Planning Report 07-1

Economic Analysis of the Technology Infrastructure Needs of the U.S. Biopharmaceutical Industry

NIST

**National Institute of
Standards and Technology**
U.S. Department of Commerce

Economic Analysis of the Technology Infrastructure Needs of the U.S. Biopharmaceutical Industry

Final Report

Prepared by

Michael Gallaher
Jeffrey Petrusa
Alan O'Connor
Stephanie Houghton
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709-2194

Economic Analysis of the Technology Infrastructure Needs of the U.S. Biopharmaceutical Industry

Final Report

November 2007

Prepared by

Michael Gallaher
Jeffrey Petrusa
Alan O'Connor
Stephanie Houghton
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709-2194

Contents

Figures

Tables

viii

Executive Summary

The purpose of this report is to quantify the biopharmaceutical industry's current expenditures on the technology infrastructure; identify key infrastructural bottlenecks across the biopharmaceutical development cycle; and estimate the drug-development, production, and surveillance cost savings an improved technology infrastructure may provide.

The biopharmaceutical industry's origins are rooted in research by Stanley Cohen and Herbert Boyer that led to a process called recombinant DNA, in which pieces of DNA from cultured cells are artificially combined in a controlled setting. Recombinant DNA was a novel technique that enabled large-scale production of biopharmaceutical products like human insulin. From these origins, the industry has grown into one that currently spends $21 billion on research and development (R&D) annually and has commercialized over 400 products.

Traditionally, drug companies have developed new products through the trial-and-error investigative processes of medicinal chemistry. They created small chemical compounds—referred to as small-molecule drugs—developed through the analysis of symptoms that characterize certain illnesses and diseases. Although small-molecule drugs still account for a significant share of the drug industry's new product pipeline, they no longer dominate the focus of new drug development.

Over the last two decades, there has been a shift in the drug development industry to larger, more complex molecular compounds that target biochemical mechanisms instead of symptoms. These "large-molecule" biopharmaceuticals take advantage of how human biological systems function. Biotechnology researchers design drugs that capitalize

on the specific, precise, and predictable attributes of subcellular molecules, like DNA and proteins, across a diverse set of species and cell types.

The dramatic growth in the number of biopharmaceuticals in clinical development can be attributed to exponential growth in our scientific understanding of the biological systems associated with human health and disease. As researchers begin to identify the biological pathways to disease, drug development research will continue to shift focus away from traditional pharmaceuticals toward complex biopharmaceuticals. The industry's ability to enhance innovation will largely be determined by several factors, including the efficiency of its R&D, quality control and assurance programs, and postmarket surveillance activities. The technology infrastructure supporting biopharmaceutical drug development—the core data, methods, and standards—must be sufficient to enable effective R&D. Information must be able to be communicated among data sets, technology platforms, and organizations. This implies that the protocols, descriptors, and assays used to acquire these data be transparent and standardized within and across organizations. Researchers must have confidence in and assurance of the accuracy and precision of the data and processes informing drug development decisions.

The purpose of this study is to investigate current expenditures and potential future efficiency gains associated with developing the technology infrastructure to inform public policy and support strategic planning at the National Institute of Standards and Technology (NIST). Study objectives were the following:

- to estimate the biopharmaceutical industry's annual spending on technology infrastructure–related investments and activities;

- to assess key areas of technology infrastructure in which industry has underinvested because of market and technical barriers or because of unattractive risk-reward ratios at the firm level; and

- to characterize the potential efficiency gains an improved technology infrastructure holds for the industry, as represented by potential drug development cost reductions, shorter time to market for new drugs, and greater probabilities of developing successful products approved by the Food and Drug Administration (FDA).

This report's scope is limited to the biopharmaceuticals industry because the diversity of and varying definitions of biotechnology make it difficult to

study as a single industry and hence at the level of clarity and specificity required by economic analyses.

ES.2 THE BIOPHARMACEUTICAL INDUSTRY'S TECHNOLOGY INFRASTRUCTURE

As the term "infrastructure" implies, the technology infrastructure of an industry refers to the tools, methods, and data that enable or support R&D, products, and services. These tools, methods, and data are considered infrastructural because they are not necessarily commercial products themselves; rather, they support or embody processes and components that make many advanced technology products and services possible.

Many elements of the technology infrastructure are unseen or taken as a given because they are deeply embedded in or underlie research methodologies and instruments. For example, techniques that control process quality or verify the accurate calibration of laboratory instruments are part of the technical infrastructure, as are standardized reference materials and data that researchers use to increase their confidence and assurance of the accuracy and precision of their work. More visible components of this infrastructure include analytical instruments and advanced software systems and algorithms.

ES.2.1 The Economic Role of the Technology Infrastructure

While laboratory instruments, computational systems, and advanced chemistries represent vital technology infrastructure employed in the direct application of science to drug discovery, production, and postmarket surveillance, a universe of unseen technologies support the application of those systems. The technology infrastructure is equally composed of the protocols, tests, and other methodologies devised to improve R&D efficiency and effectiveness. These infrastructure components are often overlooked. However, inadequate technologies such as calibrants and standard operating protocols—all of which researchers rely on everyday—have significant, cumulative economic impacts on overall industry productivity.

Gaps in the technology infrastructure are most readily apparent to researchers when they hamper productivity and collaboration and thus manifest obstacles to the development of new drugs.

Current activities to support the technology infrastructure consist of expenses incurred through internal activities; external purchases; and participation in consortia, partnerships, or other research activities:

- the purchase of measurement-related equipment, software, reference materials, or services;

- activities required for setup and validation of analytical instruments, reagents, or other research tools; examples of these activities may include developing calibration test methodologies, standard operating procedures, or process standards in measurement and manufacturing practices;

- in-house customization of technology platforms purchased from third-party vendors;

- efforts to develop interoperability between different software or equipment systems; and

- license fees or any other spending on enhancements to routinely used processes or equipment intended to increase productivity, reduce redundancy, or improve the confidence in results.

