: 12). My answer to this question is that it is an alternative to traditional pharma, where in an ideal scenario both comprise a large portion of drug discovery. I assert this position for multiple reasons, namely that the open science model used by M4K Pharma is not practical for drug markets outside of rare diseases and that the industry benefits from some big players that are characterized by profits and patents.

M4K Pharma presents a well-thought model for open source drug discovery that can work well for rare diseases, but it is doubtful that it could be applied as successfully in other markets under the current regulatory conditions. The 7 years of market exclusivity granted for orphan drugs is substantial, especially when combined with the 5 years granted for new chemical entities. However, treatments for non-rare conditions that rely only on new chemical entity would only be entitled 5 years. In my view, this is not enough incentive for companies to forego the greater profits and protection offered by patents in favor of open science ideals. Patents would furthermore be a superior alternative in these larger markets because the new chemical entity exclusivity is still available to apply for as the patent expires. In fact, it is a common practice for companies to use both patents and market exclusivities to protect their intellectual property (Morgan, Roberts and Edwards, 2018: 17).

In order for M4K Pharma's model to scale effectively to larger markets, regulatory changes would be needed to create more incentives for open source drug discovery. The writers of the whitepaper even acknowledge this and suggest two additional forms of market exclusivity. The first is granted for simply uploading preclinical or clinical trial data to an open source repository, and the incentive would protect against competitive uses of the data prior to approval of the drug developer's marketing application. The second is an "open science extension" that can be granted in addition to other market exclusivities. This extension would be granted on the condition that the drug company demonstrates that it has made its preclinical and clinical data publicly available, that it has not filed for patents, and that it has agreed to set affordable price ceilings for its drug (Morgan, Roberts and Edwards, 2018: 8).

However, even if it were practical in all markets, I would still assert that not all companies in the pharmaceutical industry should follow the open science model, despite its many virtues. The main reason for this is that the brightest minds are needed to address today's health challenges, and some of them will be motivated more by the greater potential of profits offered by patents to make their contributions. It may not necessarily be harmful, and in fact may be beneficial, to drug discovery if there are a handful of big players that lead the industry with state-of-the-art resources. After all, if all pharma companies are open source, where will in-kind contributions from the private sector come from? It is likely that the existence of some for-profit companies would lead to higher quality contributions given to open source companies like M4K Pharma.

Recommendations

While open science companies should not replace all of pharma, I assert that it is beneficial if such companies grow in popularity so that they form a sizable portion of the industry. Perhaps the majority of companies in the rare disease markets would follow M4K Pharma's model in this scenario, and it could be further scaled with help from regulation. In this ideal setting where both open source and traditional pharma are both prominent, the two business models may influence each other and compete in beneficial ways for consumers. As open science pharma become more established, the public may question why traditional pharma's prices are not as low as those charged by open science. Traditional pharma companies will receive an even greater incentive to be ethical and accountable because of the imminent comparison with open science.

Meanwhile, for-profit pharma companies should integrate more open data governance practices without necessarily becoming open source companies. Many players in the pharmaceutical industry recognize that public distrust is a problem and that transparency is the solution; however, they want data sharing to be done responsibly while respecting patients' privacy (Joana Osório, personal communication). A step pharma companies can take is by uploading their clinical trial data to open science repositories such as Vivli and YODA. These secure platforms allow for anonymized datasets to be accessed by the international research community and allow for users to request datasets from data contributors (Center for Outcomes Research and Evaluation, no date; Vivli, no date). Vivli's webpage on "Why Share Your Clinical Research Data" states that there is growing incentive for companies to follow these practices and that funders and medical journals increasingly require research data to be openly shared (no date: Paragraphs 1-2). The page also mentions that "sharing detailed clinical research data is associated with increased citation rate" (no date: Paragraph 4).

A tangential recommendation is that most funding from public sources should be conditional on affordable pricing commitments made by recipients. Many governments recognize the threat that high prices have on healthcare systems (U.S. Food and Drug

Administration, no date; The White House, 2023) and should align their funding with these values. This condition may in turn encourage firms to use open science principles to keep costs low so that lower prices will be more feasible.

Conclusion

In this paper, I have explored some of the complexities surrounding pharma's closed data governance and how they lead to the ills of traditional drug discovery. The proposed recommendations of more open science pharma companies operating, traditional pharma adopting open science principles, and public funding requiring affordable pricing commitments will lead to better and more efficient research and to lower prices for new drugs. Researchers will be able to gain learnings more easily from relevant studies thereby reducing the costs of their research progress. The public will have greater trust in pharmaceutical companies because of the transparency of their uploaded clinical studies. The reduced cost of research caused by prevailing trends of open science principles will make lower drug prices more feasible. The growing presence of open science companies will also exert a downward pressure on prices because of their affordable pricing commitments.

