

Modeling the Effects of Advanced Automation and Process Design on Cell Line Development

By

Ryan Shofnos

B.S. Aerospace Engineering, University of Maryland, 2010

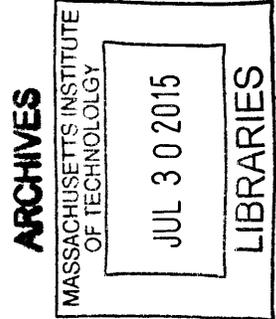
Submitted to the MIT Sloan School of Management and the Mechanical Engineering Department in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration and
Master of Science in Mechanical Engineering

In Conjunction with the Leaders for Global Operations Program at the
Massachusetts Institute of Technology

June 2015

©2015 Ryan Shofnos. All rights reserved.



The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author _____ **Signature redacted** _____
MIT Department of Mechanical Engineering, MIT Sloan School of Management
May 8, 2015

Certified by _____ **Signature redacted** _____
David E. Hardt, Thesis Reader
Ralph E. and Eloise F. Cross Professor of Mechanical Engineering

Certified by _____ **Signature redacted** _____
Retsef Levi, Thesis Supervisor
J. Spencer Standish (1945) Prof. of Management, Prof. of Operations Management

Certified by _____ **Signature redacted** _____
J. Christopher Love, Thesis Supervisor
Associate Professor of Chemical Engineering

Accepted by _____ **Signature redacted** _____
David E. Hardt, Chair, Mechanical Engineering Committee on Graduate Students
Ralph E. and Eloise F. Cross Professor of Mechanical Engineering

Accepted by _____ **Signature redacted** _____
Maura Herson, MBA Program Director
MIT Sloan School of Management

This page intentionally left blank.

Modeling the Effects of Advanced Automation and Process Design on Cell Line Development

By

Ryan Shofnos

Submitted to the MIT Sloan School of Management and the Department of Mechanical Engineering on May 8, 2015 in Partial Fulfillment of the requirements for the Degrees of Master of Business Administration and Master of Science in Mechanical Engineering

Abstract

Research and development of biologic drugs is a time- and resource-intensive process that can span several years and billions of dollars. Any improvements in the efficiency and end-to-end cycle time of this process provide value to producers in the form of reducing at-risk investment in new drug programs and improving speed to market. Cell Line Development (CLD), a major portion of the research and development lifecycle, is responsible for creating the parent cell for these new drug programs. The biotechnology industry has made great gains in CLD technologies and procedures, though many fields continue to advance and can further contribute to improved operational efficiency.

This thesis proposes a methodology for evaluating CLD systems, characterizing alternative processes and technologies, and determining the ideal investments that can maximize system efficiency and processing speed. Approaches that are currently available in the industry are reviewed and used as model inputs to determine realistic short-term gains. Furthermore, nascent technologies that may reach industrial applicability are considered for an additional potential system design. Pfizer's CLD system is used as a case study, in which it is shown that total system utilization and cycle time can be improved by 29.6% and 8.8%, respectively, through the use of currently available technologies and procedures. The costs and risks of the new approaches are reviewed and found to be significantly low when compared with these gains. As technologies continue to develop in the future, they may further improve CLD system performance. However, the majority of gains are achieved by applying currently available approaches.

Thesis Supervisor: Retsef Levi

J. Spencer Standish (1945) Prof. of Management, Prof. of Operations Management

Thesis Supervisor: J. Christopher Love

Associate Professor of Chemical Engineering

Thesis Reader: David E. Hardt

Ralph E. and Eloise F. Cross Professor of Mechanical Engineering

This page intentionally left blank.

The author wishes to acknowledge the Leaders for
Global Operations Program for its support of this work.

This page intentionally left blank.

Acknowledgements

First and foremost, I would like to thank my faculty advisors, Retsef Levi and Christopher Love, for their guidance throughout this effort. Professor Love's expertise in this topic was invaluable and helped me hit the ground running. To Professor Levi, I would like to express my appreciation for his knowledge and drive that were critical to the team's success.

I also owe much of this work to the amazing team at Pfizer that brought me in and taught me all about the biopharmaceutical world. I am incredibly thankful for the support of Alex Chin, Cecilia Cooley, Linda Francullo, Pamela Pegman, Brendan Smith, and the entire CLD team for their support (and unwavering stamina through all of my questions).

Finally, I would like to thank my family for supporting me through this research effort and the past two years in general. A special thank you to my wife Kate for continuing to inspire and motivate me, and to Mom, Dad, Matt, and all of the grandparents, aunts, uncles, and cousins. You have all supported and believed in me throughout my education, for which I will always be grateful.

This page intentionally left blank.

Table of Contents

Abstract.....	3
Acknowledgements.....	7
Table of Contents.....	9
Glossary & Acronym List	11
1. Introduction.....	13
1.1. Project Motivation	13
1.2. Project Goals	13
1.3. Problem Definition.....	14
1.4. Thesis Overview	14
2. Background and Context.....	16
2.1. Biotechnology Industry Overview.....	21
2.2. Biologic Drug Development	25
2.2.1. Cell Line Development (CLD)	25
2.2.2. Clinical Trials, Patents, and Exclusivity	29
2.2.3. Development Risk.....	30
2.3. Pfizer, Inc.....	33
3. Literature Review	34
3.1. Future Methods	34
3.2. Modeling Approaches	35
4. Modeling Current and Potential States of Development Operations	36
4.1. Model Inputs.....	36
4.2. Tradeoff Considerations	43

4.3.	Results	45
4.3.1.	Current State and Emerging Bottlenecks	45
4.3.2.	Potential Improvement via Available Inputs	49
4.3.3.	Potential Future States	52
5.	Performance and Risks of Potential Approaches	55
5.1.	Review of Underlying Technology	55
5.1.1.	Cloning	55
5.1.2.	Screening	58
5.1.3.	Scale-Up	60
5.1.4.	Clone Screen	60
5.1.5.	Outsourcing Team Responsibilities	61
5.2.	Risk Management	62
6.	Conclusions and Recommendations	65
6.1.	Conclusions for CLD at Pfizer	65
6.2.	Recommendations for Future Initiatives	66
	References	67

Glossary & Acronym List

Advanced Microscale Bioreactor (AMBR): An automated platform that mimics a full-scale production bioreactor environment for a group of cell cultures on a significantly smaller scale.

Analytical Research & Development (ARD): A research and development group that performs certain analyses of cellular populations and other biological materials.

Biopharmaceutical: A human health technology produced by harnessing the metabolic processes of living cells that, when ingested, addresses an indication of a disease or disorder.

Biosimilar: The biotechnology industry's analogue to pharmaceutical generic drugs; a biopharmaceutical produced using procedures that are similar to the brand-name producer's but may not be identical as a final product.

Biotechnology: An industry that creates its major products by harnessing the metabolic processes of living cells to generate material.

Cell Line Development (CLD): The process of inserting genetic sequences into host cells, generating and testing several of the resulting clones, and developing the master cell bank.

Clone Select Imager (CSI): An automated platform used during cell screening to assess a population's growth rate and number of cells.

Drug Substance (DS): The material produced by host cells that, when combined with other materials (e.g., buffers and solvents), becomes the complete biopharmaceutical drug.

Expanded Extensive Clonal Sequencing (EECS): A series of assays performed on the DNA and RNA of a population of cells to verify that these sequences have not significantly mutated over time.

Master Cell Bank (MCB): The final product of cell line development that will be used as the parent population of all subsequent drug substance production.

Monoclonal Antibody (mAb): A protein that is expressed by a unique parent cell line due to the transfection of the gene(s) of interest into the host cell's genomic structure.

Monoclonality: A property of a population of cells indicating that it has grown over time from one single parent cell. Monoclonal populations are more stable than the alternative and are required by many regulatory agencies for biopharmaceutical drug programs.

Orphan Drug: A medication that has been developed to treat a very rare disease that affects very small patient populations.

Pre-MCB (pMCB): A group of frozen cell line clones that, once a final clone has been identified as the best candidate, will be thawed and used to generate the master cell bank.

Sequence Variant Analysis (SVA): A series of assays on the protein produced by a population of cells. This is usually performed after a population has aged many generations to verify that the individual amino acid sequences and overall structure have not significantly mutated over time.

Size Exclusion Chromatography (SEC): One of the sub-tasks that comprise SVA, in which protein is quickly analyzed for any noteworthy changes in mass. This allows for fast characterization of molecules before proceeding into more time-consuming, extensive analyses.

Titer: The production rate of a cell or population of cells, traditionally measured in grams of drug substance per liter.

1. Introduction

The biopharmaceutical industry is generally comprised of two segments: (i) small molecule pharmaceutical drugs and (ii) large molecule biologics. Small molecules are developed and produced by chemically defined processes that synthesize the pieces of the drug together, and are typically delivered to patients orally. Biologics, on the other hand, are more complex to develop because the main drug substance is produced by harnessing the metabolisms of mammalian, bacterial, fungal, or insect cells, and are usually administered as an intravenous or subcutaneous liquid. This thesis focuses on Cell Line Development (CLD), one of the major research and development (R&D) processes for creating new biologic drugs and evaluates an operational modeling strategy for improving its efficiency and speed.

1.1. Project Motivation

The development and commercialization of biologic-based human therapies is a significant engineering and financial investment that can span close to 10 years and require investments greater than \$1 billion. A critical phase of the development lifecycle, known as CLD, is the creation and selection of a clonal cell line that will ultimately manufacture the biologic drug substance for clinical trials and commercial distribution. Reaching clinical trials quickly enables early assessment of the efficacy of the potential new drug. Therefore, generating the material needed for clinical studies is a critical step in the R&D process.

Pfizer, Inc. is a biopharmaceutical company that develops, produces, markets, and distributes both small molecule and biologic drugs worldwide. Small molecules make up the majority of its portfolio, but both Pfizer's production of biologics and the biotechnology industry as a whole are growing rapidly, as discussed in further detail below in Section 2. Therefore, Pfizer's R&D processes for biologics are becoming more vital.

1.2. Project Goals

This thesis investigates the capability of an operational modeling approach to identify upgrades to a CLD system that increase its capacity and throughput without increasing the requisite staffing. These improvements will allow CLD to develop a greater number of new drug

programs. Furthermore, due to the fact that CLD serves as the bridge between molecular discovery and clinical drug supply, they will reduce the end-to-end timeline that precedes clinical trials. We develop a model that evaluates potential methods of re-designing and automating CLD processes to realize these potential benefits at sufficiently low cost.

1.3. Problem Definition

To identify the upgrades to the CLD system that will enable these performance goals, we propose an operational modeling strategy that evaluates various approaches for each step in the process. This model can be used to identify specific system upgrades that improve overall performance and inform investment decisions accordingly. If management uses such a model to assess its R&D operations, then they can make informed updates that improve system efficiency without increasing headcount, enabling improved overall R&D performance.

1.4. Thesis Overview

Background and Context

This chapter provides the context in which this effort applies with an overview of the pharmaceutical and biotechnology industries. It highlights market factors that influence CLD operations, establishes a high level overview of the CLD process, describes subsequent and parallel activities that CLD enables and interacts with, and outlines the risks associated with drug development. This information highlights the competitive landscape in which drug producers operate and the resulting incentives for high-performing R&D operations.

Literature Review

This chapter reviews some potential future technologies that are currently being developed and may be able to further augment CLD systems. It also proposes potential modeling approaches that can be used to evaluate these systems. Monte Carlo simulation in Microsoft Excel is chosen as the most applicable method for the case study at Pfizer. Additional modeling approaches are discussed that can be used to obtain further insights in any future cases.

Modeling Current and Potential States of Development Operations

This chapter discusses the structure of the operational model, its main inputs, and the key objectives that the updated CLD system aims to achieve. We discuss the tradeoffs between the most pertinent inputs and outputs and provide results in the form of two updated system designs: one that is achievable using technologies that are currently available and another that augments the system with a potential future technology. The former updated design is predicted to improve total system utilization and cycle time by 29.6% and 8.8%, respectively. Furthermore, this new design reduces the variation of these metrics by approximately 17% and 14%. Adopting technologies that are currently in development can further improve utilization and cycle time by 0.7% and 1.6%, respectively. However, this future case depends on the assumptions that are made about nascent technologies that may come to significantly over- or under-perform relative to current predictions.

Performance and Risks of Potential Approaches

We discuss the technologies and operational characteristics of each of the approaches that the model uses as input values. This includes approaches that are currently available and can be applied in the near-term. We also highlight some of the risks associated with each approach and identify risk management frameworks that can be used to address them. Based on experimental validation of these approaches and interviews with pertinent stakeholders, we find that an updated CLD system can be designed and achieve the gains mentioned above while incurring negligible costs and operational risks.

Conclusions and Recommendations

This final chapter reviews the performance of the proposed operational model and updated CLD system. The current recommendation for Pfizer's case study is to upgrade the CLD system with the reviewed technologies that are currently available to achieve the gains stated above. Additionally, we highlight areas for further development and application of this modeling strategy, including adopting similar efforts throughout the R&D organization and continuing to evaluate the applicability of new technologies as they are developed.

2. Background and Context

In this section, we outline the current state of the biopharmaceutical industry. Market information is provided to frame the competitive landscape that Pfizer and its competitors operate in, the challenges they face, and the motivation for this research effort. We move on to describe the R&D operations that are required for drug development, from CLD through clinical trials and commercialization. Finally, we discuss the business risk associated with such drug development programs and its relation to R&D capacity and throughput.

The biopharmaceutical industry is comprised of two major components: (i) the pharmaceutical industry and (ii) the biotechnology industry. The pharmaceutical industry is the traditional medical drug development mechanism, creating small-molecule medicines by synthetically arranging molecules into a medically advantageous arrangement. On the other hand, the biotechnology industry produces several products, including agricultural compounds, industrial chemicals, and, most notably for this topic, human health technologies. These are large-molecule medications that are created by harnessing and altering the metabolic mechanisms of living organisms to produce drug substances. This biopharmaceutical industry is highly competitive, as can be seen below in Table 1, which shows the sales and relative market share of each of the top biopharmaceutical companies.

Table 1: Global sales and growth of the top 20 biopharmaceutical companies as of December 2013 (IMS Health 2013). *Sales and rank figures are in US\$ with quarterly exchange rates. Growth is in constant \$ to normalize for exchange rate fluctuations. Sales cover direct and indirect pharmaceutical channel wholesalers and manufacturers. The figures include prescription and certain over the counter data and represent manufacturer prices.*

	2013 RANK	2013 SALES (US\$ Mn)	2013 GROWTH (LC\$ %)	2012 SALES (US\$ Mn)
GLOBAL MARKET		874,611	4.5	857,710
NOVARTIS	1	50,576	1.9	50,521
PFIZER	2	44,330	-2.6	46,707
SANOFI	3	38,181	1.4	38,531
MERCK & CO	4	36,350	-7	39,891
ROCHE	5	36,146	5.3	34,958
GLAXOSMITHKLINE	6	32,544	1.5	32,736

	2013 RANK	2013 SALES (US\$ Mn)	2013 GROWTH (LC\$ %)	2012 SALES (US\$ Mn)
JOHNSON & JOHNSON	7	30,784	12.2	27,717
ASTRAZENECA	8	30,257	-2.9	31,704
TEVA	9	24,258	-1.8	24,762
LILLY	10	23,045	8.4	21,583
AMGEN	11	18,621	8.7	17,103
ABBVIE	12	18,150	2	17,881
BOEHRINGER INGELHEIM	13	17,375	5.7	16,889
BAYER	14	17,276	8.3	16,431
NOVO NORDISK	15	14,300	15	12,576
TAKEDA	16	13,399	-9.5	15,909
ACTAVIS	17	12,742	2.9	12,375
MYLAN	18	11,087	7.4	10,325
BRISTOL-MYERS SQB.	19	11,023	-12.9	12,756
GILEAD SCIENCES	20	11,011	14.8	9,540

The pharmaceutical industry as we know it today traces its roots to the early 20th century, though pharmaceuticals have been used in some fashion throughout recorded history. It includes organizations that develop and manufacture small-molecule compounds through chemical synthesis. These drug products are then combined with solvents or other materials to facilitate introduction into the human body. Small molecules are the traditional medical substances and represent approximately 90% of the drugs currently in the market (Bayer Healthcare 2014). One major advantage of small-molecule pharmaceuticals over large-molecule biologics is that they are often easily processed into orally-accepted formats, rather than requiring intravenous or subcutaneous injection. Another advantage is the relatively low cost and high reliability of development and production. They are available in the market as either patented or generic drugs. The largest source of revenue is available to the developer of a drug before the patent expires, after which point other organizations may create generic versions at much lower cost. This allows them to sell the drug at a much lower price. Therefore, most of the value of a new drug program is realized during the time period after market introduction and before revenue losses after patent expiration (Gibson 2013). For example, Pfizer's

exclusivity rights expired for Lipitor® in 2011 and Lyrica® in 2013, causing a drop in revenues in the following year of \$5.7 billion (59%) and \$105 million (51%), respectively (Pfizer, Inc. 2014).