The methods, techniques, and data discussed above form a complex technology infrastructure that enables productivity and efficiency in each of the major stages of drug development and provision related to biopharmaceuticals (such as basic and applied research, clinical trials, commercial manufacturing, and market development). Improvements in this infrastructure can have numerous potential economic impacts, including the following:

- Cost reductions
    - Lower labor and materials costs for discovery, development, and production of a given therapy
    - Lower transaction costs associated with marketing new products and meeting postmarket tracking/assessment requirements

- Accelerated time to market
    - Shorter time between discovery and FDA approval

- Quality improvements
    - Detection of drug failures in earlier clinical trial phases
    - Investigational New Drugs (INDs) have greater probability of receiving FDA approval
    - Reduced uncertainty in drug efficacy and safety in the population targeted for prescription, longer shelf lives, and less restrictive storage and handling needs

ES.2.2 The Biopharmaceutical Product Development Process

For the pharmaceutical industry as a whole, only 1 in roughly 10,000 compounds screened during the drug discovery stage succeeds through the clinical trial stage and results in an FDA-approved drug (PhRMA, 2005). Of drugs entering clinical trials, FDA reports only about 8% ever reach the market to recoup R&D expenditures. Furthermore, the failure rate in the last and most expensive Phase III stage is nearly 50% (FDA, 2006). Revenue from successful products must also cover the costs from failed efforts as companies seek an overall adequate rate of return on investment.

Recent studies have suggested that biopharmaceutical development costs are higher in preclinical stages and the drugs may be more difficult to manufacture, but that clinical trials require fewer patients and have higher success rates (see Reichert [2004]). Compared with a traditional success rate of around 20%, the Tufts University research group reports 30% to 35% success rates for biopharmaceuticals that entered clinical trials between 1990 and 1997 (Tufts, 2005).

FDA approved the first biopharmaceutical for marketing in 1982. As of 2006, the biopharmaceutical industry accounted for over 260 biotechnology-based therapeutics and vaccines approved for over 380 indications (BIO, 2006). The length of time from initial discovery through approval for most biopharmaceuticals is 8 to 15 years, with most costs incurred during clinical trials (BIO, 2004).

The specific R&D and manufacturing processes differ among different categories of biopharmaceutical products. However, all these drugs start with a foundation in basic research and pass through the following major stages:[1]

- drug discovery,
- preclinical development and testing,
- clinical trials,
- scale-up and commercial manufacturing, and
- postmarket surveillance.

In a recent study, DiMasi (2002) explored the effect that shortening development times could have on pharmaceutical drug R&D costs. His

---

[1]These five stages are presented in chronological order, with the exception that scale-up to commercial manufacturing is undertaken simultaneously with preclinical development and clinical trials.

results suggest that uniform reductions clinical trial stages of 25% would lower the future-value total cost per approved drug by 16%. DiMasi found that an increase in the overall success rate for compounds entering clinical trials from 21.5% to 33.3% would lower the present value cost per approved drug by up to 30.2% (DiMasi, 2002).

ES.2.3 Technology Focus Areas for this Study

This study investigated the expenditures and potential benefits from an improved technology infrastructure in six specific technology fields, termed technology focus areas (TFAs). RTI worked with technical experts to identify potential infrastructure improvements that could be feasibly achieved within the next 10 years. RTI then estimated potential efficiency gains that achieving these improvements would convey to the biopharmaceutical industry.

To assess potential efficiency gains for discovery, preclinical, and clinical (Phase I through Phase III) activities, four TFAs were selected:

1. **Enhanced bioimaging techniques.** Bioimaging refers to the use of imaging technologies (such as CT, MRI, and ultrasound) to assess biological structures and phenomena. Imaging studies are increasingly required in regulatory and research reviews of new drugs.

2. **Standards and metrology in gene and protein expression analysis.** Gene and protein expression analyses seek to acquire information on whether certain genes and proteins are active or in-active in biological processes. Expression analysis is essential to the screening and analysis of potential drugs and their targets.

3. **Improved bioinformatics and *in silico* predictive modeling.** Bioinformatics is a synonym for computational biology: the application of advanced computing to create and analyze vast databases of biological information. *In silico* modeling refers to using complex computer simulations of organisms to predict how a given compound may affect the organism.

4. **Identification and validation of molecular biomarkers.** The term "biomarkers" refers to any biological measurement that yields information on disease progression, pharmacology, or safety. Biomarkers can be used as the basis for decision making in drug development. Cholesterol and blood pressure are two traditional biomarkers. Technology advances in areas such as bioimaging and protein detection are expanding the number of useful biomarkers. Researchers hope to use biomarkers to shorten clinical trials, increase overall research efficiency, and provide improved diagnostic techniques. Before the benefits of novel biomarkers can be realized, researchers must overcome significant challenges related to validation methods of new biomarkers and regulatory acceptance of drug study results that are based on biomarker measurements.

These infratechnologies have applications throughout the drug development process and cut across a number of different therapeutic categories. In addition to the four drug-development TFAs, two process-based TFAs were included:

5. Infratechnologies to enhance **scale-up and commercial manufacturing**, including improvements in upstream and downstream processing and process monitoring/quality assurance.

6. Infratechnologies to support **postmarket surveillance** activities, including product surveillance, tracking, and Phase IV clinical trials.

ES.3   ANALYSIS METHODOLOGY AND RESULTS

The economic analysis has two objectives: (1) estimate current private-sector spending on the biopharmaceutical technology infrastructure and (2) calculate the savings an improved technology infrastructure could yield the industry.

The first objective entailed surveying the industry to capture spending and then extrapolating the results to estimate industry-level expenditures. To meet the second objective, RTI leveraged recent drug development cost studies and built an economic model that recalculated the cost of bringing a new drug to market under a future scenario with an enhanced technology infrastructure. The model included both the cost of failed INDs and the time value of money. Experts estimated how costs, stage length, and the probability of success would improve given the feasible infrastructure improvements RTI identified.

Industry representatives who participated in the study represented firms whose combined annual R&D spending accounted for 42% of total industry R&D spending[2] and 49% of annual product sales.[3]

---

[2]RTI calculated market share for the technologies applied in the R&D phases by dividing company-reported R&D expenditures in 2004 by the total R&D expenditures for all publicly traded biotechnology firms. Market share for the manufacturing and postmarket segment was calculated by dividing net sales of biopharmaceuticals of the individual firm by the total sales for all commercialized biopharmaceuticals using data for 2004.

[3]Responses to the Internet survey were anonymous and hence could not be linked to R&D expenditures or sales. Thus, Internet respondents are not represented in the market share figures. As a result, the market share of firms participating in this study is greater than reported here. For example, the 12 commercial manufacturing firms that completed the Internet survey produce 30 of the 264 domestically approved biopharmaceutical drugs.