References:

Balasegaram, M. et al. (2017) ‘An open source pharma roadmap’, PLOS Medicine, 14(4), p. e1002276. Available at: https://doi.org/10.1371/journal.pmed.1002276.

Beaulieu, A. and Leonelli, S. (2021) Data and Society: A Critical Introduction. Los Angeles: SAGE Publications.

Bountra, C., Lee, W.H. and Lezaun, J. (2017) A New Pharmaceutical Commons: Transforming Drug Discovery. Oxford. Available at: https://www.oxfordmartin.ox.ac.uk/downloads/academic/Transforming_Drug_Discovery.pdf.

Center for Outcomes Research and Evaluation (no date) Welcome to the YODA Project. Available at: https://yoda.yale.edu/welcome-yoda-project (Accessed: 20 May 2023).

DiMasi, J.A., Grabowski, H.G. and Hansen, R.W. (2016) ‘Innovation in the pharmaceutical industry: New estimates of R&D costs’, Journal of Health Economics, 47, pp. 20–33. Available at: https://doi.org/10.1016/j.jhealeco.2016.01.012.

Golder, S. et al. (2016) ‘Reporting of Adverse Events in Published and Unpublished Studies of Health Care Interventions: A Systematic Review’, PLOS Medicine, 13(9), p. e1002127. Available at: https://doi.org/10.1371/journal.pmed.1002127.

International Council for Science et al. (2015) Open Data in a Big Data World: An international accord. Paris. Available at: https://council.science/wp-content/uploads/2017/04/open-data-in-big-data-world_long.pdf (Accessed: 20 May 2023).

Ipsos Mori (2016) Wellcome Trust Monitor Report Wave 3. London.

M4K Pharma (no date) About M4K Pharma. Available at: https://m4kpharma.com/about/ (Accessed: 20 May 2023).

Morgan, M.R., Roberts, O.G. and Edwards, A.M. (2018) ‘Ideation and implementation of an open science drug discovery business model – M4K Pharma’, Wellcome Open Research, 3, p. 154. Available at: https://doi.org/10.12688/wellcomeopenres.14947.1.

OpenPharma (no date) Position Statement on Open Access. Available at: https://www.openpharma.blog/position-statement-on-open-access/ (Accessed: 20 May 2023).

Phillips, C. and Williams, S. (2023) Pharma knocks tech off the top spot as most trusted industry in Ipsos’ latest global report on Trust, Ipsos. Available at: https://www.ipsos.com/en-uk/pharma-knocks-tech-off-top-spot-as-most-trusted-industry-trustworthiness-monitor (Accessed: 20 May 2023).

Scannell, J.W. et al. (2012) ‘Diagnosing the decline in pharmaceutical R&D efficiency.’, Nature reviews. Drug discovery, 11(3), pp. 191–200. Available at: https://doi.org/10.1038/nrd3681.

The Royal Society (2012) Science as an open enterprise. London. Available at: https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/ (Accessed: 20 May 2023).

The White House (2023) FACT SHEET: Seniors Across the Country Are Saving Millions of Dollars in Health Care Costs Because of President Biden’s Prescription Drug Law. Available

at: https://www.whitehouse.gov/briefing-room/statements-releases/2023/03/15/fact-sheet-seniors-across-the-country-are-saving-millions-of-dollars-in-health-care-costs-because-of-president-bidens-prescription-drug-law/ (Accessed: 20 May 2023).

UCL Institute for Innovation and Public Purpose (2018) The people's prescription: Re-imagining health innovation to deliver public value. London. Available at: https://www.ucl.ac.uk/bartlett/public-purpose/sites/public-purpose/files/peoples_prescription_report_final_online.pdf (Accessed: 20 May 2023).

U.S. Food and Drug Administration (no date) Generic Drugs: Questions and Answers. Available at: https://www.fda.gov/drugs/frequently-asked-questions-popular-topics/generic-drugs-questions-answers (Accessed: 20 May 2023).

Vivli (no date a) About Vivli: Overview. Available at: https://vivli.org/about/overview/ (Accessed: 20 May 2023).

Vivli (no date b) Why share your clinical research data. Available at: https://vivli.org/why-share-your-clinical-research-data/ (Accessed: 20 May 2023).

Walker, A. (no date) Pharmaceutical Patents: an overview, Alacrita. Available at: https://www.alacrita.com/blog/pharmaceutical-patents-an-overview (Accessed: 20 May 2023).