The biotechnology industry is much younger than the pharmaceutical industry, with its beginnings considered to be in the 1970s, when Genentech was founded and began commercializing the use of recombinant DNA (“Biotechnology” 2014). Medicines produced via biotechnology are created using this recombinant DNA (Karcher 2013), which refers to artificially altering the DNA of an organism. These drugs are known as large-molecules and biologics, among other names. Biologic drugs are created using a drug substance that is produced and secreted by living organisms, which have been transfected with recombinant DNA. This drug substance may fall into one of three of the following categories:

1. Monoclonal antibody (mAb): a protein that is expressed by a unique parent cell line due to the transfection of the gene(s) of interest into the host cell’s genomic structure. This is the most common type of biologic drug (Ossipow and Fischer 2014).
2. Fusion protein: a combination of parts of several naturally-occurring proteins, resulting from transfecting multiple gene sequences that code for different molecules.
3. Antibody-drug conjugate (ADC): a targeted therapy consisting of a complete or partial mAb combined with a synthetically-produced small-molecule drug. This drug is often cytotoxic, enabling targeted and potent cancer treatments.

After the drug substance is produced and collected, it is purified and formulated into a final product that is usually introduced into the human body via injection. Biologic drugs face a similar patent structure as pharmaceuticals, though the nascent market for biosimilars, the corollary of small-molecule generics, is still in its infancy. Regulations regarding biosimilar development have only recently been established and the market is still dominated by brand-name biologics.

The combined biopharmaceutical industry remained relatively small until it was accelerated by several groundbreaking discoveries, such as the antibiotic properties of penicillin and the pharmacodynamics and -kinetics of insulin. Growth in the United States closely followed that

of the domestic GDP until the 1980s, when drug discovery and development was significantly accelerated by regulatory reform. It experienced rapid growth over the following few decades, though this growth has slowed relatively in recent years, as seen below in Figure 1.

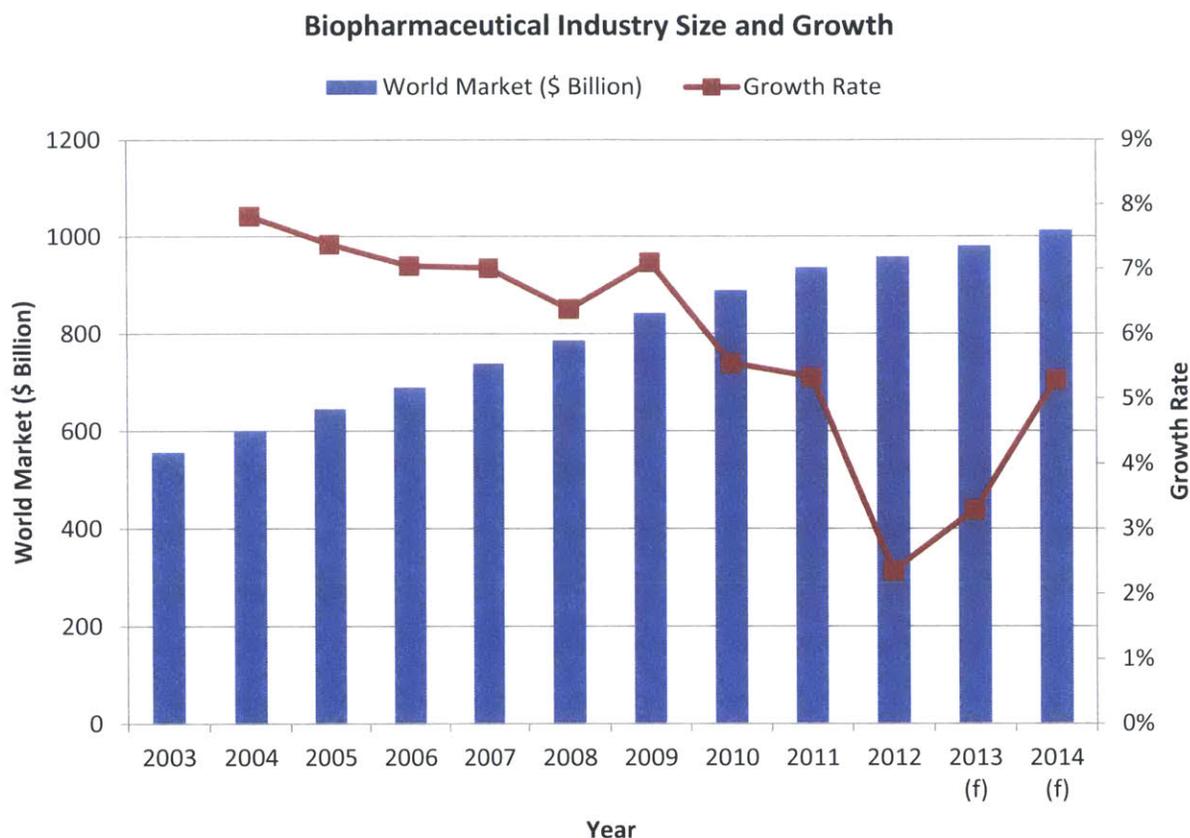
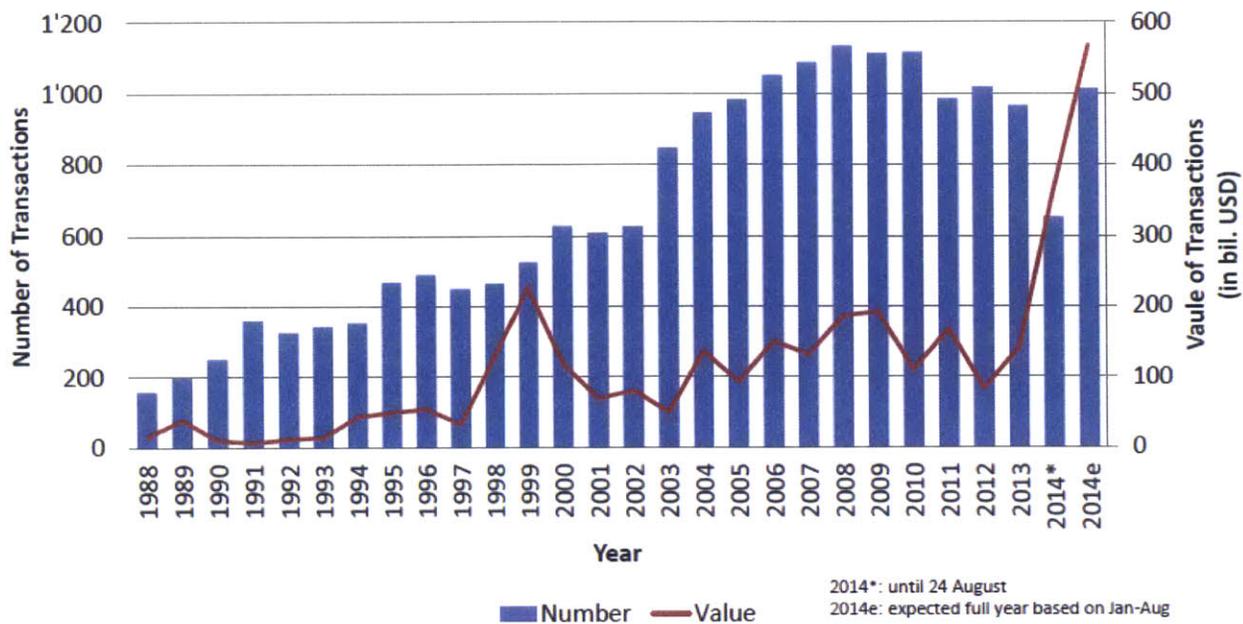


Figure 1: Global biopharmaceutical market size and growth from 2003-2014 (IMS Health 2013).
Worldwide revenue has increased from \$558 billion to just over \$1 trillion while the annual growth rate has decreased from 8% to 5%. (f) = forecasted.

The biopharmaceutical industry continues to grow but at a progressively decreasing rate. The resurgence after 2012 is due in part to the advent of biosimilar approval pathways, an aging global population, an increasing standard quality of life worldwide, and recent technological advancements (e.g., bioinformatics and synthetic biology). The general market resurgence and advent of technological advances makes it imperative for biologic drug producers to make the most efficient use of their development capacity.

Despite the close relationship between these two industries, there are significant underlying differences in market dynamics between the pharmaceutical and biotechnology industries. Biotechnology has experienced significant recent growth that far exceeds that of the broader industry. The biotechnology industry as a whole has grown 10.8% on average from 2009 through 2014, and is expected to continue to grow at an average of 9% per year through 2019 to \$444.9 billion. This is primarily due to increasing investor confidence, an aging global population, increased R&D activities, and an increasingly more educated global workforce (Phillips 2014).

On the other hand, the relatively mature small-molecule market has caused the combined biopharmaceutical industry to experience slower overall growth. It has grown at a 6.5% annual rate from 2009 through 2014 and is anticipated to grow 3.7% per year through 2019 to \$1.4 trillion. Although the small-molecule industry is also driven by the aging population, it is experiencing increasing competition from more effective and efficient biologic drugs, as well as increasing scrutiny from regulatory agencies (Turk 2014). Additionally, several patents for top-performing blockbuster drugs have recently expired, including Pfizer's Lipitor in 2011, which was the most prominent product in the United States in terms of sales (Frost & Sullivan 2012). It should be noted that despite their differences, the pharmaceutical and biotechnology industries are beginning to merge. Their initial distinction was a product of the fact that large pharmaceutical companies shunned the emerging technology in the 1980s and 1990s due to its lack of established credibility. However, these established organizations are now embracing biotechnology, as evidenced by the large amount of recent mergers and acquisitions (M&A). Such actions have steadily increased over the past couple of decades and experienced a large jump in value in 2014, as seen below in Figure 2.



Source: Thomson Financial, Institute of Mergers, Acquisitions and Alliances (IMAA) analysis

Figure 2: Total number and value of biotechnology and pharmaceutical mergers and acquisitions (Institute of Mergers, Acquisitions and Alliances (IMAA) 2014). *M&A actions peaked around 2007 though are experiencing a resurgence in 2014. M&A value experienced a large jump in 2014.*

The large amount of mergers and acquisitions can also be attributed to the patent cliff that began in 2011. The biopharmaceutical market stands to lose more than \$200 billion in revenue between 2013 and 2018 due to patent expirations, which greatly increases the attractiveness of smaller companies that are focused on new treatments for rarer conditions. By absorbing these smaller players, established biopharmaceutical can augment their portfolios without incurring the costs and risks of internal R&D (Silver 2013).

2.1. Biotechnology Industry Overview

The biotechnology industry is relatively young and fragmented, though major players are beginning to emerge. Roche remains the largest entity with 9.2% of the global biotechnology market, followed by Gilead (7.6%), Amgen (6.8%), and Merck (4.2%) (Phillips 2014). As seen above, the majority of the larger players are increasing their footprint through mergers and acquisitions (for example, Pfizer’s acquisition of Wyeth in 2009). The industry’s total revenue growth is expected to remain positive but continue decreasing, as seen below in Figure 3.

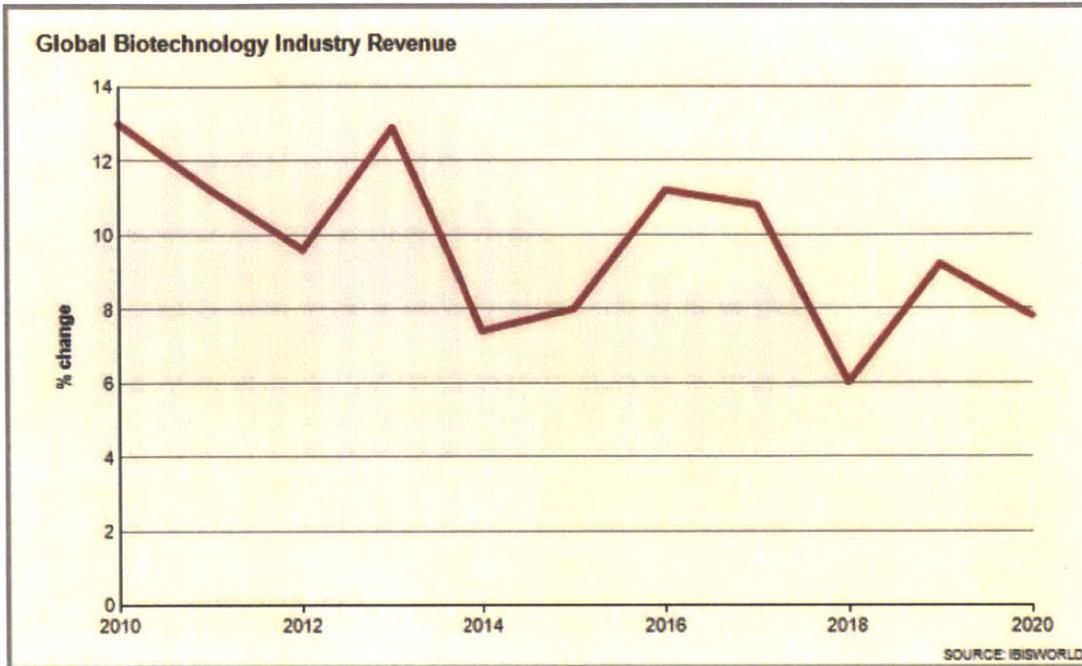


Figure 3: Recent and projected biotechnology industry revenue (Phillips 2014). *Growth has been largely positive and is expected to remain so, though the industry will most likely approach maturation over the next decade.*

Though the industry includes markets such as industrial & environmental, agricultural, and animal health technologies, the largest market is human health technologies (biologic medications), as seen below in Figure 4.

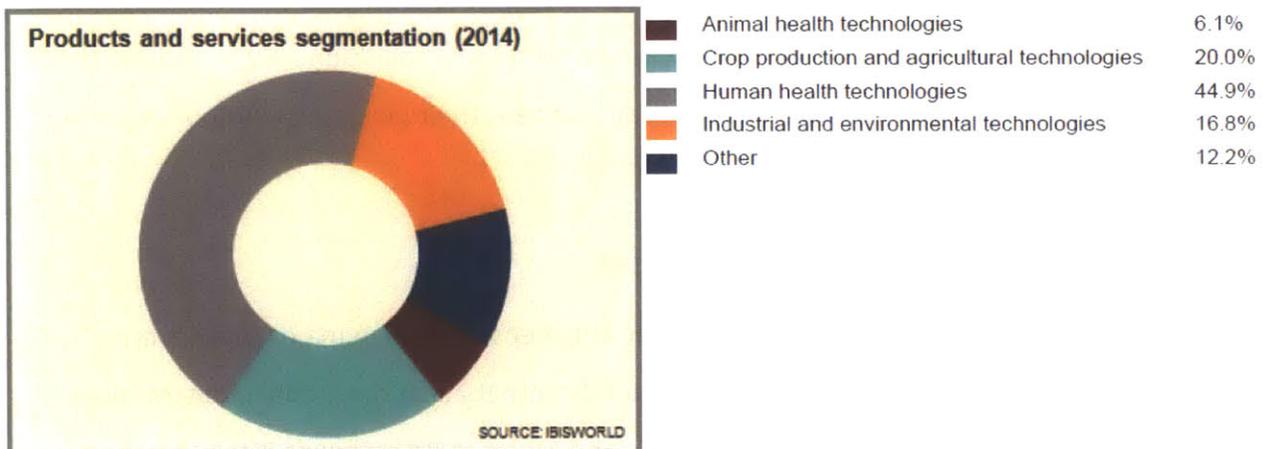


Figure 4: Biotechnology industry product and service segmentation (Phillips 2014). *Human health technology is the primary offering of the biotechnology industry.*

The total market performance is driven by:

- An increasing global population aged 65 and older;
- Increasing healthcare spending per capita, especially in member nations of the Organization for Economic Cooperation and Development (OECD);
- An increase in the number of students pursuing education in the life sciences;
- Increased public and private investment in R&D; and
- Advances in technology, such as next generation sequencing.

On the other hand, current risks to market growth include the increasing cost of medications, additional scrutiny from payer organizations, and increased regulatory burden.

The advent of biosimilar medications does represent a risk in the sense that they will cut into the revenues of legacy products, but they are also an opportunity to access less affluent markets with fewer resources available for medical treatment. Biosimilars are biotechnology's analogues to the generic medications of the small-molecule industry. However, unlike generics, they are not identical to the previously patented drug. Instead, they achieve high levels of similarity with the established product by following the same development procedure. Biosimilarity is defined by the Food and Drug Administration (FDA) as "highly similar to the U.S.-licensed reference biological product notwithstanding minor differences in clinically inactive components; and that there are no clinically meaningful differences between the biologic product and the reference product in terms of the safety, purity, and potency of the product" (FDA 2010). This enables the creation of a drug that performs similarly without requiring the longer, more costly approval pathway for innovator products. The nature of biologic drug development and the current level of technology make it prohibitively expensive, and perhaps impossible, to determine the entire chemical structure of every biologic dose with complete certainty.

However, biosimilars pose a risk to established biologic drug producers because they can provide a much lower-cost alternative medication to the market. The lower costs are possible because the biosimilar developer does not need to fund and maintain a discovery and

development pipeline for completely new drugs. These lower costs translate into lower prices for patients, requiring the established producer to significantly reduce prices or face a significant reduction in active patients.