ES.3.1 Estimating Current Technology Infrastructure Expenditures by the Biopharmaceutical Industry

Current private-sector expenditures on technology infrastructure were estimated by surveying biopharmaceutical drug developers and their supply chain, including service vendors, consultants, and academics who support them. Research directors and managers at biopharmaceutical companies estimated the relative proportions of labor, capital, and materials expenses comprising a category of infrastructure expenditure and the distribution of those costs among the TFAs that RTI identified. They also discussed the timing of costs within the product development stage.

Aggregated responses by stage were scaled to national expenditures using activity measures relevant for each drug development. Study participants provided several measures for comparing and aggregating their data with other participants. Measures used to aggregate responses included

- the number of scientists and engineers comprising the business unit's research staff,
- the number of FDA-approved biopharmaceutical products, and
- the percentage of sales corresponding to the unit for which they are responding.

RTI estimates that the biopharmaceutical industry currently spends $1,219 million annually on technology infrastructure-related products and services, including $884 million in support of drug R&D-related activities, and $335 million to support commercial manufacturing and postmarket surveillance activities. These estimates represent current biopharmaceutical industry expenditures.

In reviewing the data in Table ES-1, we can make a number of observations concerning infrastructure spending in the R&D segment:

- Gene expression systems and biomarkers accounted for over half of total technology infrastructure spending in the R&D segment at 30% and 24%, respectively.
- Bioimaging accounts for 15% and informatics for 22%.
- The remaining 8% of the $884 million is distributed among all other technology areas in the R&D segment.[4]

---

[4]Some respondents were reluctant to classify infrastructure expenditures in the TFA categories provided and hence classified them as "other."

Table ES-1. Annual Technology Infrastructure Expenditures by Technology Focus Area, 2005

Technology Focus Area	Annual Technology Infrastructure Spending (millions)	Percentage Distribution	Relative Spending
Bioimaging	$136	15%	$4,011 per scientist
Biomarkers	$212	24%	$6,240 per scientist
Bioinformatics	$198	22%	$5,813 per scientist
Gene expression analysis	$265	30%	$7,800 per scientist
Other	$73	8%	$2,136 per scientist
Subtotal of R&D Activities	**$884**	100%	
Commercial manufacturing	$162	48%	$613,000 per approved drug
Postmarket surveillance	$173	52%	$656,000 per approved drug
Subtotal of Commercial Activities	**$335**	**100%**	
Industry Total	**$1,219**		

Source: RTI estimates.

The commercial segment expenditures were $335 million, broken out as follows:

- Commercial manufacturing accounted for 48% of total infrastructure expenditures in the commercial segment.

- Postmarket surveillance accounted for 52% of total infrastructure expenditures in the commercial segment.

ES.3.2 Estimating Potential R&D Efficiency Gains from an Improved Technology Infrastructure

Rather than ask the industry how expenditures could have been different in 2005, this study took the approach of evaluating how a specific set of improvements could increase efficiency going forward. The same experts who provided data to estimate current infrastructure-related spending quantified the impact a series of feasible infrastructure improvements would have on R&D, manufacturing and post-FDA approval activities. They offered their views on how specific improvements to the technology infrastructure could

- lower the development cost of the average biopharmaceutical drug,

- increase the probability the drug would be approved by FDA by enhancing data quality and analytical methods,

- shorten the drug's time to market;

- lower the ongoing costs for manufacturing that drug and improve manufacturing tolerances, and

- make the postmarket surveillance infrastructure more efficient and responsive.

The metric most commonly used for assessing industry's spending is the "R&D cost per approved drug."[5] The economic model recalculates average R&D cost per approved drug using the baseline costs from a 2007 study by Joseph DiMasi and Henry Grabowski. Baseline estimates included biopharmaceutical drug development costs, times, and the probability a candidate drug moved from one R&D stage to the next.

The cost model offered two output measures:

- the change in actual R&D expenditures, per investigational new drug (IND) and per FDA-approved drug, and

- the change in the present value of these R&D expenditures, including and excluding the cost of failures.

The time-value of money concept—where the present value of $1 invested 10 years ago is worth more than the same nominal $1 invested today—takes into account investment options, inflation, and other time-based factors affecting the value of money. Compressing the schedule of the drug development cycle greatly affects the true cost of drug development. Thus, the model was developed to show both changes in actual and present-value R&D expenditures.

Table ES-2 presents the overall changes in success and failure rates across clinical trial phases predicted by survey respondents who were asked to assess how the typical distribution of failed INDs would change under the improved infrastructure scenario. Identifying poor drug candidates (as measured by the probability of FDA approval) sooner in the pipeline has the most significant impact on costs.

The results in Table ES-2 are in percentage terms because it is not possible to predict the absolute number of INDs that would enter clinical trials under the improved infrastructure scenarios. In terms of the distribution of INDs that fail during clinical trials, the percentage of

---

[5]RTI followed this approach, rather than looking at annual R&D costs at the firm level, for multiple reasons. First, and most importantly, the models for R&D cost per approved drug build in the cost of failure—a very important component of the industry's total R&D spending and a major focus of efforts to streamline drug development. Second, private firms may respond to reductions in drug failure rates, not by lowering their R&D spending but by producing more drugs. If this is the case, total industry R&D spending would remain unchanged, but the R&D cost per approved drug would decrease.

Table ES-2. Estimated Potential New-Drug Approval and Failure Rates

Technology Focus Area	IND Approval Probability	For INDs Failing in Clinical Trials, Percentage Failing by Phase[a]				Probability of Recall
		Phase I	Phase II	Phase III	Total	
Baseline	*30.2%*	*23.4%*	*52.4%*	*24.2%*	*100.0%*	*0.40%*
Individual Scenarios						
Biomarkers	41.0%	39.2%	37.0%	23.8%	100.0%	0.30%
Bioinformatics	40.0%	30.0%	40.5%	29.5%	100.0%	0.30%
Gene expression	45.0%	37.5%	35.5%	27.0%	100.0%	0.10%
Combined Scenarios						
Lower bound	40%	30%	41%	30%	100%	0.30%
Upper bound	45%	39%	37%	24%	100%	0.10%

[a]If 100 INDs enter clinical trials, the IND approval probability suggests that ~30 INDs will eventually receive FDA approval. The remaining 70 INDs will then fail in one of the three clinical trial phases. During Phase I, ~16 INDs or 23% of the 70 would fail. In Phase II, an additional 37 INDs or 53% of 70 are likely to fail. The remaining 17 or 24% of the 70 IND failures would occur during Phase III.

