Thesis Portfolio

A Novel Wheat-Germ-Agglutinin Barcoding Approach for Diverse Mass Cytometry Samples

(Technical Report)

The Impact of Artificial Intelligence Tools in Pharmaceutical Research

(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia – Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Brendan Cox

Spring, 2024

Department of Biomedical Engineering

Table of Contents

Sociotechnical Synthesis

Technical

Barcoding is a powerful tool enabling the batch analysis of cell populations. Barcoding, specifically in mass cytometry, labels cells originating from different samples with a unique configuration of metal isotopes such that all samples can be processed concurrently and later differentiated in data outputted by a mass cytometer. The abundant mass channels available in mass cytometry or Cytometry by Time of Flight (CyTOF) support per-sample barcoding due to a large number of differentiable mass/charge ratios. The increase in throughput and reduction in experimental variability resulting from barcoded experiments have made it a widely used tool in multiplexed cytometric studies. Most live cell barcoding approaches, however, are restricted by species, cell lineage, or fixation state. In this Capstone project we demonstrate a live cell barcoding method that mitigates those restrictions, increasing the diversity of samples that can be concurrently measured in a CyTOF experiment. Thiolated wheat germ agglutinin (tWGA) has affinity for most membrane bound cells and is stable enough to be conjugated with metal isotopes. We demonstrate the utility of tWGA conjugates as a live cell barcode before fixation and across cell lineage. Our comparator is a percentage of CyTOF events falling within defined Mahalanobis and barcode separation distance ranges called percent good debarcoded (PGD). We show that this tWGA barcode approach has a PGD of 73% with a z-score of -0.87, which we claim to be comparable to 11 other barcoding methods tested with an average PGD of 81%.

STS

Artificial intelligence (AI) is an emerging tool in drug development with the potential to fundamentally change business strategy for established pharmaceutical companies. The commercialization of a novel drug requires enormous upfront investment and the assumption of significant risk. The so-called “Valley of Death” between a preclinical drug candidate and a labeled, marketable product can be prohibitive for companies outside big pharma, where profitable portfolios can support clinical trials. The result is an industry where established companies are empowered to acquire smaller startups, profiting from innovation without being internally innovative themselves. This STS thesis investigates the future role of AI in pharmaceutical development, the changes it’s likely to bring to the industry, and the impact of those changes on society.

The technical portion of this work is a literature review detailing what technologies are available to biopharma companies, how they might be used, and their impacts on productivity. The report will describe relevant sources of data, such as Protein-Protein or Drug-Protein Interaction databases, and how they are used to train machine learning models to facilitate drug discovery. The types of neural networks frequently used in drug development software, such as generative adversarial networks, will also be explored in detail.

The rest of this work focuses on the downstream impact of AI tools, divided into three main areas of thought: monopolization, accountability, and social impact. Monopolization refers to the persistent oligopoly that characterizes the global

pharmaceutical industry, and the ability of these massive companies to continue to consolidate market share. AI, as a disruptive and transformative technology, has the potential to be a significant threat to the current distribution of power. It may enable new entrants to compete more effectively by reducing the cost of new product development, optimization of clinical trials, and reducing risk of novel R&D. Accountability will focus on the ways AI may increase transparency in the pharma industry. Generally speaking, *in silico* research— meaning research done digitally or computationally— is more easily replicated. AI also has a wide array of applications in data management and analysis for regulatory agencies such as the FDA. Finally, the section on Social Impact will characterize how the other two analyses impact the end user of a pharmaceutical product. More competition as a result of AI-driven tools may drive down costs, increase the number of available medications, or even impact the ratio of generic to branded medications.

Synthesis

Whether looking at AI-driven research tools or a new barcoding method, the rate and scale of medical research is often bottlenecked by the physical properties or requirements of the material being studied. In our Capstone project, we develop a new approach to cytometry experiments that significantly reduces labor, time, and cost. AI utilization in pharma is largely driven by the same goal: reduce time, money, and labor spent on physical experimentation in favor of computational modeling. Whether these advances are viewed as good or bad largely depends on the identity of the primary beneficiary of the technology. Increased efficiency should be pursued with an understanding of impact and stakeholder identity.