Biosimilars are a relatively new market consideration, especially in the United States. The FDA only recently established a unique, abbreviated pathway for biosimilar approval in March 2010. As a result, only two approved biosimilar drugs were available in the U.S. market as of October 2014. Europe, on the other hand, approved 21 biosimilars from 2006 through the same point in time (Fazzolare and Brougher 2014). Biosimilar applications in the United States have continued to trickle in, as seen in Figure 5 below. Commercial investment in biosimilars is expected to increase over the next several years (IMS Health 2012), though their future long-term prevalence is still a matter of much debate. The Congressional Budget Office estimates that they may save consumers up to \$25 billion, though this savings would represent a small portion of the entire \$2.7 trillion U.S. healthcare system (Silver 2013).

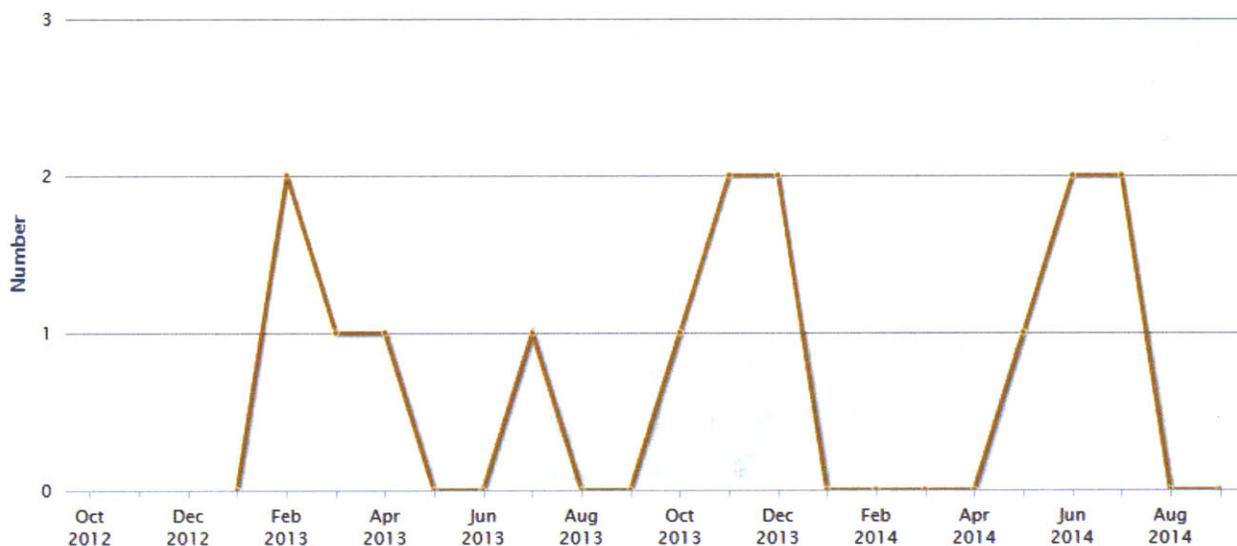


Figure 5: Number of biosimilar IND applications received per month in the United States (FDA 2014). *Application rates continue to be low in the United States but may increase due to patent expiration and further regulatory development.*

Overall, medications produced via biotechnology rather than traditional small-molecule methods have the advantage of targeting the underlying causes of medical conditions by using human proteins rather than treating symptoms with synthetic chemicals. This mechanism is

enabled by the more complete understanding that resulted from mapping the human genome. Biologic drugs have experienced their greatest success in the realms of oncology, diabetes, inflammation & arthritis, neurological disorders, HIV, and more.

2.2. Biologic Drug Development

The discovery and development of new biologic drugs is a long and costly process. It begins with the discovery of a new molecule, and its encoding genetic sequence, that can treat an indication associated with a certain disorder. There are usually a large number of potential candidates that have been discovered, but it is impractical to develop all of them. Instead, producers must evaluate each option based on estimated likelihood of clinical success and market dynamics, among other considerations.

After discovering a viable molecule and determining that commencing a drug development program for it would be worthwhile, the drug producer proceeds into CLD, which is described in detail below in Section 2.2.1. Additional R&D activities include determining the ideal parameters for large-scale production and processes for downstream processing, as these two aspects can vary between drug programs.

2.2.1. Cell Line Development (CLD)

After deciding to move forward with a particular genetic sequence and corresponding protein, CLD can commence. This is the process by which an ideal cell is created to produce the protein of interest, and, at Pfizer, includes 57 individual steps over a six- to twelve-month timeframe. The details of these processes need to be obscured for confidentiality purposes but their ultimate goal is to create the best possible host cell that has incorporated the gene(s) of interest. Upon the completion of the development process, this host cell establishes the master cell bank (MCB), which becomes the parent for all future production of the drug.

To create this ideal parent cell, a number of host cells are transfected with the gene(s) of interest. This host cell is most commonly a Chinese Hamster Ovary (CHO) cell, due in part to its proven track record over the past decades, its ability to grow in suspension cell cultures, its low capability to propagate human viruses, and its ability to grow in serum-free, chemically defined

media, which enables reproducibility of the development process (Lai, Yang, and Ng 2013). However, other cells may be used, such as bacterial cells (e.g., *Escherichia coli*), yeasts, and viruses. These alternatives tend to have certain advantages, such as the rate at which they reproduce. The doubling time of CHO cells varies between specific strains but tends to be on the order of 22 hours (Invitrogen 2007). *E. coli*, on the other hand, can double in as little as 30 minutes (Expression Technologies 2003).

After selecting the host cell type, a group these cells are gathered and transfected with the gene(s) of interest, then separated into single-cell populations. This step is critical in confirming monoclonality, i.e., the fact that a drug is produced by a population of cells that has grown from a single parent cell. Without this confidence, there is a possibility that multiple populations within one cell culture could mutate and diverge from each other, compromising the integrity of the drug product and potentially interrupting supply. Therefore, the regulatory agencies require high levels of confidence in the monoclonality of biologic drugs.

After separating single cells into their own containers, each cell is grown up to a size that is sufficient to analyze its growth rate (usually measured in doubling time) and titer (the production of the protein of interest, measured in grams of protein per liter of cell culture). The number of cellular populations to evaluate and the thresholds for growth rate and protein production are levers that the developer can vary. For example, increasing the number of populations to evaluate will increase the likelihood of creating the ideal clone, but requires a much larger investment in time and resources.

The next step is confirming that the highest-performing clonal populations have continued to express the gene(s) and protein of interest, i.e., that they have not significantly mutated. The eventual clone that is nominated to move forward for production must be stable in this sense for at least the amount of time that the production runs will take in the manufacturing facility. Production runs can take several months and include seeding successively larger bioreactors as the cellular population grows before extracting and purifying the expressed protein. This stability verification process is one of a few interactions between CLD and Analytical Research &

Development (ARD), another R&D organization within Pfizer. This group is responsible for analyzing cellular populations and other biological materials for various groups within R&D.

Before moving on to the MCB phase, the clonal populations need to prove that they are capable of performing highly not just at the micro-scale in a laboratory, but also at larger scales that are more representative of the eventual production environment. Therefore, the CLD scientists scale up the top-performing clones to a sufficient volume for bioreactor evaluations, in which each clone is assessed based on titer, growth rate, and additional specific metabolites. After this performance (along with longer-term stability) has been confirmed, the highest-performing clone can advance and become the MCB.

The entire CLD process can be broken down into several high-level phases, from transfection through master cell banking, as shown below in Figure 6. Each individual process step that occurs over the course of CLD fits into one of these phases.

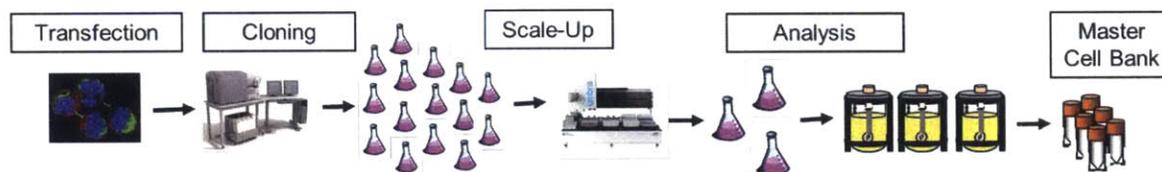


Figure 6: Overview of the cell line development process.

All of these tasks are performed at Pfizer by a group of approximately 20 scientists and four managers within the CLD group. They work with several other groups within R&D to develop each drug program, including ARD, the groups that determine ideal production and downstream purification procedures, and others.

Typically, current cell line development approaches take approximately six to twelve months and are considerably time- and capital-intensive. Most biopharmaceutical companies use CLD technologies based on methotrexate amplification or glutamine synthetase. A more detailed overview of such a CLD process is shown below in Figure 7. There are many technological options for each stage of this procedure. For example, single clones can be separated into their own vessels after transfection and selection via several approaches,

including limiting dilution, fluorescence-activated cell sorting (FACS), colony picking, and laser-enabled analysis and processing (LEAP) (Lai, Yang, and Ng 2013; Sleiman et al. 2008; Sharfstein 2008).

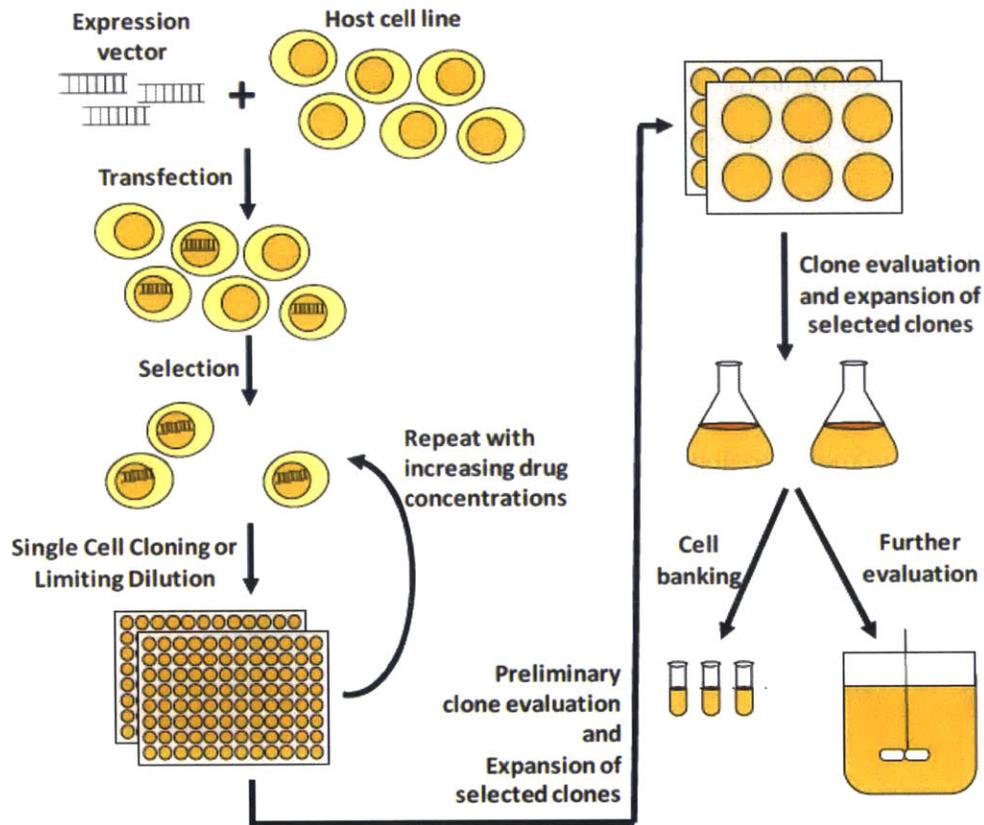


Figure 7: Overview of a typical cell line development process (Lai, Yang, and Ng 2013).

The transfection phase that precedes cell selection and cloning does present an opportunity for advances. Specifically, recent developments in synthetic biology have allowed the identification and manipulation of specific segments of a host cell's genome. Historically, transfection would occur randomly throughout the host cell's DNA and would vary significantly throughout the population of transfected cells. Such heterogeneity is one reason that a large number of cells needed to be screened. Due to the randomness of the transfection, many more clones needed to be generated and screened to find the rare high-performer. In a more homogenous population, each cell is very similar to the rest of population by design, which may someday require reduced screening. Some drug developers have been able to create such populations with reduced heterogeneity. Using recently-adapted enzymes, they are able to insert genes of

interest precisely where they may be most effective. This targeted transfection approach provides the benefit of improved protein production as well as homogeneity within the population of transfected cells (Lanza, Cheng, and Alper 2012).

After transfection, selection, and cloning, the single-cell colonies are scaled up and evaluated based on their growth rates and protein production levels. This can be performed manually, though semi- and fully-automated approaches are feasible. For example, cell cultures may be passaged using a hand-held pipette or performed by a liquid handling machine, such as SciRobotics' Pickolo or Hamilton's Star, Vantage, or Nimbus solutions. Additionally, growth rates can be estimated visually or by using an automated imager, such as Molecular Devices' Clone Select Imager (CSI) or Solentim's Cell Metric. Newer imagers are also beginning to incorporate additional functionality, such as on-board assays to automatically determine growth rates and titers, such as SynenTec's Cellavista system. Once the cell colonies have grown to sufficient size, additional options become available for cellular assays, such as Gyros' Gyrolab, Beckman Coulter's Vi-Cell, and TAP Biosystems' advanced microscale bioreactor (AMBR) (Keil 2014).

2.2.2. Clinical Trials, Patents, and Exclusivity

After a final confirmatory test to ensure that the drug is not toxic and successful approval of an Investigational New Drug (IND) application, clinical trials can begin. The purpose of these trials is to determine whether a new drug is safe and effective, and if so, what the dosing schedule should be. There are three phases of required clinical trials (Stilling 2013):

1. Primarily targeted at confirming safety, phase I trials determine how to administer the drug to patients, how the human body processes the drug, the maximally tolerated dose, and any significant side effects. These trials are run on relatively small populations of very ill patients (approximately 20 to 100).
2. Phase II trials are conducted on several hundred patients who all have measureable indications of the target disease. The goal of this phase is to determine if the drug successfully treats the disease.

3. Finally, several hundred or thousand patients undergo phase III trials to determine if the new drug is a significant improvement over currently available treatments.

In the United States, upon successful completion of all three phases of clinical trials, the developer can submit a Biologic License Application (BLA) for the new drug (as opposed to a New Drug Application, or NDA, for small-molecule drugs). Upon approval, the developer may market, transport, and sell the new biologic drug in the United States.

At any point during this process, the developer may create a patent for the drug, provided that it is an innovative drug rather than a biosimilar. Such a patent will last for 20 years in the United States (FDA 2015). The timing of patent submission is up to the discretion of the drug developer; patenting earlier supports exclusivity during development but can reduce the period of time during which the drug is commercially available and still patent-protected. However, regardless of the timing of patent submission, the developer is guaranteed 5 years (for new drugs) or 7 years (for orphan drugs) of exclusivity upon approval from the FDA. This stipulation is provided independently of, and can overlap with, any patent protection (FDA 2015). Because of this limited period of time during which to recuperate the costs of R&D, any way to reduce the development timeline is highly valuable to drug development organizations. Furthermore, all of this pre-clinical development work is performed, in a sense, at risk. That is, the product may fail to pass the IND application, any of the three phases of clinical trials, or the BLA.

2.2.3. Development Risk

R&D processes, which include everything from initial molecular discovery through clinical trials, can span several years with no guarantee of success. This large time investment makes R&D in this industry a high-risk, high-cost endeavor. Commercial drug production, which includes large scale manufacturing and distribution of the finished drug, is the payoff that results from undertaking this risk. Therefore, R&D serves as a delay after discovering a potential new drug and before beginning to earn revenue for its production. This makes R&D operations a significant focus for biopharmaceutical firms, and is part of the reason why market prices for commercialized drugs can be perceived to be very high. They need to enable the producer to

recover the costs of developing not only the drug that successfully made it to market, but all of the other drugs that failed to successfully pass the BLA phase.

The probability of success at each stage of the development timeline varies based on many factors. As seen below in the third row of Figure 8, new drugs under investigation have a relatively high chance of failure at each stage in the process.