Note: Comparable data for bioimaging could not be calculated; thus, bioimaging was excluded from this table. RTI estimates based on DiMasi and Grabowski (2007).

failures shifted toward earlier failures. That is, a greater proportion of failures would occur in Phase I, which in turn leads to a lower rate of failures in Phase II where efficacy is evaluated for the first time. The proportion that fail in Phase III remains relatively constant or increases slightly in the improved infrastructure scenarios relative to the baseline. Such muted variation in Phase III failures between scenarios may reflect the limitations of our current scientific knowledge base of the interactions between chemical compounds and the human body. An increased understanding of biological systems and processes would likely have a greater impact on Phase III failures than an improved technology infrastructure.

The baseline expected R&D expense for an approved biopharmaceutical is $560 million; however, when this figure is adjusted to account for the time value of money and firms' opportunity costs (i.e., the true present-value expense), the expected R&D expense is $1,241 million.

Tables ES-3 and ES-4 illustrate how improvements in the technology infrastructure for specific TFAs could affect the cost, development time, and success rates for a new drug under several scenarios. These scenarios include one for each TFA individually and two scenarios with combined results from all TFAs. The most and least optimistic gains in

Table ES-3. Potential R&D Cost Reductions in Biopharmaceutical Development with an Improved Technology Infrastructure

Technology Focus Area	Estimated *Actual* Cost per Approved Drug (millions)	Percentage Change from Baseline	Estimated *Present-Value* Cost per Approved Drug (millions)	Percentage Change from Baseline	Development Time (months)
Baseline	*$559.6*	—	*$1,240.9*	—	*133.7*
Individual Scenarios					
Bioimaging	—	—	—	—	—
Biomarkers	$347.9	–38%	$676.9	–45%	108.2
Bioinformatics	$375.0	–33%	$746.3	–40%	116.6
Gene expression	$345.8	–38%	$676.0	–45%	111.9
Combined Scenarios					
Lower bound	$421.2	–25	$869.6	–30	122.4
Upper bound	$289.2	–48	$533.1	–57	98.1

Note: See Table 5-4 for estimated changes in FDA approval, distribution of IND failures within clinical trials, and probability of a recall of an FDA-approved drug. The period between completion of Phase III clinical trials and FDA approval is assumed to be 16 months in present-value calculations. Source: RTI estimates based on DiMasi and Grabowski (2007).

Table ES-4. Potential Manufacturing Efficiency Gains from an Improved Technology Infrastructure

Phase/Activity Cost	Baseline Production Costs		Potential Change in Cost by Phase/Activity		
	Percentage of Total[a]	Baseline Total (millions)	Percentage Change	Change in Cost (millions)	Costs under an Improved Infrastructure (millions)
Preproduction	30%	$1,900	–29%	–$551	$1,349
Upstream processing	20%	$1,267	–18%	–$228	$1,035
Downstream processing	40%	$2,533	–22%	–$557	$1,976
Process monitoring and quality assurance testing	10%	$633	–23%	–$146	$491
Total commercial manufacturing costs		**$6,333**		**–$1,482**	**$4,851**

[a]From Frost and Sullivan (2004).

Source: RTI estimates.

time, cost, and approval rates were collected from each TFA to estimate this range.

The counterfactual technology infrastructure could potentially

- increase the probability of FDA approval for the average IND from 30.2% to 40%;

- reduce the expected direct R&D cost for a new FDA-approved drug to between $289 million and $421 million, which is 25% to 48% less than the baseline cost estimated by DiMasi and Grabowski (2007); and

- reduce the average time to take a candidate from discovery through Phase III clinical trials from 11 years to 10 years under the lower bound scenario and to slightly more than 8 years under the upper bound scenario.

Reducing R&D expenses while simultaneously increasing the likelihood of FDA approval generates significant savings. Given that R&D expenses for an approved drug can accrue over a period of 12 years, changing both the amount and timing of expenses can have a significant impact on benefits calculations (see Figure ES-1).

Figure ES-1. Baseline and Counterfactual Estimated Cost per Approved Drug

C_0 Baseline expected R&D costs per approved drug

C_1 Expected R&D costs per approved drug with an improved technology infrastructure
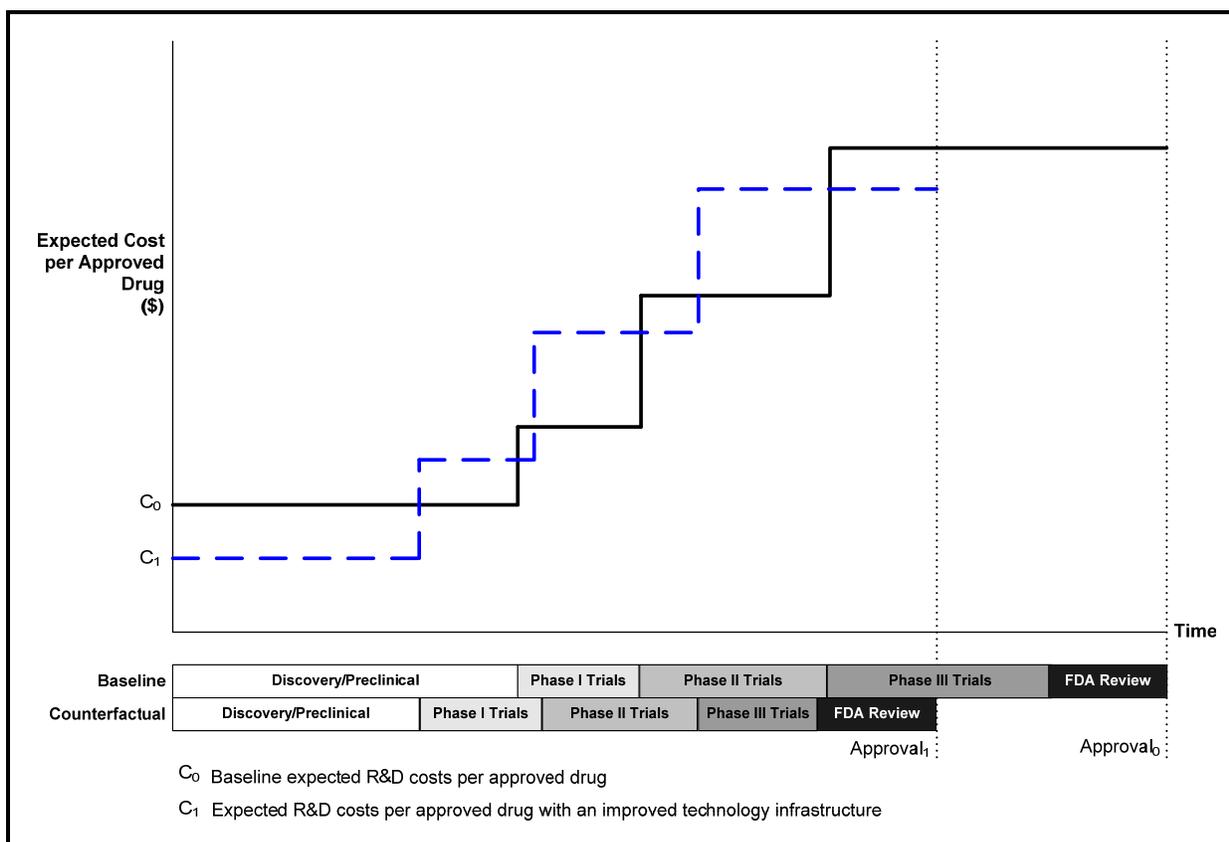
A 50% improvement is significant and would require both substantial and broad-based advances in a range of technical infrastructures, particularly gene expression analysis and biomarkers. However, these results are consistent with a 2004 FDA report in which one expert suggested that biomarkers could reduce the cost of developing a new drug by 50% (FDA, 2004b, p. 19).