FDA classification ^e	Phase 1 to phase 2				Phase 2 to phase 3				Phase 3 to NDA/BLA				NDA/BLA to approval			
	Total in phase ^a	Advanced or suspended ^b	Phase success ^c	Phase LOA ^d	Total in phase ^a	Advanced or suspended ^b	Phase success ^c	Phase LOA ^d	Total in phase ^a	Advanced or suspended ^b	Phase success ^c	Phase LOA ^d	Total in phase ^a	Advanced or suspended ^b	Phase success ^c	Phase LOA ^d
All indications	2,541	1,918	64.5%	10.4%	3,743	2,268	32.4%	16.2%	1,554	975	60.1%	50.0%	908	659	83.2%	83.2%
NMEs	1,585	1,218	64.2%	7.5%	2,375	1,470	28.6%	11.6%	831	515	53.2%	40.7%	425	293	76.5%	76.5%
Biologics	572	411	68.4%	14.6%	819	464	37.9%	21.3%	320	182	63.2%	56.1%	159	116	88.8%	88.8%
Non-NMEs	218	168	66.7%	20.0%	355	226	45.1%	29.9%	321	234	75.6%	66.3%	293	227	87.7%	87.7%
Lead indications	1,770	1,336	66.5%	15.3%	2,070	1,247	39.5%	23.1%	1,009	633	67.6%	58.4%	664	472	86.4%	86.4%
NMEs	1,094	848	65.2%	12.0%	1,275	791	36.4%	18.3%	497	300	61.7%	50.3%	283	185	81.6%	81.6%
Biologics	362	257	75.1%	20.8%	403	216	44.0%	27.7%	182	106	71.7%	63.1%	106	75	88.0%	88.0%
Non-NMEs	167	124	66.9%	23.2%	232	153	49.0%	34.6%	254	186	79.0%	70.7%	246	189	89.4%	89.4%

Figure 8: The probability of success at each clinical trial phase and NDA/BLA approval, based on 4,451 drugs investigated between January 2003 and December 2011. *The total likelihood of approval (LOA) for a new biologic drug entering phase 1 trials is approximately 14.6% (Hay et al. 2014).*

In other words, for each new biologic drug that receives approval for its BLA, the following number of drugs must enter each clinical trial phase (see the third column of Table 2, below):

Table 2: Number of new drugs that must enter each phase of clinical trials to result in one successful, commercial drug. *For each successful drug, approximately seven must begin the clinical trial process.*

Phase	Phase Success Rate	Number of New Drugs Entering Phase per BLA Approval
Phase I Trial	68.4%	6.9
Phase II Trial	37.9%	4.7
Phase III Trial	63.2%	1.8
BLA Approval	88.8%	1.1

Later stage (i.e., phase III) drugs tend to be held to a higher standard than phase I drugs in terms of performance and host cell line suitability. This fact combined with the risk associated with developing new drugs serves as a major incentive for manufacturers to make early-stage

development as fast and low-cost as possible. One way in which this occurs is by postponing as much work as possible during phase I development and revisiting it if the drug program does in fact proceed to phase III. Given that approximately 26% of drugs entering phase I will reach phase III (68.4% x 37.9%), a producer can achieve savings by delaying assays that determine phase III suitability until after the results of phase I or phase II trials. The downside is that this may result in the need to redevelop cell lines for a portion of the 26% of programs that do move on to phase III but have slight phase III suitability issues. However, this may be outweighed by the savings that result from never performing these resource-intensive phase III suitability assays and associated administrative work on the 74% of new drug programs that never make it to phase III.

Drug producers are further incentivized to streamline early-stage development by the large first-mover’s advantage in the biopharmaceutical industry. It has been shown that drug programs that reach commercial markets sooner are more valuable than higher-performing drugs for the same indications that arrive on the market later. This fact is shown below in Figure 9.

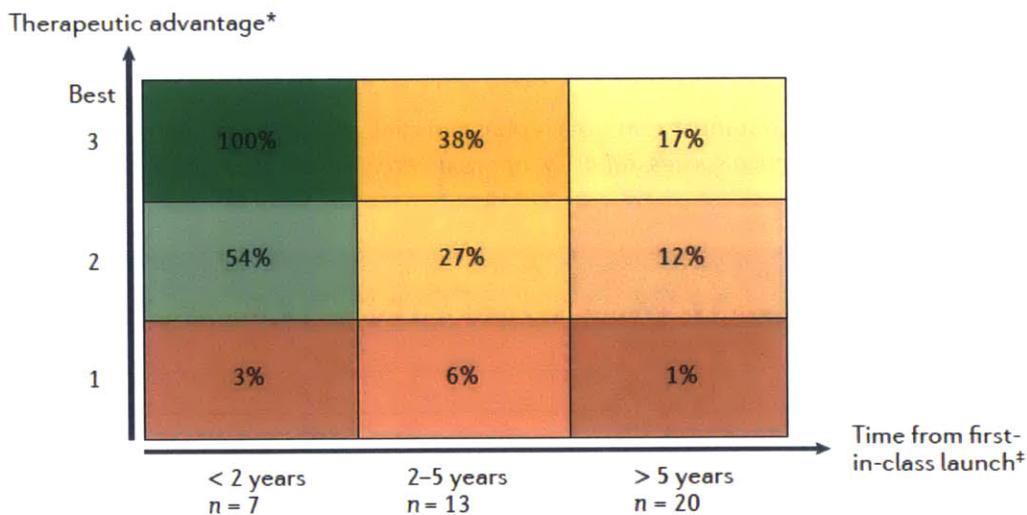


Figure 9: Value captured by follow-on entrants as a function of time of market entry and therapeutic advantage. *The value of a drug program decreases more significantly due to delays reaching market than it does for decreases in therapeutic advantage. *Compared to others in same mechanistic class. [‡]Defined by date of approval by the US FDA (Schulze and Ringel 2013).*

The risk associated with developing new biologic drugs extends down to the most granular level as well. For example, individual operations within the CLD system can have varying degrees of success. Each step usually proceeds more or less according to platform specifications, but occasional delays can occur due to poor gene incorporation, cellular growth, protein production, and more. Some of this variation is observed and discussed below in Section 4.1. In the worst case, a process step could fail due to an inherent characteristic of the drug program, which necessitates reevaluating the entire program based on molecular structure, transfection mechanism, or other consideration.

2.3. Pfizer, Inc.

Pfizer, Inc. is a pharmaceutical and biotechnology company headquartered in New York, NY that discovers, develops, manufactures, markets, and distributes medications for many types of diseases. The majority of its products are small-molecule pharmaceuticals, though it has recently entered the biologics market and is growing in this space. Focusing on human health technologies, Pfizer's most successful products include Advil[®], Enbrel[®], Lipitor[®], Lyrica[®], Plevnar[®], and Viagra[®]. Founded in 1849, Pfizer remains one of the world's largest producers of human health technologies, with 2013 revenue of \$52 billion and 77,700 employees (Pfizer, Inc. 2014).

Although Pfizer is one of the largest biopharmaceutical organizations in the world, it has a relatively low share of the biologics market due to its relative recent entry into this market. Historically an exclusively small-molecule pharmaceutical company, Pfizer entered the market in 2009 when it acquired Wyeth. However, biologics are becoming a more integral part of Pfizer's portfolio, as its contribution to total revenues has increased from 17% (\$9.9 billion) in 2011 to 22% (\$10.7 billion) in 2013 (PMLiVE 2014).

3. Literature Review

As discussed above in section 2.2, operational efficiency in R&D is incredibly valuable to biopharmaceutical organizations. Any possible timeline reductions in the CLD process can help achieve this goal. However, there is a point of diminishing returns that is established by the natural lifecycle of the host cell lines. That is, there is currently no reliable method for increasing the multiplication rate of host cells, resulting in a minimum time requirement simply to achieve enough cell mass to perform the requisite tests and establish an MCB. However, similar value can be derived by enabling a CLD operation to process a greater number of drug candidates, with the anticipated result being a greater number of successful drug approvals. This section discusses several future technologies that may enable such a capability without incurring significant additional cost. It also discusses potential modeling approaches that are capable of identifying which of these upgrades are the most beneficial.

3.1. Future Methods

In addition to the approaches outlined above in section 2.2.1, several upcoming technologies may be able to further improve the efficiency of CLD operations. For example, microfluidic cellular processing and assessment can relieve some of the resources currently required, while also providing superior cell culture properties (Lin et al. 2011). However, microfluidic platforms are still being developed and have not been notably adapted to industrial environments yet.

On the other hand, nascent microfluidic approaches have demonstrated the capability of performing several processes that comprise CLD operations, including transfection, cell sorting, and growth assays (Adamo et al. 2013; Bryan et al. 2014; Kimmerling et al. 2014; Sharei et al. 2013; Sun, Kovac, and Voldman 2014). These approaches are able to process and evaluate large numbers of cells either partially- or fully- automatically, making them promising in situations in which large cell populations are required. This ability to operate at such a large scale give them the potential to improve the efficiency in cases of random genetic integration discussed in Section 2.2.1. Their value is less fully realized when small populations of cells are needed, as is

the case with targeted transfection, as the workload associated with traditional methods does not escalate as drastically when fewer numbers of clones are needed.

3.2. Modeling Approaches

Many promising technologies that have been developed were briefly outlined above in Sections 3.1. Many of them have been hypothesized to improve the efficiency of CLD procedures, yet there is very little in the literature that illustrates their operational performance (i.e., their effect on total development timelines and resource requirements) as part of the end-to-end development system. Therefore, we aim to evaluate such effects of these and other approaches. Several software solutions are capable of providing these insights, each with their own advantages and disadvantages. Microsoft Excel has been shown to be successful for operational assessments in biopharmaceutical organizations, and has the advantages of low capital expenditure and the fact that most managers have a strong working knowledge of it (Heyman 2010). However, more powerful tools do exist and have been shown to effectively model operational environments, such as Simio, though the organization employing them would need to invest considerably more resources into acquiring them and training staff. Such tools have been successful in modeling large-scale manufacturing and service operations environments but little has been done in more complex environments, such as biologic drug development (Guo-giang and Ding-zhong 2013; Yang et al. 2013). Therefore, we select Microsoft Excel as the modeling tool for this investigation. We also add Monte Carlo simulation via Oracle Crystal Ball to accommodate the uncertainty in some of the input parameters, as described below in Section 4.

4. Modeling Current and Potential States of Development Operations

In this section, we model the explored procedures and technologies for CLD processes to determine the optimal system from an operational perspective. Interviews with pertinent stakeholders identified that the primary goals include reducing employee utilization while increasing the number of molecules developed. These objectives support the overall goal of enabling Pfizer to develop a larger number of new drug programs without increasing labor costs or sacrificing quality. Reducing cycle time is a secondary objective, as it is always beneficial but limited by the natural lifecycle of host cells, as discussed above in Section 3. These parameters are evaluated as functions of the technologies and procedures employed, the number of drug development programs that the CLD team must develop host cell lines for, the additional responsibilities of CLD personnel, and other inputs discussed below in Section 4.1. Therefore, the magnitude of any improvements will be measured as the incremental reductions in two parameters: (i) total system utilization per drug program and (ii) total end-to-end cycle time for a new drug development program's CLD process.

The model is first compiled using the details of the current development environment at Pfizer, Inc. We validate the structure of the model by comparing its original outputs to the current state in terms of project throughput and employee utilization. Then, we update various input values to determine their effect on the overall CLD system and identify a new system design using the evaluated approaches that achieves the maximum improvement. Furthermore, a potential future technology is added to this new system design to observe its potential incremental benefit.

4.1. Model Inputs

The inputs for the operational model include the 57 individual steps in the CLD process, the various technologies and procedures that can be applied to them, various administrative factors, and financial considerations. Covering all of these areas enables a detailed analysis of the end-to-end development process as well as a cost-benefit analysis for future investments.

In this model, the inputs covering the available technologies and procedures are broken into ten distinct phases:

1. **Preparation through Recovery:** covering the beginning of the CLD process, including host cell and DNA preparation, through post-transfection recovery (the period of time in which the transfected population is isolated from other materials and grown to a sufficient size for cloning);
2. **Cloning:** the separation of transfected populations into single-cells for purposes of establishing monoclonality, as discussed above in Section 2.2.1;
3. **Screening:** evaluating the populations that have grown from these single cells for growth rate and protein production, then selecting a group of the top performers to advance through the next steps of the process;
4. **Scale-up:** growing the colonies of top-performing clones to achieve sufficient material for further analysis;
5. **Clone Screen:** the head-to-head comparison of each top-performing cell line in a bioreactor-equivalent environment;
6. **Initial Stability Study:** ensuring that top-performing cell lines remain genetically stable, meaning that they do not significantly mutate over the course of an organizationally-determined number of generations. Such stability is critical to maintain production quality and drug supply over the duration of production runs in full-scale manufacturing environments;
7. **Analysis & Verification:** at several points throughout the process, assays are performed to evaluate various metrics that determine the quality and stability of the cell line, including the mass and makeup of the drug substance, the genetic makeup of the host cell line, and the RNA sequences that translate the genetic code into the protein(s) of interest;
8. **Pre-MCB (pMCB):** the frozen bank of the top-performing cell lines that, once a final clone is nominated, will be expanded and used to create the MCB;

9. **Bioreactor Evaluation:** an additional production run in relatively large-scale bioreactors to confirm that the performance of the top-performing cell lines does not significantly degrade at scale; and
10. **Complete Stability Study:** the extended stability study for additional generations to establish the maximum recommended lifetime of the MCB in production bioreactors before significant deteriorations in product quality occur, such as a reduction in protein expression or loss of stability.

In the operational model, these phases are represented as shown below in Table 3. The phases are shown in the left column, while active drop-downs on the right enable the user to select the procedure and/or technology being used for each phase.

Table 3: Procedural and technological input parameters for the operational model.

Procedural and Technological Inputs	
Phase	Approach
Preparation through Recovery	Select from drop-down
Cloning	Select from drop-down
Screening	Select from drop-down
Scale-Up	Select from drop-down
Clone Screen	Select from drop-down
Begin Stability Study	Select from drop-down
Analysis & Verification	Select from drop-down
Pre-MCB	Select from drop-down
Bioreactor Evaluation	Select from drop-down
Finish Stability Study	Select from drop-down

Each of the 57 individual operations that comprise the CLD process fall into one of these phases. For each phase, the model user selects from a list of options, which includes several current (as described above in section 2.2.1) and potential future (as described above in section

3.1) approaches. The overall utilization and cycle time contributions of each approach for each individual procedure are determined using four methods:

1. Direct observation of current drug development programs;
2. Performing and monitoring continuous improvement studies;
3. Interviewing experienced scientists both within Pfizer and at various vendors; and
4. Where necessary, estimating operational performance based on published specifications.

Each individual operation experiences some degree of variation in terms of its contribution to utilization and cycle time. This can be due to natural variation in the performance of each step that does not have any noticeable detriment to the overall process. However, as discussed above in Section 2.2.3, variation can occur due to systemic issues that can derail the entire drug program. Though not observed during the course of this investigation, such events should be considered and evaluated in parallel with the results of such modeling efforts.

The variation for many of the individual operations that comprise each phase in the current process is negligible, though several operations do have interesting distributions for these metrics. The average and standard deviation of these operations are shown below in Figure 10, and are used as the basis for the inputs in the model.

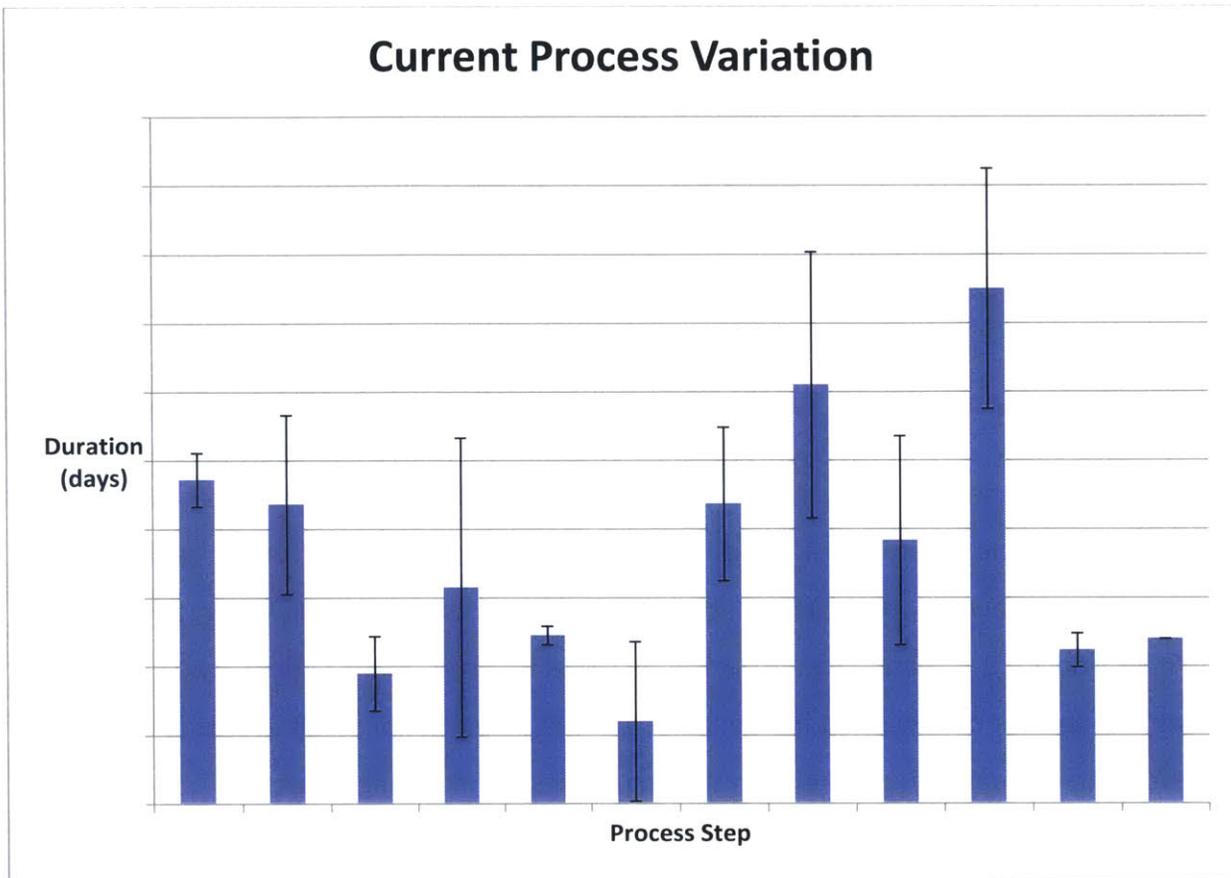


Figure 10: The average (blue columns) and standard deviation (error bars) in processing time for several of the process steps, as recorded for historical new drug programs. *Specific names of process steps and duration values are omitted for confidentiality purposes.*

However, due to the long duration and infrequent nature of these operations, a sufficient number of observations were only available for one of them. This particular operation very often follows standard procedures and has a cycle time close to the average. However, there are unique scenarios in which it is significantly delayed. The negative binomial distribution fits the gathered data regarding its cycle time most closely with a X^2 value of 0.3364, so this input is modeled accordingly as shown below in Figure 11.

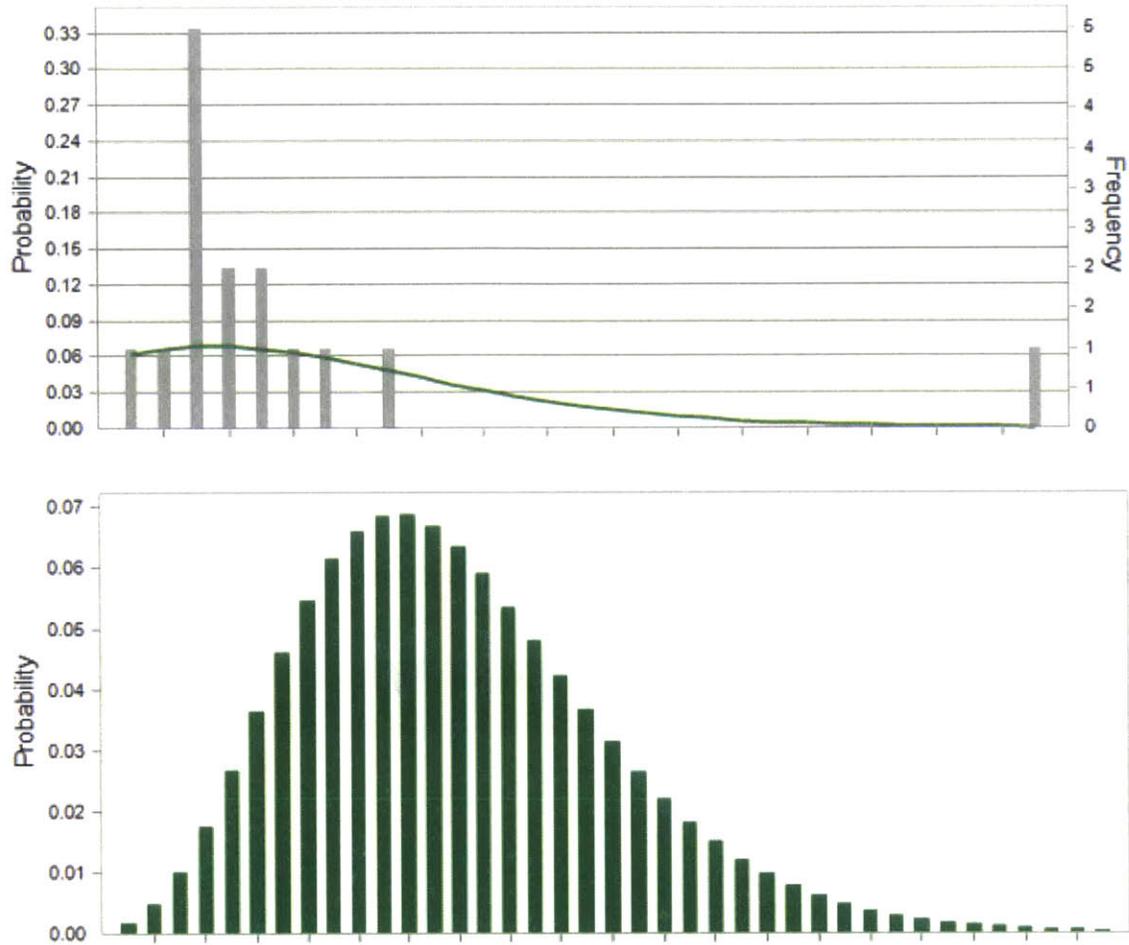


Figure 11: The distribution of cycle times observed for operation number 19 (top) and the resulting probability distribution (bottom). *The specific procedure's name and the units of its cycle time are obscured for confidentiality purposes.*

The remaining operations that exhibit variability are modeled as normal or lognormal distributions, though these assumptions should be reassessed in the future after a sufficient number of observations (i.e., at least 15) have been recorded. The operations with coefficients of variation greater than or equal to 0.3 are modeled as lognormal distributions due to the elimination of negative values and some heaviness in the right tail that captures the extreme worst-case scenarios (de Treville 2014). The remaining operations are modeled as normal distributions. The coefficient of variation for new procedures and technologies that were not able to be tested is assumed to be equal to that of the current approach for each phase. This assumption allows calculation of the variability in the total duration and hands-on time of the new approaches:

$$\sigma_{new\ approach} = \mu_{new\ approach} \times c_{v,current\ approach}$$

Where σ is the standard deviation, μ is the mean, and c_v is the coefficient of variation.

The administrative and financial inputs are captured in a second table on the user screen. These inputs include the total number of new development programs (“new starts”) per year, the number of molecules included in each new start, the timeline of the upstream discovery activities, the number of full time employees (FTEs) within CLD, the planned number of hours in a standard work week, and the material costs per new start (enumerated below in Table 4). These values are dynamically varied by the user.

Table 4: Administrative and financial input parameters for the operational model. *Exact values are obscured for confidentiality.*

Administrative and Financial Inputs	
Total New Starts / Year	X
Molecules / Program	X
Discovery: phase 1 (weeks)	X
Discovery: phase 2 (weeks)	X
CLD FTEs	X
Work Week (hours / week)	X
Material Costs (\$ / new start)	X
Late Stage Assessment Trigger	X

The discovery phase timelines are important because they establish the time points at which some molecules may be eliminated from the development process. For example, after a discovery program completes the exploratory toxicology (eTox) study and determines the level of safety of each molecule, an organization could decide to keep only some or one of the most promising, safest molecules. Such actions reduce the workload on the CLD team, though they do decrease the number of attempts at developing a successful biologic.

Furthermore, the model incorporates the CLD team’s responsibilities that are in addition to developing cell lines for new drug programs. This will vary between organizations, but can include reevaluating later-stage programs against more stringent criteria before commencing the relatively costly phase III trials, as captured above with the “Late Stage Assessment Trigger” input. This input can take the value of “New Start”, “Phase I”, or “Phase II”, meaning that the late stage assessment will be scheduled either when a new drug program begins, when it passes phase I clinical trials, or when it passes phase II clinical trials.

Other additional responsibilities can include institutional workload (e.g., meetings and training) and continuous improvement efforts. We capture these parameters as assumptions in the operational model that determine how much of an FTE’s time is unavailable for work related to new drug program development.

Finally, some of the CLD team’s responsibilities may be partially or fully outsourced. An additional input for these cases is added to capture these decisions by either keeping or eliminating their workload and adjusting their cycle times if needed. However, such approaches are not without costs, as detailed below in Section 4.2.

4.2. Tradeoff Considerations

All of the inputs outlined above in Section 4.1 have the potential to reduce total system utilization and drug development cycle times. However, when making such changes to a CLD operation, certain tradeoffs need to be considered. The operational model highlights the following tradeoffs:

Table 5: Summary of inputs and associated tradeoffs that need to be considered.

Input	Tradeoff Consideration
Procedural and technological approaches	Advancements improve utilization and cycle time in exchange for the cost of acquisition and training, as well as any risks outlined below in Section 5.

Input	Tradeoff Consideration
New drug development programs per year	A greater number of programs increases system utilization due to the additional new start and late-stage workload. However, the result is additional chances of reaching commercialization, per the success rates detailed above in Section 2.2.3.
Number of molecules developed per drug program	Developing multiple molecules for a single drug program allows CLD operations to begin before discovery operations are complete, resulting in an overall cycle time reduction. There is also the potential for increasing clinical trial success rates by selecting the molecule with the best-performing host cell line. Additionally, the producer can choose to update the number of molecules being developed during the course of CLD operations as discovery assays identify more information about them. However, this causes an increase in CLD utilization due to the additional workload of developing additional molecules.
Number of FTEs	Hiring additional employees relieves workload on the team as a whole and improves system utilization, but in exchange for increased labor costs. For the purposes of this project, the CLD team is constrained to the current number of FTEs.
Work week	Increasing the standard amount of time that each employee works would improve total system utilization at the cost of reduced employee morale.
Late stage assessment trigger	As discussed above in Section 2.2.3, postponing late stage assessments alleviates over-utilization of the CLD team. The tradeoff is an increased risk of needing to redevelop cell lines for some drug programs when they reach later clinical trials.

Input	Tradeoff Consideration
Outsourced workload	Contracting external organizations to perform development work or any of the CLD team's additional responsibilities improves total system utilization in exchange for outsourcing costs, supply chain risk, and intellectual property (IP) risk.

4.3. Results

The inputs and objectives described above in Section 4.1 are streamlined and presented in a single user interface, as shown below in Figure 12. This dashboard allows the user to manipulate each input and dynamically observe the effect on the outputs of interest (i.e., total system utilization and new drug development cycle time).

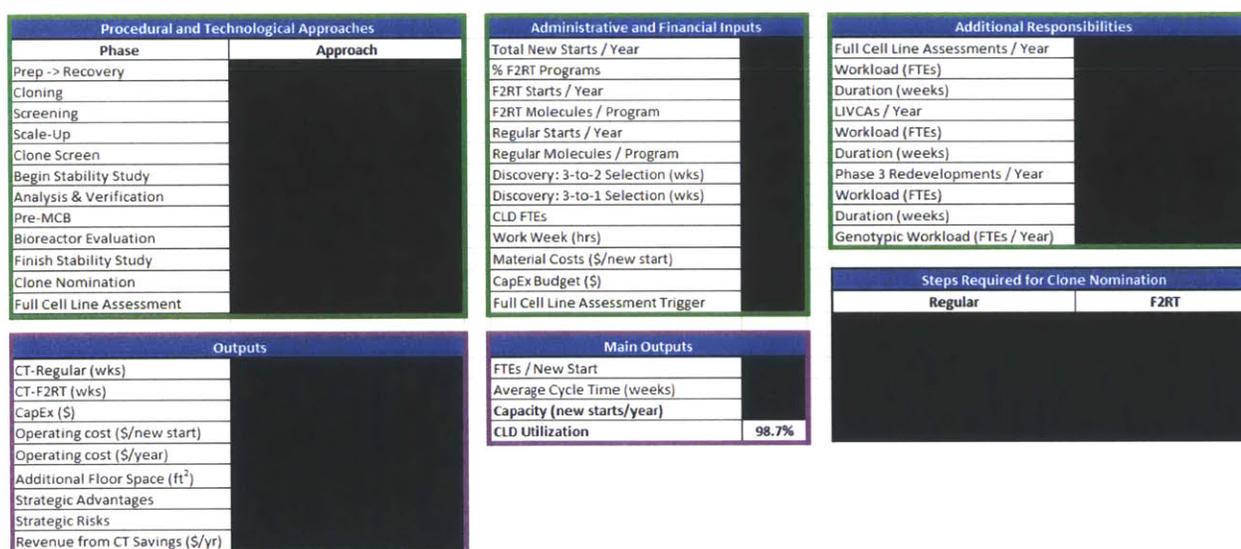


Figure 12: The operational model dashboard interface. *Specific input and output values are redacted for confidentiality purposes. Input fields are highlighted in green and outputs in purple. The additional field at the lower right allows the user to specify which of the 57 individual operations provide a specific deliverable that is needed for clone nomination.*

4.3.1. Current State and Emerging Bottlenecks

The model is first validated using the current state of CLD operations. However, in this case study, there is no previous data at this level of detail. Therefore, the output of the model is

verified by comparing it to the perceived state as determined by stakeholders in various positions throughout Pfizer’s R&D organization. The output of the model, shown below in Figure 13, Figure 14, and Figure 15, matches the current state as evaluated by these pertinent stakeholders.

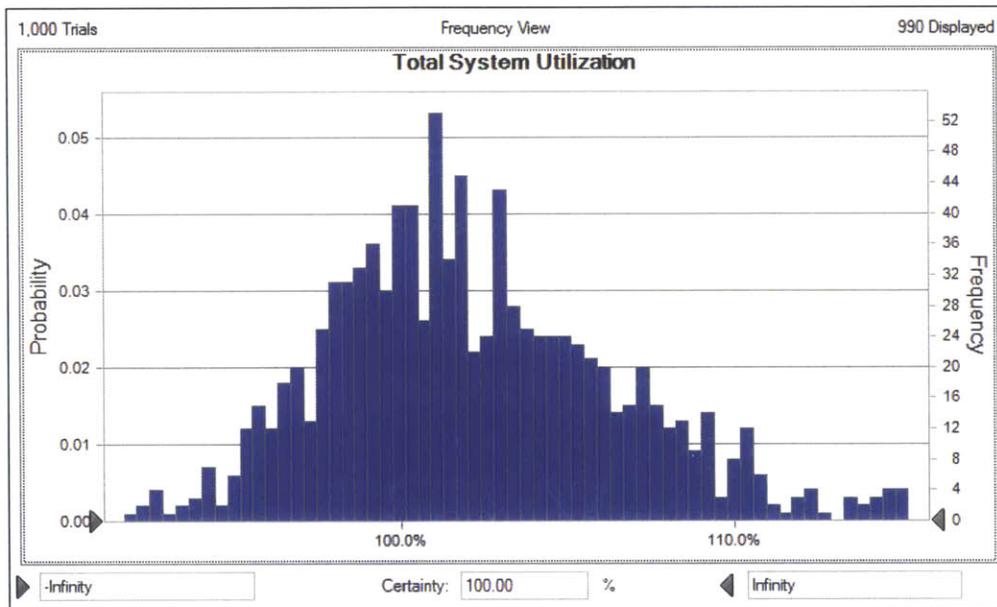


Figure 13: The total utilization of the CLD team as calculated by the operational model. *The average utilization is 102.3% with a standard deviation of 4.6%. This confirms the general consensus among stakeholders that the CLD team is slightly over-utilized yet still able to perform its responsibilities.*

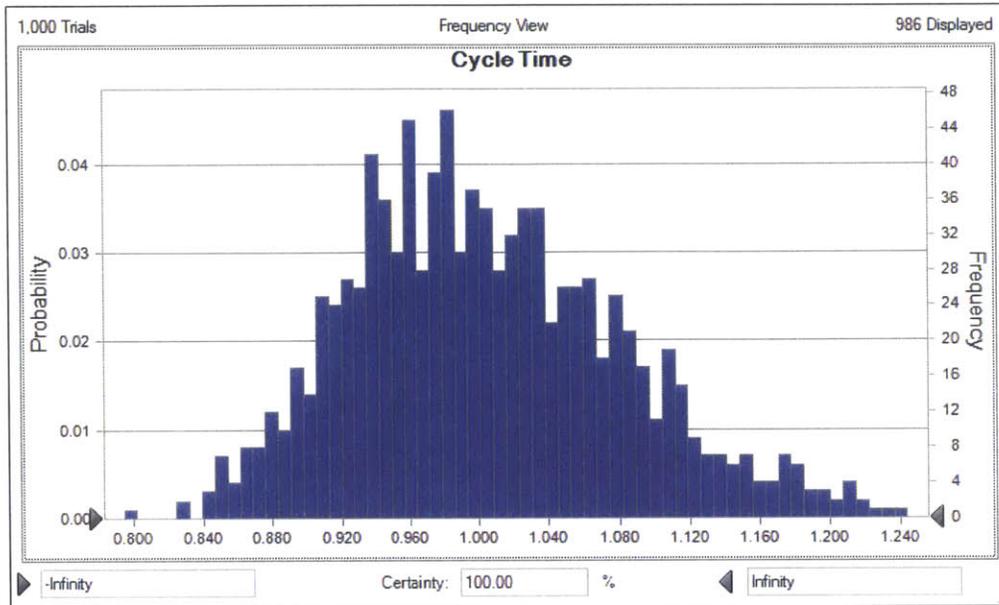


Figure 14: The average cycle time for all types of new drug development programs. *The values are normalized to a unit-less scalar representation for purposes of confidentiality. A value of 1.0 represents the current published cycle time for creating a host cell line for a new drug program at Pfizer. The base case average value is 1.007 with a standard deviation of 0.085.*

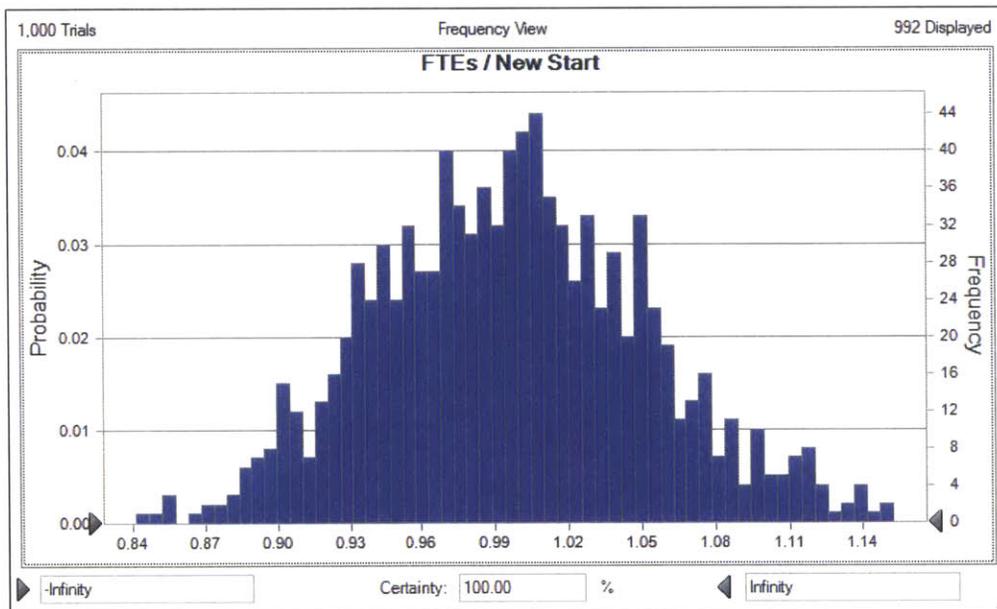


Figure 15: The number of FTE resources required to run a new drug development program through CLD. *The values are normalized to a unit-less scalar representation for purposes of confidentiality. A value of 1.0 represents the current resources needed to create a host cell line for a new drug program at Pfizer. The base case average value is 1.0 with a standard deviation of 0.06.*

These values for utilization, cycle time, and FTEs per new drug development program are confirmed with the current stakeholders that are familiar with the process and the current state of the CLD system.