ES.3.3 Estimating Potential Manufacturing Efficiency Gains from an Improved Technology Infrastructure

Whereas benefits related to drug development activities from an improved infrastructure were measured relative to the average cost per IND and per FDA-approved drug, for manufacturing and postmonitoring activities, the analysis estimated efficiency gains relative to *total* industry manufacturing expenses. The estimated efficiency gains in this report for manufacturing and R&D are therefore not additive. This is because information on average or "typical" product costs was not publicly available, requiring RTI to measure potential gains relative to estimated total industry manufacturing costs.

Study participants provided estimates of how an improved infrastructure would reduce manufacturing costs by phase. They estimated a 29% reduction in preproduction costs and a 22% reduction in downstream processing costs, among other impacts, that would result from the improved infrastructure (see Table ES-4).

ES.4 CONCLUDING REMARKS

Implicit in this study's findings is that each biopharmaceutical company's technology infrastructure varies. The relative sophistication of any one firm's infrastructure is a function of the amount of intellectual capital it has invested as much as its acquired resources such as instruments and software. Differences across firms result from different strategies and abilities for overcoming individual technical barriers in the absence of well-coordinated industry technical assistance and standardization. In addition, these differences go beyond R&D stages and are manifest in varying quality assurance and control programs (QC/QA) as well as adverse event reporting.

Most experts interviewed over the course of this study conceptualized technology infrastructure expenditures into two general categories. The first are expenditures that constitute an investment in current and future R&D efficiency. The labor effort, systems, and instruments expended are

essential, unavoidable, and integral to a firm's primary economic activity. The second category are expenditures that represent costs incurred to develop workarounds and overcome technical barriers stemming from a pervasive lack of industry-wide technical assistance and standardization.

Challenges in the biopharmaceutical technology infrastructure and the differences in how organizations respond to these challenges are rooted in

- the historical trial-and-error approach to drug discovery and development;
- development times averaging 12 years, which are compounded by changes in regulatory requirements, information systems, and procedures;
- variability in methodologies and protocols that acquire information and variations in how that information is described and characterized;
- few industry standards for ontologies, data formats, and data communications systems;
- the rapid introduction and adoption of data acquisition technologies (which far outpace the development of industry's ability to manage, communicate, analyze, and synthesize data); and
- changes in the regulatory environment in a diverse set of countries and foreign languages (see Table ES-5 for stakeholders' recommendations).

Interviewees and survey respondents stated that NIST participation in standardization activities for biopharmaceuticals would be welcome. These experts believe that NIST has a natural role, given its status as an independent, neutral body, its greater access to consistent funding, and its mission to provide measurement-intensive and other technical infrastructure to industry. Companies' reporting requirements to FDA add a regulatory driver to nearly all research initiatives. Thus, senior scientists and directors at biopharmaceutical companies cited FDA's regulatory authority as a potential constraint and suggested that NIST and FDA collaborate to identify and develop standards that are congruent with the Critical Path initiative.

Although the entire biopharmaceutical industry would benefit from an improved infrastructure, emerging companies that have yet to adopt or develop an internal infrastructure stand to gain the most. Such small start-up firms should have greater chances for success with access to a deeper and more efficient technical infrastructure. from the resulting overall increase in operating efficiency should enable these firms to

Table ES-5. Stakeholders' Comments on Technology Infrastructure Needs

Technology Focus Area	Within the TFAs Studied in this Report, Company Representatives, Academics, and Government Researchers Recommended that Needed Improvements to the Technology Infrastructure:
Bioimaging	• Include consistent taxonomies for medical and anatomical regions of observations • Standardize image labeling procedures and ontologies • Develop formats for exchanging imaging data among data systems • Improve the image archival, retrieval, and management infrastructure • Improve access to imaging technology, including image capture and interpretation systems
Biomarkers	• Address the need for greater sensitivity in detection of protein expression levels • Develop traceable standards for currently known immunoassayed biomarkers • Standardize existing protocols for generating gene expression results • Develop standardized methods and tools to hasten validation of technology platforms • Standardize statistical methodologies for data analysis in biomarker validation studies
Bioinformatics	• Improve data visualization and analysis techniques, • Develop common (neutral) data formats and analysis tools • Set standard protocols for investigating and transforming data • Standardize ontologies for characterizing data
Gene Expression	• Make available reference materials that mimic the biological complexity of tissue and blood samples • Establish sample quality standards, including tools to evaluate the extent to which samples may have degraded • Create sample acquisition, handling, and preparation techniques given the amount of time samples may spend in transit between research sites • Establish systems, data, and analysis mechanisms to benchmark microarray performance • Develop calibration tools and techniques for scanning equipment • Provide standard calibration curves for genes as well as standard control techniques, assays, protocols, and investigative algorithms
Commercial Manufacturing	• Develop standardized data formats for production equipment and instrumentation • Create on-line measurement methodologies to improve process understanding and establish standard QA/QC measures • Improve inspection and validation methodologies • Develop reference standards analogous to cellular material for future production cell and gene therapies
Postmarket Surveillance	• Standardize protocols and descriptions adverse event data to engender greater efficiency in ongoing safety and efficacy monitoring and FDA reporting • Standardize the syntax and interchange between clinical safety databases • Improved statistical methodologies to enable multivariate analysis of safety data • Develop uniform standards for data formats for clinical records

avoid the obstacles and costs associated with surmounting technical barriers, many of which would likely otherwise prevent the company from succeeding.