Based on this initial assessment of the current state, several individual operations establish themselves as bottlenecks in the process and as highly FTE-intensive (strongly contributing to both cycle time and the high utilization level). Specific procedure names and metrics are obscured for confidentiality, but it is evident that the operations that occur in the middle of the CLD process are the most restrictive in terms of hands-on workload, as shown below in Figure 16.

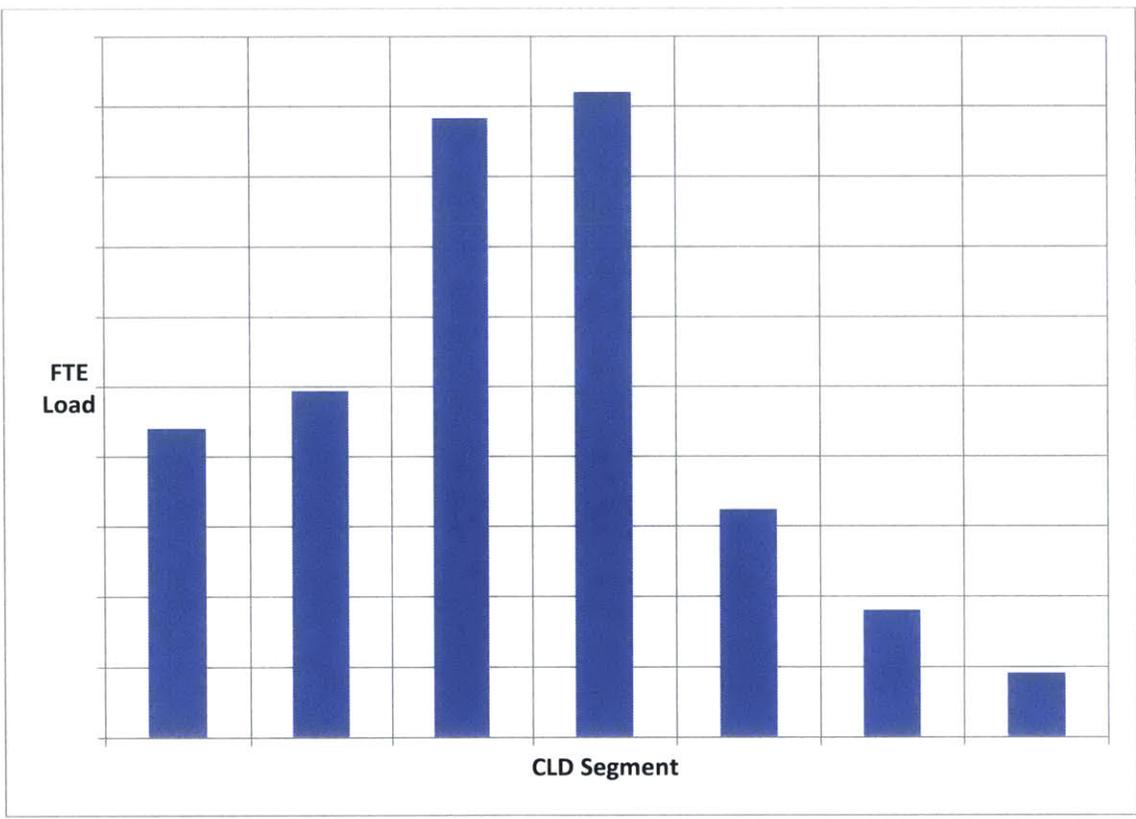


Figure 16: Hands-on workload as a function of the CLD process segment. *The horizontal axis represents the flow of time from the initiation of the CLD process through clone nomination and MCB. Workload spikes in the middle of the process before experiencing a significant drop towards the end of the process. Specific values for time segment and FTE load are omitted for confidentiality purposes.*

This distribution identifies the procedures in the middle of the CLD process as the primary points of interest for investigation of automation options. However, we do consider several approaches for individual procedures throughout the entire process.

4.3.2. Potential Improvement via Available Inputs

The dynamic nature of the model allows a user to easily identify which approaches have beneficial effects on the system and what combination of approaches results in the ideal updated state of CLD operations. However, finding the optimum system design becomes more difficult as the number of potential input approaches increases. Therefore, we create a Visual Basic script employing a random number generator to evaluate many possible permutations of approaches and identify the one with the best performance (i.e., the lowest total system utilization). This script identifies an improved state that uses a combination of available technologies and approaches. The specific approaches employed are obscured for confidentiality purposes, but the outputs of the resulting updated system design are shown below in Figure 17, Figure 18, and Figure 19.

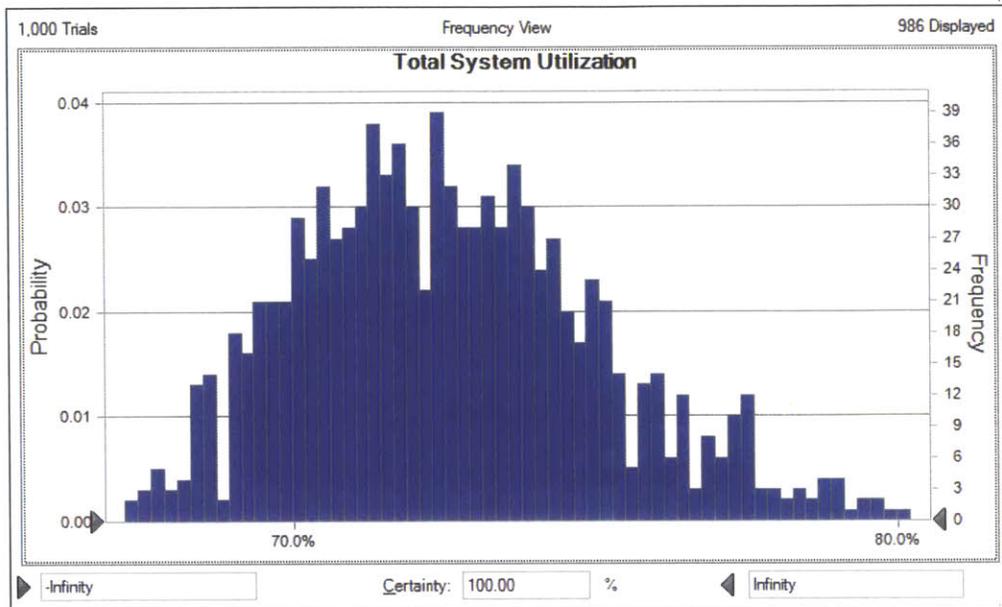


Figure 17: The total utilization of the CLD team using the highest-performing combination of inputs. *The average utilization is now 72.7% with a standard deviation of 2.7%.*

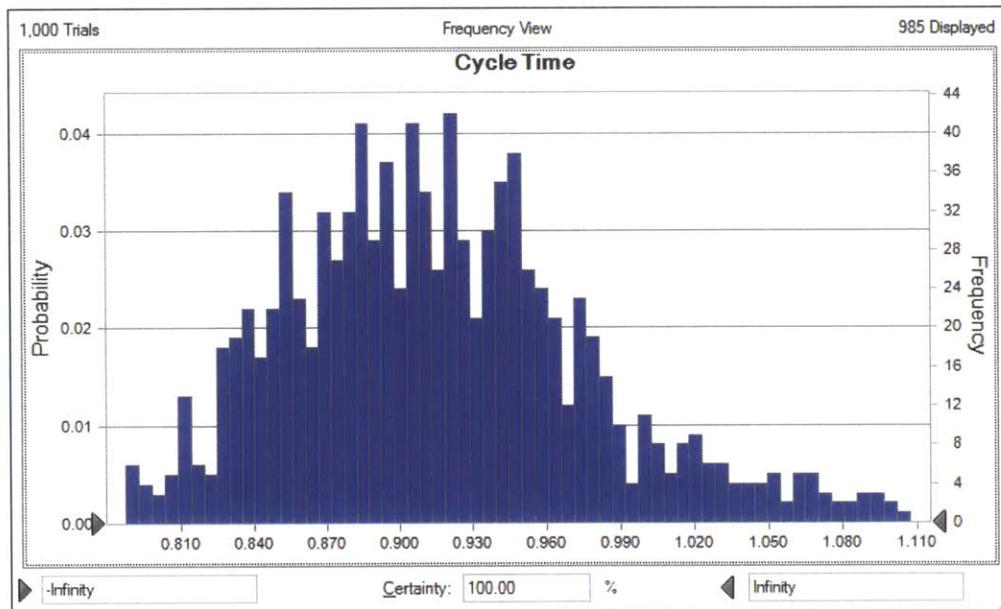


Figure 18: The average cycle time using the highest-performing combination of inputs for all types of new drug development programs. *The values are normalized to a unit-less scalar representation for purposes of confidentiality. A value of 1.0 represents the current published cycle time for creating a host cell line for a new drug program at Pfizer. The average value is now 0.919 with a standard deviation of 0.067.*

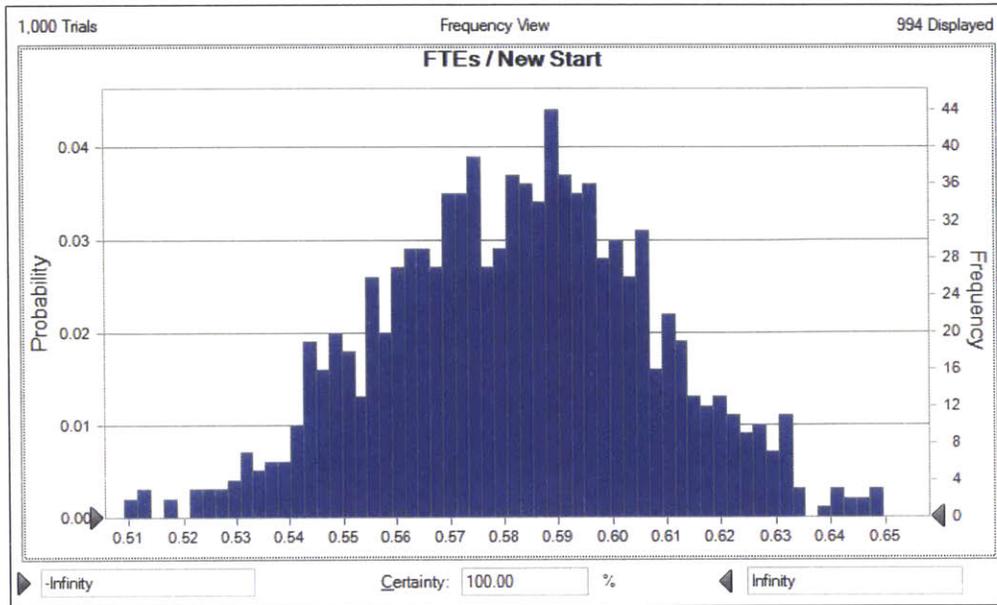


Figure 19: The number of FTE resources required to run a new drug development program through CLD using the highest-performing combination of inputs. *The values are normalized to a unit-less scalar representation for purposes of confidentiality. A value of 1.0 represents the current resources needed to perform create a host cell line for a new drug program at Pfizer. The average value is now 0.58 with a standard deviation of 0.03.*

This updated system design for CLD operations significantly reduces the main output of interest, which is the total system utilization. The operational model predicts that investing in certain new technologies and developing new procedures will enable the CLD team to work more efficiently, creating the opportunity to develop additional drug programs or perform other functions. The improvements of this new system design over the current state are summarized below in Table 6.

Table 6: Summary of base-case and updated output parameters. *The new system design provides a 29.6% improvement in total system utilization and an 8.8% reduction in cycle time. It also significantly reduces the variability of all three metrics.*

Average Values	Current State	New Design	Improvement
Total System Utilization	102.3%	72.7%	29.6%
Cycle Time	1.007	0.919	8.8%
FTEs / New Start	1.00	0.58	42.0%

Coefficient of Variation	Current State	New Design	Improvement
Total System Utilization	0.045	0.037	17.4%
Cycle Time	0.084	0.073	13.6%
FTEs / New Start Variation	0.060	0.052	13.8%

It should be noted that a large portion of the improvement in utilization stems from outsourcing one of the CLD team's additional responsibilities. This is a possibility that Pfizer is currently investigating and can be achieved at relatively low cost. The new system design without outsourcing this function results in equal outputs for cycle time and FTEs per new start as the previous case, but the total system utilization increases back up to 91.3%, or an 11% improvement over the current state. The cycle time and FTEs per new start are relatively unaffected because, as an additional responsibility rather than a new drug development operation, this particular function is not on the critical path to clone nomination. However, it clearly represents a significant portion of the workload that must be performed by CLD personnel.

4.3.3. Potential Future States

The operational model can also be used to evaluate potential approaches for future use based on nascent technologies. It is difficult for the user to know the operational performance of a new technology with a high degree of certainty without being able to test it, though estimates can be made based on published functionality. For example, a microfluidic approach that is ostensibly capable of transfecting cells, separating them into monoclonal colonies, monitoring their growth, and assessing their protein production, would theoretically reduce the cycle time all the way down to the minimum. The minimum cycle time in this case is the amount of time required for the selected host cell to grow from one cell to the number required to create the MCB. Furthermore, a platform that is capable of performing these operations fully automatically would reduce the hands-on workload of the CLD scientists during these operations to zero, aside from setup and post-processing activities.

Using these assumptions, we add a hypothetical microfluidic approach to the updated system design that was outlined above in Section 4.3.2 to evaluate the potential marginal improvement. Such an approach takes effect in the cloning and screening phases. The new processing time for post-cloning cell recovery is estimated to be the minimum time needed for the cells to grow, as discussed above. Both the normal and lognormal distributions are inappropriate for modeling this processing time, as it is nearly scientifically impossible for the cells to grow much faster than their natural metabolism allows. We therefore employ a Pareto distribution, with the location of the distribution equaling this minimum time. The shape of the Pareto distribution is calculated using the same coefficient of variation as the current state of recovery operations as follows (UAH 2015):

$$\sigma = \sqrt{b^2 \frac{a}{(a-1)^2(a-2)}}$$

Here, σ is the standard deviation and is calculated in the same manner as described above in Section 4.1. The parameter ‘a’ is the shape of the distribution and ‘b’ is the location. We know the location to be the minimum possible processing time for these operations, allowing us to determine the shape of the distribution. Employing this hypothetical microfluidic approach results in the following performance:

Table 7: Summary of microfluidic output parameters. *Adding a microfluidic approach to the new system design provides a 30.3% improvement in total system utilization and a 10.4% reduction in cycle time relative to the base case. However, incremental improvements over the new design are limited.*

Average Values	Microfluidic Approach	Improvement over Base Case	Improvement over New Design
Total System Utilization	102.3%	30.3%	0.7%
Cycle Time	1.007	10.4%	1.6%
FTEs / New Start	1.00	42.0%	0.0%

Coefficient of Variation	Microfluidic Approach	Improvement over Base Case	Improvement over New Design
Total System Utilization	0.03	22.8%	6.5%
Cycle Time	0.068	20.0%	7.3%
FTEs / New Start Variation	0.05	13.8%	0.0%

We see a similarly significant improvement over the current state of CLD operations in terms of all three main outputs of interest. However, the improvement over the new design described in Section 4.3.2 is much less significant. For example, it improves total system utilization by only 0.7% when compared with the previously proposed approaches for transfection, single cell separation, growth monitoring, and protein production assays.

5. Performance and Risks of Potential Approaches

This section provides an overview of the underlying technologies of a selection of the inputs discussed above in Section 4.1. It also investigates potential risks associated with the current and proposed approaches to CLD operations, as well as risk management frameworks that can be used to address these risks.