The broader biopharmaceutical and biotechnology industry would benefit from greater efficiency and effectiveness with a nationally coordinated standardization effort supported by an independent research organization with proven technical expertise, technology transfer abilities and access to financial and technical resources. The ultimate beneficiaries are patients who gain access to a broader array of novel therapies where development is supported by an effective technology infrastructure.

1 Introduction

Traditionally, drug companies developed new pharmaceutical products through the trial-and-error investigative processes of medicinal chemistry. They created small chemical compounds—referred to as small-molecule drugs—developed through the analysis of symptoms that characterize certain illnesses and diseases.[1] Although these small chemical entities still account for a significant share of the drug industry's development pipeline, they no longer dominate the focus of new drug development.

Over the last 2 decades, the drug development industry has shifted toward larger more complex molecular compounds that target biochemical mechanisms instead of symptoms. These "large-molecule" drugs are called biopharmaceuticals because they take advantage of how human biological systems function. Researchers capitalize on the specific, precise, and predictable attributes of subcellular molecules, like DNA and proteins, across a diverse set of species and cell types to develop medicines.

Biopharmaceuticals are medicines produced from the living cells of mammals, plants, viruses, and bacteria. They consist of products like recombinant proteins like human insulin, monoclonal antibodies like several cancer therapies, and vaccines like those for measles, mumps, and rubella. A distinguishing characteristic of biopharmaceuticals is that

---

[1]An alternate, more technical definition is that traditional pharmaceutical drugs are characterized as small chemical compounds with a molecular weight typically less than 500 Daltons that were developed through medicinal chemistry and are metabolized in the liver.

they are not tablets that patients take orally[2] but are intravenous and subcutaneous products administered by injection (King, 2006).

The term "biotechnology industry" is widely used to describe a cluster of firms using a diverse set of techniques from engineering and the life sciences to enable novel products and services, such as new drugs and medical diagnostic devices. The Biotechnology Industry Organization (BIO), the leading trade association that spans the many industry sectors engaged in biotechnology-related activities, reports that there were nearly 1,500 such firms employing over 198,000 people in the United States in 2004 (BIO, 2004).

The biopharmaceutical industry's origins are rooted in research by Stanley Cohen and Herbert Boyer that led to a process called recombinant DNA, in which pieces of DNA from cultured cells are artificially combined in a controlled setting. Recombinant DNA was a novel technique that enabled large-scale production of biopharmaceutical products like human insulin. From these origins, the industry has grown into one that currently spends $21 billion on R&D annually and has commercialized over 400 products.

The dramatic growth in the number of large-molecule drug candidates in clinical development can be attributed to exponential growth in our scientific understanding of the biological systems associated with human health and disease. This increase in scientific knowledge allowed researchers to expand the list of potential disease targets and advance genetic/protein engineering and modification technologies (Reichert and Valge-Archer, 2007).

An emerging research model for biotechnology products is systems biology, an integrated approach to analyzing biologic interactions and understanding how biological systems function. Systems biology integrates scientific disciplines and the computational power of information technologies to analyze the biological components and bioprocesses associated with human development, wellness, disease, and aging. By expanding our scientific understanding of human systems and their functional interaction, it is anticipated that systems biology will enable more effective innovation in diagnostics and therapeutics. Leroy Hood, president and cofounder of the Institute for Systems Biology,

---

[2]Biotechnology has enabled the development of novel delivery mechanisms that carry small-molecule drugs to their therapeutic target. These delivery mechanisms greatly improve drug efficacy by preventing of- target metabolism of the drug, which results in more of the drug dose reaching the intended therapeutic target.

predicts that this powerful combination of knowledge and tools will ultimately enable a new era of predictive, preventative, and personalized medicine. Systems biology has the potential to increase the productivity of R&D for both biotechnology-based drugs and traditional small-molecule drugs.

In the biopharmaceutical industry, systems biology has the potential to lower the cost of drug development and shorten the time it takes to bring a drug to market. Having a clearer understanding of the mechanism of disease will enable researchers to optimize drug efficacy. Moreover, better biological information about patients and potential toxicity indicators will allow researchers to streamline their clinical trials, ultimately reducing the number of costly clinical trial failures. For patients, systems biology approaches may reveal new ways to monitor health, identify and characterize transitions from healthy to diseased states, and provide information that helps patients and clinicians take preventative measures.

However, the technology infrastructure—the core data, methods, measurements, and technologies that enable effective R&D—must be sufficiently developed to accommodate widespread, effective use of systems biology approaches. Identifying and quantifying relationships and interactions between biological components and processes will require dynamic measurement technologies and the ability to process terabyte quantities of experimental data, representing an enormous investment in computing systems and laboratory instrumentation and software. Information must be able to be communicated among data sets, technology platforms, and organizations. This implies that the protocols and assays used to acquire these data be transparent and standardized within and across organizations. Researchers must have confidence in and assurance of the accuracy and precision of the data and processes informing drug development decisions. There is a need for a technology infrastructure that will improve the resolution, precision, and certainty around the experimental data these analytical tools create.

The purpose of this study is to inform the National Institute of Standards and Technology's (NIST's) Strategic Planning and Economic Analysis group of

- the biopharmaceutical industry's annual expenditures on technology infrastructure–related investments and activities;

- key areas within the technology infrastructure in which industry has underinvested because of market and technical barriers or because of unattractive risk-reward ratios at the firm level; and

- the potential efficiency gains an improved technology infrastructure offers the industry, as represented by potential cost reductions, shorter time to market for new drugs, and greater probabilities of developing successful products approved by the Food and Drug Administration (FDA).

1.1 THE NATURE AND ROLES OF TECHNOLOGY INFRASTRUCTURE

The concept of technology infrastructure warrants an explanation because the technology infrastructure's attributes are this study's principal topic. Chapter 2 delves deeply into infrastructural issues, but a top-level understanding is imperative for understanding the scope of this report.

As the term "infrastructure" implies, the technology infrastructure refers to the tools, methods, measurements, and data that enable or support R&D, products, and services. These are considered infrastructural because they are not necessarily commercial products themselves; rather, they support or embody processes and components that make many advanced technology products and services possible.

Many elements of the technology infrastructure are unseen or taken as a given because they are deeply embedded in or underlie research methodologies and instruments. For example, techniques that control process quality or verify the accurate calibration of laboratory instruments are part of the infrastructure, as are standard reference materials and measurements that researchers use to increase their confidence and assurance of the accuracy and precision of their work. More visible components of the technology infrastructure include analytical instruments and advanced software systems and algorithms.