5.1. Review of Underlying Technology

The number of technological options modeled varies between each phase of the CLD process. This section highlights a few approaches for some of the more pertinent phases and provides an overview of the technology that drives them.

5.1.1. Cloning

The single cell separation phase, commonly referred to as “cloning,” is the first phase with several available approaches. They include flow cytometry, colony picking, and limiting dilution. Flow cytometry has been developed into several automated approaches, including FACS.

A FACS machine takes an input of a population of cells and distributes single cells from that population into separate wells. It selects single cells by forming a stream of cells, interrogating each cell with a fluorescent light source, and examining the scattering pattern of light reflected off each cell. The user inputs the desired scattering pattern, which is known to correlate with several desirable cell characteristics (e.g., healthy metabolisms, high quality growth rate, etc.). The machine then applies a charge to the stream of cells when the specified criteria are met, allowing cells of interest to be deflected by a voltage differential into the output plate while discarding all other cells. Designs can vary between manufacturers, but they are typically oriented as shown below in Figure 20.

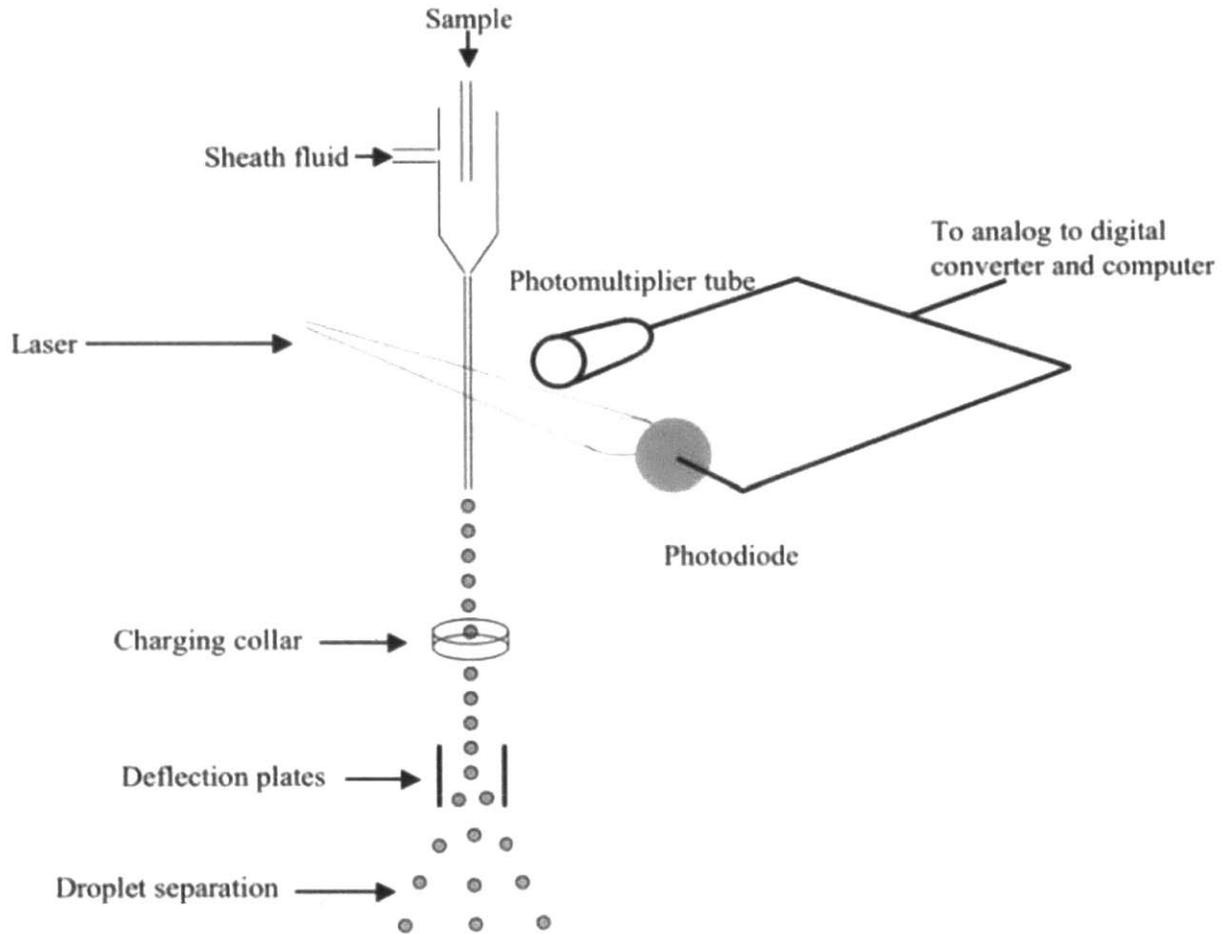


Figure 20: A representative design of a fluorescence-activated cell sorting (FACS) machine (Macey 2007).

This approach is superior to the previously-required manual approaches, due to its potential ability to automate this phase. However, real-world applications of this technology can require a user to monitor and control the entire process, limiting its upside. It does have the benefit of being a thoroughly-reviewed process that regulatory agencies are comfortable approving when confirming monoclonality of new biologics.

Colony picking is an alternative approach for single cell cloning. In this approach, populations of cells are deposited into a semi-solid medium after undergoing transfection. Over the course of a period of time afterwards, the cells are allowed to grow and are evaluated based on their expression of the protein of interest. This evaluation is performed by adding a target capture antibody that has been designed to pair with the protein of interest and an

optically detectable fluorophore. After identifying the highest-performing colonies, they can move on to the screening and scale-up phases (CDI Bioscience 2015).

The colony selection approach has the potential advantage over flow cytometry of eliminating a preceding step in the CLD workflow. Because each colony can be evaluated based on expression of the protein of interest, it may not be necessary to have a selection step after transfection and before cloning. That is, any cells that did not successfully transfect will be eliminated in the cloning step due to their inability to produce the drug substance. On the other hand, the probability of monoclonality is more difficult to establish in this approach. As opposed to flow cytometry, which evaluates and deflects single cells with high confidence (though not necessarily 100%), colony picking relies on the statistical probability of selecting colonies that have grown out from single cells based on the shape of the colony, its distance from other colonies within the semi-solid media, and several other parameters. Due to the regulatory agencies' recent focus on higher certainty of monoclonality, the colony picking approach does represent a significant risk.

Another cloning approach, limiting dilution, shows promise in both its ability to automate cloning and to provide a high probability of monoclonality. This process involves increasingly diluting a population of cells to lower and lower concentrations until a statistically significant determination of monoclonality can be made. For example, the final population could exist in many separate wells, each with a volume of 1 mL and a concentration of 0.01 cells per mL, meaning that each well has a 99% probability of containing zero cells, a 1% probability of containing one or more cells, and a 0.01% probability of containing two or more cells. The next step in the CLD process, screening, can identify and eliminate the vast majority of wells that do not have cells in them. The concentrations, volumes, and number of wells can be customized by the operator to ensure a statistically significant probability of monoclonality (Science Gateway 2015).

Cloning by limiting dilution has the potential to improve CLD operational efficiency because it can be automated using a variety of liquid handling machines and can be customized to provide the requisite degree of certainty of monoclonality. The accompanying risk is the fact

that this approach only works efficiently at scale when it is performed automatically. It is dependent upon the automated protocol and can create delays in the process in the event of any disruption to the liquid handling machinery.

5.1.2. Screening

The primary goal of this phase is to determine the post-cloning growth rate of each colony and select top performers to scale up. A secondary goal, though it is becoming more important, is being able to optically prove monoclonality. To do so, the screening approach must have a high enough resolution to be able to clearly show a single cell, which tends to have an average diameter of approximately 15 μm , and distinguish it from any condensation or other debris that may be in the well (Han et al. 2006). Several approaches based on similar technology have been mentioned as screening options: Molecular Devices' CSI, Solentim's Cell Metric, and SynenTec's Cellavista system. All of these systems perform cell screening by using optical information to assess the growth rate of various clones. The underlying technology is similar for all three approaches, though their operational capabilities vary.

The CSI operates by scanning a microtiter plate and evaluating the colony size in each well. It does so by taking four magnified, white light images per well at 3.7 μm per pixel and examining the portion of the well that is covered by cells, known as the confluence. The confluence at a number of time points is used to calculate the growth rate of each colony to enable selection of the fastest-growing clones. The advantages of such a device include the ability to image colonies without using potentially invasive biomarkers, relative high speed (approximately three minutes per 96-well microtiter plate) when compared to manual assays, and onboard calculation of growth rates. Additionally, Molecular Devices has developed an automated approach for high throughput screening in which a robotic arm can retrieve plates from a co-located incubator, place them in the CSI for imaging, and replace them in the incubator. On the other hand, the risks of devices such as these include reliance on the shape of each colony's growth due to the fact that only confluence is measured (as opposed to specific cell counts), resolution below what is needed to prove monoclonality, and, in the case of the

automated approach, down-time and maintenance due to several pieces and moving parts (Molecular Devices 2011).

Solentim's Cell Metric system provides an alternative that is based on the same technology but with some improved operational functions. It has an improved resolution of 2 μm per pixel and takes a much higher 50 images per well using a 10x objective lens. This comes at the cost of a slower processing time of three to four minutes per 96-well microtiter plate, though this is mitigated by the onboard plate stacker that automatically processes up to 10 plates at a time. Furthermore, the slightly increased cycle time of this operation is negligible when compared to the several month total CLD process, and is most likely an acceptable tradeoff for fully automatic processing. The improved resolution and images per well also allow CLD scientists to confidently establish the monoclonality of each clone. The downtime and maintenance cost risks still exist but are reduced due to fewer moving parts and the elimination of the need to interface between multiple devices (Solentim 2015).

We also evaluate an additional approach for imaging colonies during the screening phase. The Cellavista system provides the highest resolution of the three options at 1.86 μm per pixel by using up to a 20x objective lens, though the number of images per well is reduced to 16. An independent experiment confirmed that it is still capable of confirming monoclonality. It provides the potential for a fully automated approach by integrating with liquid handling devices from other vendors, though this reintroduces the risk of connecting multiple devices together with a higher number of moving parts. A unique benefit is the capability to perform onboard assays, enabling a CLD team to automatically obtain additional information such as protein production at this stage. These onboard assays can potentially improve overall system utilization by eliminating the need for subsequent assay steps in the development process. On the other hand, the Cellavista is the slowest of the three screening approaches described here, as it scans one 96-well microtiter plate in approximately four minutes. However, this negligible increase in processing time is a small tradeoff when compared to the ability to prove monoclonality and automatically perform onboard assays (SynenTec Bio Services 2015).

5.1.3. Scale-Up

In this section, we investigate manual passaging and automated scale-up via liquid handlers for scaling up the selected clones to larger volumes. Manual passaging involves storing containers of the clonal populations in incubators, while intermittently transferring them to new, larger vessels. This approach has the advantage of being relatively easy to establish with little upfront investment and training. It is also flexible, in that scientists can passage the clones at varying times as their schedules permit. One risk associated with manual passaging is long processing times at scale. Another is non-repeatability between passages; each scientist may perform the process differently, or one scientist may inadvertently alter the process from one passage to another.

Alternatively, clonal populations can be scaled up using automated liquid handling equipment, such as Hamilton's Star line. Such an approach can be semi- or fully-automated. In the semi-automated case, a scientist would remove the vessels from the incubator, place them in the liquid handler, then walk away to perform other tasks while the machine passages the cells. In a fully-automated scenario, a liquid handler can be paired with an incubator, allowing a robotic arm to retrieve and replace vessels at specified intervals. These automated approaches have the advantage of eliminating the hands-on workload for the CLD FTEs, enabling a development operation to operate more efficiently or to scale and process more drug candidates. The risk of such approaches is similar to that of limiting dilution cloning described in Section 5.1.1: the operation would be dependent upon the automated machinery and any maintenance issues could cause delays.

5.1.4. Clone Screen

After the top several clones have been scaled up to larger volumes, they undergo a clone screen to predict their relative performance in a bioreactor environment. Here, we explore a manual procedure for doing so and compare it to an automated approach. The purpose of this phase of the CLD process is to narrow down a select group of clones and choose the optimal candidates for the MCB based on their projected performance in a full-scale bioreactor.

The production performance of the clones can be estimated manually by scaling them up to a larger volume (established by the specific CLD team) and maintaining them in incubators with shake plates. This environment mimics some of the aspects of large-scale bioreactors, e.g., the constant agitation by the impeller. The major advantages are true for many manual processes: low cost and flexibility to operational changes, making it a good candidate for low volume processes. However, it becomes expensive and resource-intensive when CLD work scales up. It also poses a similar non-repeatability risk as that of the scale up process described above in Section 5.1.3.

Automating the clone screen helps address these risks and is achievable using TAP Biosystems' AMBR instrument. This system contains many scaled-down vessels that accurately mimic the environment of a production bioreactor. The vessels can be either 15 mL or 250 mL in volume, depending on the needs of the specific CLD team. Each vessel is effectively a miniature bioreactor, allowing the accurate assessment of many clones with a smaller footprint than more traditional methods. The vessels contain impellers, controlled gas intake regulators, and optical pH and dissolved oxygen sensors. There are also additional modules that can further automate parallel assays. Such a system has the advantage of simultaneously evaluating many clones, enabling higher throughput at scale. It also provides the potential for complete hands-off operation, enabling scientists to perform other tasks, though variability between drug programs could make this difficult to achieve (TAP Biosystems 2015; Zhou and Kantardjieff 2014).

5.1.5. Outsourcing Team Responsibilities

Finally, some of the CLD team's responsibilities that are in addition to new drug program developments can be outsourced. The benefit is clear in that it relieves FTE utilization, but it does introduce some new risks. These include process disruption, as the team now relies on an external provider that can have its own operational delays. Another risk is any additional costs associated with maintaining the integrity of the organization's IP, in addition to the fees for service. These considerations need to be weighed by the specific manufacturer when contracting out its procedures.

5.2. Risk Management

The pharmaceutical industry has recognized the importance of risk management in development, manufacturing, distribution, inspection, submission, and review processes. It has established procedures guiding drug producers to build quality into all aspects of their processes to eliminate risks wherever possible. The general model proposed for risk management, shown below in Figure 21, can be implemented via several tools. These include Failure Mode Effects Analysis (FMEA), Fault Tree Analysis (FTA), and Preliminary Hazard Analysis (PHA) (ICH 2005).

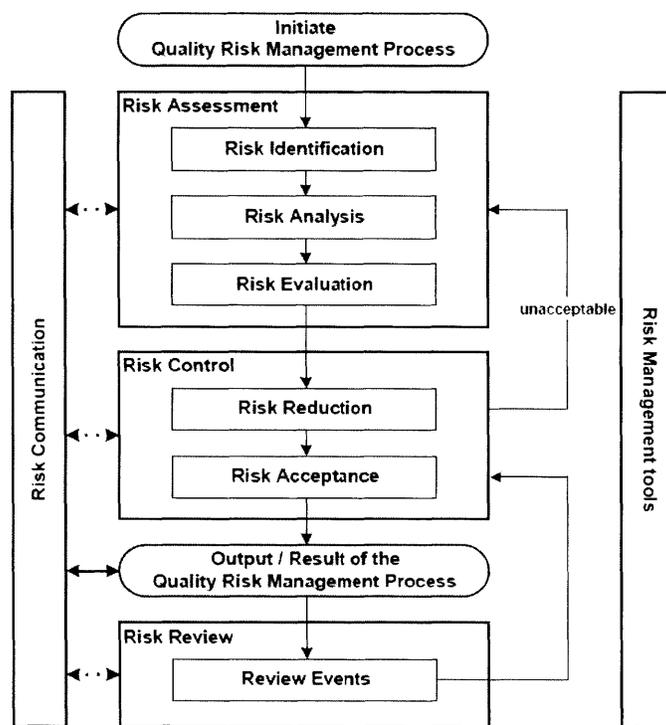


Figure 21: The risk management process established by the International Conference on Harmonisation (ICH 2005).

We have outlined several risks associated with both the current procedures and the potential updates to the CLD process. Many of the new approaches introduce risks related to automation, namely downtime, maintenance costs, and training costs. Others are specific to certain approaches, such as the monoclonality risk of using a colony picking approach in the cloning phase, described above in Section 5.1.1. Though outside the scope of this modeling

effort, an organization that is redesigning their CLD procedures should assess all of these risks before adopting a new approach. The performance and risks of each individual approach will depend upon the specific system design of the drug producer in question. However, it may be infeasible to independently test a new piece of equipment to the extent necessary for a full risk analysis before investing in it. In these cases, developers can use specifications published by vendors or other industry operators to inform their decision before performing their own confirmatory tests and building their risk management approaches. Each developer will then determine how to best address the risks of their updated process. The goal should be to minimize the combination of business process interruptions and risk mitigation activity costs, as illustrated below in Figure 22.