The technology infrastructure enables productivity and efficiency in each of all major stages of drug development. Although some researchers may not notice their presence, gaps in the technology infrastructure are often readily apparent to other researchers because they hamper productivity and collaboration and thus present additional obstacles to the development of new drugs.

Technology infrastructure expenditures may consist of expenses incurred through internal activities; external purchases; and participation in consortia, partnerships, or other research activities:

- the purchase of measurement-related equipment, software, reference materials, or services;

- activities required for setup and validation of analytical instruments, reagents, or other research tools; examples of these activities may include developing calibration test methodologies, standard operating procedures, or process standards in measurement and manufacturing practices;

- in-house customization of technology platforms purchased from third-party vendors;

- efforts to develop interoperability between different software or equipment systems; and

- license fees or any other spending on enhancements to routinely used processes or equipment intended to increase productivity, reduce redundancy, or improve the confidence in results.

1.2 THE BIOPHARMACEUTICAL INDUSTRY

This report's scope is limited to the biopharmaceutical industry because the diversity and varying definitions of biotechnology make it difficult to study at the level of clarity and specificity required by an economic analysis.

1.2.1 Finding a Focus within Biotechnology: Approaching a Broad Industry Category

The biotechnology industry spans many disciplines from agriculture and the environment to health care and industrial applications. It eludes categorization according to traditional classification systems like the North America Industry Classification System (NAICS) because of the diversity of firms' research applications and the breadth of their market foci. No one grouping of homogeneous firms or organizations defines this rapidly evolving industry (Toole, 2003).

The industry is often differentiated by the applications its R&D seeks to enhance: therapeutics (e.g., drugs and devices), diagnostic/detection applications (e.g., medical, national security), chemicals (e.g., pesticides, insecticides, and new chemicals), agricultural (e.g., seed, plant, and animal applications), food and cosmetics, environmental, and energy (e.g., biomass).

Researchers in the biopharmaceutical industry and other knowledgeable stakeholders will note that the report also has a focus on protein-based biopharmaceuticals. The decision was made to recognize how the diversity of these drugs translates into diverse needs and practices, but also to narrow the product focus to streamline technology infrastructure discussion. Many of the underlying research tools and methods are generic in their application across the different categories of biopharmaceuticals. Given that recombinant proteins and monoclonal antibodies represent the majority of compounds currently in the development pipeline, this report focuses more on the development processes for these therapies.

1.2.2   Biopharmaceutical Products

Biopharmaceuticals address a diverse set of biological problems associated with very different parts of the human anatomy. It is useful to define these categories of drugs to set the stage here for the more technical description of the technologies that support their discovery and development that follows in later chapters. Currently, there are inconsistent definitions, taxonomies, and classification schemes applied to the term "biopharmaceutical." Inconsistent terminology and definitions make it difficult to analyze this industry: the number of approved biotech products is highly inconsistent among industry sources that include BIO and PhRMA.

For the purposes of this report, we define biopharmaceuticals as the subset of biotechnology-based products explicitly used for therapeutic purposes in humans. Biopharmaceuticals include the more traditional therapeutic proteins derived from recombinant DNA technologies, as well as monoclonal antibodies, vaccines, gene therapies, and cell therapies. Table 1-1 presents one potential categorization of all approved biopharmaceuticals marketed as of 2005.

Therapeutic Proteins

Therapeutic proteins were the first class of biopharmaceuticals to be commercially developed and approved for therapeutic use in humans. Therapeutic proteins include enzymes, hormones, and proteins that are naturally produced by the human body under normal conditions. Therapeutic proteins are frequently labeled as biologics because they are intended to replicate or replace proteins, enzymes, or hormones that would otherwise be produced naturally within the human body.

Table 1-1. Number of Biopharmaceutical Class Types in 2005

Biopharmaceutical Class	Type	Number of Approved Products
Recombinant DNA products	Therapeutic Proteins	9
	Monoclonal antibodies	19
	Gene therapies	1
	Monoclonal antibodies	13
Non-recombinant DNA products	Vaccines	128
	Toxins	4
	Enzymes	19
	Cultured cells and tissues	10
	Human blood products	108
	Animal blood products	24

Source: Adapted from Rader, 2005.

Human insulin is one example of a therapeutic protein that is used to treat diabetes in patients who can no longer produce the enzyme naturally. Epotein is another example of a therapeutic protein used to stimulate the production of red blood cells in the body for patients who suffer from anemia or low red blood cell count. Epotein is intended to replace erythropoietin, the naturally occurring protein produced in the kidney that regulates red blood cell production. Therapeutic proteins are the product of genetic engineering of human or animal cells and produced through recombinant DNA technologies.

Monoclonal Antibodies

Monoclonal antibodies (mAbs) are the second category of biopharmaceuticals that capitalizes on cell receptors present on the surface of almost all cells to effect changes in cell growth or reproduction. Antibodies or other proteins can bind to the cell receptor and cause changes to occur in the cell. In the cases of tumor cells, antibodies can slow the growth or stop the production of new cancer cells.

Rituximab is one example of a monoclonal antibody that binds to a cell receptor on B-cell non-Hodgkin's lymphomas (tumor cells) that causes the tumor cells to disintegrate and in some cases prevents the production of new tumor cells. Herceptin is another example of a monoclonal antibody that is used in breast cancer patients when the

cancer has spread or metastasized. Herceptin binds to a cell surface receptor HER2, present on approximately 33% of all breast cells, to block tumor growth and development. Both, Rituximab and Herceptin are produced using molecular cloning and recombinant DNA technology.

Vaccines

Vaccines are preventative treatments that build an immune response to disease by introducing nonlethal doses of the viruses to the body to stimulate the production of natural antibodies that kill the invading disease agent. Vaccines have traditionally been given as a series of subcutaneous injections. Vaccines can be prophylactic or can prevent or modify future infections from a natural pathogen (a foreign agent that enters the body), or therapeutic as is the case with vaccines for diseases such as AIDS, HPV, and hepatitis B.

Traditional vaccines containing killed or live attenuated viruses stimulate an immune response in the body. The immune system recognizes the foreign pathogen and develops antibodies that will destroy it. Many common vaccines for viruses such as influenza, cholera, yellow fever, and measles are created using the killed or live attenuated vaccine method. Tetanus and diphtheria vaccines are examples of toxoid vaccines that include an inactive toxic agent from a specific virus rather than the entire virus.