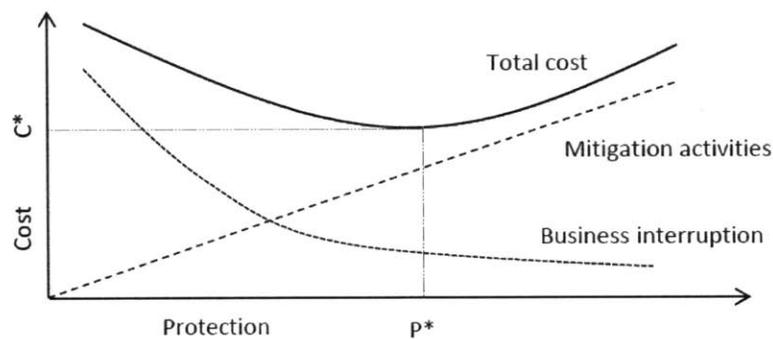


Figure 22: Minimizing the combination of business process interruptions and mitigation activity costs (Kleindorfer and Saad 2005; Linders 2013).

The risks of both the current and new approaches should be characterized to determine the optimal system design for CLD. We believe that an expanded FMEA suits the evaluation of biologic drug development systems, as it has been shown to be successful when assessing automated manufacturing processes (Braaksma et al. 2012). FMEA entails identifying potential failure modes, the effects of their failure on the finished product, and risk reduction approaches to mitigate these failures. In this case, the drug developer would evaluate each phase and machine in the CLD system, the effects of breakdowns or other issues on the drug program and the overall development process, and ways to reduce or eliminate these risks (e.g., preventative maintenance, training programs, or quality checks). This analysis can be extended to a Failure Mode, Effects and Criticality Analysis (FMECA) to quantify the probability,

detectability, and severity of each failure. Doing so allows the organization to align resources with the most critical system nodes (Guebitz, Schnedl, and Khinast 2012).

6. Conclusions and Recommendations

6.1. Conclusions for CLD at Pfizer

The modeling strategy we have outlined shows that significantly improved system designs for CLD can be achieved. Such a tool can be used to evaluate CLD systems anywhere with varying results depending on the current state, though this case study has shown that currently available technologies and procedural changes can achieve a utilization improvement of almost 30% and a cycle time reduction of almost 9%. Additionally, these updates reduce the variation of these metrics by approximately 17% and 14%, respectively. Further augmenting this upgraded system with a nascent technology can, if we make certain assumptions about its performance, increase this performance even further. However, the majority of the gains in this case study are achieved even without adopting this last approach. Available improvements in cycle time are diminishing to zero due to the lifecycle of the host cell being used, but additional future approaches that can further automate the process may add additional capacity to the system. Therefore, we propose that the CLD system at Pfizer be augmented with these approaches that are currently applicable.

We notice that a large portion of the improvement in this case is due to outsourcing some of the CLD team's additional responsibilities, which introduces unique risks. The management should review these risks, along with those of all of the other updates to the system, to ensure that they are sufficiently mitigated. Most of the approaches proposed are adaptations of existing technologies and therefore pose less significant risks to the system. The approaches that are completely new to Pfizer's CLD operation do have certain costs and other risks associated with them, including acquisition and training costs (discussed above in Sections 4.2 and 5). However, the new proposed approaches have been evaluated and there is a high degree of confidence that their benefits significantly outweigh these risks.

The modeling strategy we employ successfully demonstrates its ability to find the critical nodes of this system and evaluate potential ways to address them. Adopting these new approaches improve overall system capacity while also improving development cycle times.

Pfizer can take advantage the updated system design to develop more new drugs simultaneously while also reaching market approval more quickly. This modeling strategy can be applied to additional R&D functions to further improve the end-to-end drug development process. It can also be used in the CLD operations of other biologic drug producers, as well as other R&D, manufacturing, or service environments in any industry.

6.2. Recommendations for Future Initiatives

In addition to investing in the proposed approaches to achieve a more efficient CLD operation, Pfizer should continue to evaluate new technologies with its own independent experiments. The results of such experiments will enable the operational model to remain up to date and able to evaluate approaches as they are identified. Furthermore, a robust data collection system for each step of the CLD process will facilitate the elimination of some of the assumptions made above, most notably, that the processing times for the individual steps are distributed normally. This will provide a more thorough output from the model and, in turn, a greater degree of certainty in future investments. Furthermore, Pfizer should consider exploring ways to achieve similar gains without outsourcing some of the CLD team's functions. Doing so will eliminate such risks as business process interruption, outsourcing costs, and IP loss. Keeping these duties in house may be achievable by cross-training additional resources to perform these responsibilities or by evaluating the external providers for any applicable best practices that can be applied to Pfizer's own systems.

Finally, this operational model or other similar approaches should be employed elsewhere in the R&D organization. The approach explored here is accurate and sufficient for this case study, though processes with broader scopes and greater complexity may warrant investment in software and training to employ more complex modeling efforts. Any team that operates a resource-constrained procedure with constantly advancing technology should objectively evaluate those technologies based on their effects on the operation as a whole before investing in system-wide changes.

References

- Adamo, A., O. Roushdy, R. Dokov, A. Sharei, and K.F. Jensen. 2013. "Microfluidic Jet Injection for Delivering Macromolecules into Cells." *JOURNAL OF MICROMECHANICS AND MICROENGINEERING* 23 (3).
- Bayer Healthcare. 2014. "Small and Large Molecules: Drugs on a Chemical and Biological Basis." <http://www.bayerpharma.com/en/research-and-development/technologies/small-and-large-molecules/index.php>.
- "Biotechnology." 2014. *Encyclopædia Britannica*.
<http://libproxy.mit.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ers&AN=89403862&site=eds-live>.
- Braaksma, A.J.J., A.J. Meesters, W. Klingenberg, and C. Hicks. 2012. "A Quantitative Method for Failure Mode and Effects Analysis." *International Journal of Production Research* 50 (23): 6904–17. doi:10.1080/00207543.2011.632386.
- Bryan, Andrea K., Vivian C. Hecht, Wenjiang Shen, Kristofor Payer, William H. Grover, and Scott R. Manalis. 2014. "Measuring Single Cell Mass, Volume, and Density with Dual Suspended Microchannel Resonators." *Lab on a Chip* 14 (3): 569–76. doi:10.1039/c3lc51022k.
- CDI Bioscience. 2015. "Technical Resource | Clone Selection Using ClonePix FL." Accessed January 18. <http://www.cdibiotech.com/clone-selection.php>.
- De Treville, Suzanne. 2014. "Inventory Models." Lecture presented at the MIT Sloan School of Management, Massachusetts Institute of Technology, February 10.
- Expression Technologies. 2003. "E. Coli Growth Media." <http://www.exptec.com/Expression%20Technologies/Bacteria%20growth%20media.htm>.
- Fazzolare, David A., and Joanna T. Brougher. 2014. "What the Patent Dance of the Biosimilars Act Means for Biosimilars." *Journal of Commercial Biotechnology* 20 (4): 31–36. doi:10.5912/jcb678.
- FDA. 2010. "Biosimilars - Information for Healthcare Professionals (Biosimilars)." WebContent. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/Biosimilars/ucm241719.htm>.
- . 2014. "Number of Biosimilar Investigational New Drug Applications (INDs) Received in the Month." September 30. <http://www.accessdata.fda.gov/FDATrack/track?program=cder&id=CDER-RRDS-Number-of-biosimilar-INDs&fy=all>.
- . 2015. "Drug Development & Approval Process." WebContent. Accessed January 8. [http://www.fda.gov/Drugs/DevelopmentApprovalProcess/ucm079031.htm#What is the difference between patents and exclusivity?](http://www.fda.gov/Drugs/DevelopmentApprovalProcess/ucm079031.htm#What%20is%20the%20difference%20between%20patents%20and%20exclusivity%3F)
- Frost & Sullivan. 2012. "Frost & Sullivan." *Changing Dynamics in the Pharma and Biotech Industry*. June 18. <http://www.frost.com.libproxy.mit.edu/c/10024/sublib/display-report.do?searchQuery=biotech&ctxixpLink=FcmCtx1&ctxixpLabel=FcmCtx2&id=9837-00-41-00-00&bdata=aHR0cDovL3d3dy5mcm9zdC5jb20vc3JjaC9jcm9zcy1jb21tdW5pdHktd2Vhcm>

- NoLmRvP3NIYXJJaFR5cGU9YWRyJnF1ZXJ5VGV4dD1iaW90ZWNoJng9MCZ5PTBAfkBTZW
FyY2ggUmVzdWx0c0B%2BQDE0MjA2Mzk4MjE0NDc%3D.
- Gibson, Jennifer L. 2013. "Pharmaceuticals and Medications Industry." *Salem Press Encyclopedia*.
<http://libproxy.mit.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ers&AN=89088199&site=eds-live>.
- Guebitz, Brigitte, Hubert Schnedl, and Johannes G. Khinast. 2012. "A Risk Management Ontology for Quality-by-Design Based on a New Development Approach According GAMP 5.0." *Expert Systems with Applications* 39 (8): 7291–7301.
doi:10.1016/j.eswa.2012.01.089.
- Guo-giang, Pan, and Feng Ding-zhong. 2013. "A Simulated SIMIO-Based Optimization Approach to a Stochastic Motor Assembly Line." *Industrial Engineering and Management* 18 (2): 65–70.
- Han, Yi, Xing-Mao Liu, Hong Liu, Shi-Chong Li, Ben-Chuan Wu, Ling-Ling Ye, Qu-Wei Wang, and Zhao-Lie Chen. 2006. "Cultivation of Recombinant Chinese Hamster Ovary Cells Grown as Suspended Aggregates in Stirred Vessels." *Journal of Bioscience and Bioengineering* 102 (5): 430–35. doi:10.1263/jbb.102.430.
- Hay, Michael, David Thomas, John Craighead, Celia Economides, and Jesse Rosenthal. 2014. "Clinical Development Success Rates for Investigational Drugs." *Nature Biotechnology* 32 (1): 40–51.
- Heyman, Jeffrey B. (Jeffrey Brian). 2010. "Simulation Modeling to Predict Drug Pipeline Throughput in Early Pharmaceutical R&D." Thesis, Massachusetts Institute of Technology. <http://dspace.mit.edu/handle/1721.1/59167>.
- ICH. 2005. "ICH Harmonised Tripartite Guideline - Quality Risk Management Q9, Step 4 Version." International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.
http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q9/Step4/Q9_Guideline.pdf.
- IMS Health. 2012. "Featured Insight, Global Report - Generics, Biosimilars." July 23.
———. 2013. "IMS Health: Top-Line Market Data."
<http://www.imshealth.com/portal/site/imshealth/menuitem.18c196991f79283fddc0ddc01ad8c22a/?vgnnextoid=6521e590cb4dc310VgnVCM100000a48d2ca2RCRD&vgnnextfmt=default>.
- Institute of Mergers, Acquisitions and Alliances (IMAA). 2014. "Statistics on Mergers & Acquisitions." November 17. http://www.ima-institute.org/statistics-mergers-acquisitions.html#M&A_Ind_Biotechnology&Pharmaceuticals.
- Invitrogen. 2007. "FreeStyle CHO-S Cells." March 28.
<http://www.lifetechnologies.com/search/global/searchAction.action?query=cho+doubling+time&resultPage=1&resultsPerPage=15&autocomplete=>
- Karcher, Susan J. 2013. "DNA: Recombinant Technology." *Salem Press Encyclopedia of Science*.
<http://libproxy.mit.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ers&AN=89551674&site=eds-live>.
- Keil, Gregory. 2014. "High-Throughput Automation Solutions in Bioprocess Development." presented at the Bioprocessing Summit, August 14.

- Kimmerling, Robert, Gregory Szeto, Nigel Chou, Mark Stevens, Sungmin Son, Darrell Irvine, and Scott Manalis. 2014. "A Toolset for Multigenerational Lineage Characterization of Single CD8+ T Cells." June 10.
- Kleindorfer, Paul R., and Germaine H. Saad. 2005. "Managing Disruption Risks in Supply Chains." *Production & Operations Management* 14 (1): 53–68.
- Lai, Tingfeng, Yuansheng Yang, and Say Kong Ng. 2013. "Advances in Mammalian Cell Line Development Technologies for Recombinant Protein Production." *Pharmaceuticals* 6 (6): 579–603. doi:10.3390/ph6050579.
- Lanza, Amanda M, Joseph K Cheng, and Hal S Alper. 2012. "Emerging Synthetic Biology Tools for Engineering Mammalian Cell Systems and Expediting Cell Line Development." *Current Opinion in Chemical Engineering* 1 (November): 403–10. doi:10.1016/j.coche.2012.09.005.
- Linders, David (David Robert). 2013. "Development of a Risk Management System for Consumables Used in Biopharmaceutical Manufacturing." Thesis, Massachusetts Institute of Technology. <http://dspace.mit.edu/handle/1721.1/81723>.
- Lin, Jr-Lung, Shih-Siou Wang, Min-Hsien Wu, and Chih-Chin Oh-Yang. 2011. "Development of an Integrated Microfluidic Perfusion Cell Culture System for Real-Time Microscopic Observation of Biological Cells." *Sensors (14248220)* 11 (9): 8395–8411. doi:10.3390/s110908395.
- Macey, Marion G. 2007. *Flow Cytometry: Principles and Applications*. Totowa, N.J. : Humana Press, c2007.
- Molecular Devices. 2011. "Clone Select Imager Operator Manual." <http://www.moleculardevices.com/systems/cell-screening-and-bioproduction/cloneselect-imager>.
- Ossipow, Vincent, and Nicolas Fischer. 2014. *Monoclonal Antibodies : Methods and Protocols*. Methods in Molecular Biology: 1131. New York : Humana Press, 2014.
- Pfizer, Inc. 2014. "Pfizer Financial Statement 2013." http://www.pfizer.com/investors/financial_reports/financial_reports.
- Phillips, Jocelyn. 2014. *Global Biotechnology*. IBISWorld Industry Report L6724-GL. Melbourne, Australia: IBISWorld Services. <http://clients1.ibisworld.com.libproxy.mit.edu/reports/gl/industry/default.aspx?entid=2010>.
- PMLiVE. 2014. "PMLiVE Top Pharma List." January 13. http://www.pmlive.com/top_pharma_list/pharma_companies/pfizer#infographic.
- Schulze, Ulrik, and Michael Ringel. 2013. "What Matters Most in Commercial Success: First-in-Class or Best-in-Class?" *Nature Reviews Drug Discovery* 12 (6): 419–20. doi:10.1038/nrd4035.
- Science Gateway. 2015. "Cell Biology Protocols." Accessed January 18. <http://www.sciencegateway.org/protocols/cellbio/cell/cbld.htm>.
- Sharei, A., J. Zoldan, A. Adamo, Woo Young Sim, Nahyun Cho, E. Jackson, S. Mao, et al. 2013. "A Vector-Free Microfluidic Platform for Intracellular Delivery." *Proceedings of the National Academy of Sciences of the United States of America* 110 (6): 2082–87. doi:10.1073/pnas.1218705110.

- Sharfstein, Susan T. 2008. "Advances in Cell Culture Process Development: Tools and Techniques for Improving Cell Line Development and Process Optimization." *BIOTECHNOLOGY PROGRESS* 24 (3): 727–34.
- Silver, Steven. 2013. "Industry Surveys: Biotechnology." Standard & Poor's.
- Sleiman, Robert J, Peter P Gray, Martin N McCall, Joe Codamo, and Noelle-Ann S Sunstrom. 2008. "Accelerated Cell Line Development Using Two-Color Fluorescence Activated Cell Sorting to Select Highly Expressing Antibody-Producing Clones." *Biotechnology And Bioengineering* 99 (3): 578–87.
- Solentim. 2015. "Cell Metric™ CLD." Accessed January 22. <http://www.solentim.com/products/cell-metric-cld/>.
- Stilling, Glenn Ellen Starr, M.A., M.L.S. 2013. "Clinical Trials." *Magill's Medical Guide (Online Edition)*. <http://libproxy.mit.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ers&AN=87690477&site=eds-live>.
- Sun, Tao, Joseph Kovac, and Joel Voldman. 2014. "Image-Based Single-Cell Sorting via Dual-Photopolymerized Microwell Arrays." *Analytical Chemistry* 86 (2): 977–81. doi:10.1021/ac403777g.
- SynenTec Bio Services. 2015. "Cellavista." Accessed January 22. <http://www.synentec-bioservices.com/24/>.
- TAP Biosystems. 2015. "ambr15 - System Overview." Accessed January 21. http://www.tapbiosystems.com/tap/cell_culture/ambr.htm#.
- Turk, Sarah. 2014. *Global Pharmaceuticals & Medicine Manufacturing*. IBISWorld Industry Report C1933-GL. Melbourne, Australia: IBISWorld Services. <http://clients1.ibisworld.com/reports/gl/industry/default.aspx?entid=720>.
- UAH. 2015. "The Pareto Distribution." January 15. <http://www.math.uah.edu/stat/special/Pareto.html>.
- Yang, Chenxi, Li Luo, Qunrong Ren, Shijun Tang, and Renrong Gong. 2013. "Using SIMIO for Laparoscopic cholecystectomy (LC) Surgery Simulation." Hong Kong, China. doi:10.1109/ICSSM.2013.6602576.
- Zhou, Weichang, and Anne Kantardjieff. 2014. *Mammalian Cell Cultures for Biologics Manufacturing*. Advances in Biochemical Engineering/Biotechnology: 139. Heidelberg : Springer, 2014.