Gene Therapies

Gene therapies are defined as the insertion of a corrected or normal gene into individual cells and tissues to replace abnormal disease-causing genes. A viral vector such as an adenovirus or retrovirus is used to deliver the normal gene to individual cells. Once inside the cell, the normal gene stimulates the production of proteins needed to restore the cell to a normal healthy state.

Gene therapies were originally delivered directly to targeted cells outside of the body and then delivered to the body through some surgical procedure. More recently, researchers have inserted corrected modified DNA into adenovirus vectors, which is used to deliver the gene therapy to targeted cells. One significant problem with gene therapies is immune response. Similar to organ transplant, when corrected genes are introduced to the cell, the immune system no longer recognizes the cells as normal and begins to destroy them.

Cell Therapies

Cell therapies, similar to gene therapies in approach, deliver new normal cells or tissue to a diseased part of the body. Cell therapies to date have focused on treating hereditary diseases. Forms of cell therapies include the transplantation of stem cells and the transplantation of functional or mature cells.

Limitations to Biopharmaceuticals

Limitations to biopharmaceuticals include delivery, immunogenicity, and cost. Because of their higher molecular weight and sensitivity to harsh conditions, many biopharmaceuticals cannot be ingested like small-molecule drugs since their proteins or antibodies denature in gastric acids. Hence, these drugs are typically administered through intravenous or subcutaneous injection. In the case of cell therapies, delivery of new cells to a host system such as the brain requires an invasive surgical procedure.

Immunogenicity, which is an undesired immune system response to a biopharmaceutical once inside the body, is also a major limitation of biopharmaceuticals. Biopharmaceutical antibodies typically use an antibody known as an adenovirus to locate and bind to specific cell surface receptors. In some cases, the body's immune system may create antibodies that try to destroy the foreign adenovirus. This immunogenic response can reduce or eliminate the drug's efficacy. If the immunogenic response is nonspecific, the immune system may destroy other proteins in the body that are similar to the biopharmaceutical antibody but needed for other functions. Under some conditions, such an immune response may result in major organ failure and death.[3]

Cost is another limitation of biopharmaceuticals. Typically biopharmaceuticals are more expensive than their counterpart small-molecule drugs because of special storage and handling requirements, smaller patient populations, and high administration costs. Biopharmaceutical injections may require the assistance of skilled clinicians or even surgical procedures in the case of cell therapies.

Another key issue impacting biopharmaceuticals in the near future is follow-on biologics. Follow-on biologics are "generic" formulations of an

---

[3]A noted example of this was Jesse Gelsinger, who died in 1999 as a result of immunogenic response during a gene therapy clinical trial to treat a rare genetic liver disease. He died 4 days after suffering a massive immunogenic response to the adenovirus that carried the gene therapy to his liver cells.

existing biologic drug; approval and regulation of follow-on biologics is a key issue facing the industry. Follow-on biologics depend on the same mechanism of action as their predecessor but may or may not be exact replicates of the original version. Similar to a class of drugs like statins, follow-on biologics are innovative products that closely resemble each other. Measurement needs related to follow-on biologics center primarily on the need to characterize the molecular profiles of these drugs. The high molecular variability that can occur during the manufacturing process has raised safety concerns, and measurement technologies that more accurately characterize biologic profiles are not well developed.

1.2.3 Defining the Biopharmaceutical Industry

As of 2006, the biopharmaceutical industry accounted for over 260 therapeutics and vaccines approved for over 380 indications (BIO, 2006). FDA approved the first biopharmaceutical for marketing in 1982, marking the inception of the biopharmaceutical industry 25 years ago.

Conventional pharmaceutical firms have seen R&D spending continually increase, while the number of new drug applications (NDAs) has declined over the same period. In light of the declining R&D productivity of more conventional "small molecule" research, biopharmaceutical leaders such as Amgen, Genentech, Millennium Pharmaceuticals, Biogen, Genzyme, and others offer the promise of improving R&D productivity through large molecule-based biologics. However, advances in the technology infrastructure are needed to improve the probability of success for this class of compounds, ultimately improving attrition/failure rates at each stage of the R&D process and thereby lowering the per-drug cost of R&D.

The length of time from initial discovery of a biopharmaceutical through approval can average 8 to 15 years, with most costs incurred during clinical trials (BIO, 2004). In addition, revenue from successful products must cover the costs from failed efforts as researchers seek to recoup investment costs.

For the pharmaceutical industry as a whole, only 1 in roughly 10,000 compounds screened during the drug discovery stage succeeds through the clinical trial stage and results in an FDA-approved drug (PhRMA, 2005). Of drugs entering clinical trials, FDA reports only about 8% ever reach the market to recoup those expenditures. Furthermore, the failure rate in the most expensive Phase III stage is nearly 50% (FDA, 2006).

Similar statistics are not readily available for the biopharmaceutical sector alone, but some studies suggest clinical trial success rates are higher for biopharmaceutical therapeutics.[4] Even if attrition is less of a problem for biopharmaceuticals, the development stages are more complex. These factors translate to a high price for prescription drugs, and according to recent reports, spending on biopharmaceuticals "is growing twice as fast as traditional prescription drugs" (Anand, 2005).

In 2004, the average biotechnology company spent more than 24% of its revenues on R&D, compared with 13.3% for large pharmaceutical manufacturers (Ho and Gibaldi, 2003). (see Table 1-2).[5] Many biopharmaceutical companies are still in the development stage and have yet to commercialize drugs that take more than a decade to develop.

Table 1-2. 10 Largest Biopharmaceutical Companies by Net Sales in 2004

Company Name	NAICS	R&D Total Expenditures ($million)	Net Sales ($million)	Relative R&D Intensity (R&D/Sales)
Amgen	325414	$2,582	$10,550	24%
Genentech	325412	$816	$4,621	18%
Serono	325414	$595	$2,458	24%
Biogen-IDEC	325414	$688	$2,212	31%
Genzyme	325414	$646	$2,201	29%
Chiron	325412	$441	$1,605	27%
Gilead	325414	$224	$1,325	17%
Medimmune	325414	$357	$1,141	31%
Millennium Pharmaceuticals	325413	$266	$448	59%
Intermune	325414	$81	$151	54%

Source: Compustat North America.

Note: NAICS Industry Descriptions:
   325412—Pharmaceutical Preparation Manufacturing
   325413—In-Vitro Diagnostic Substance Manufacturing
   325414—Biological Product (except Diagnostic) Manufacturing

---

[4]See, for example, Pavlou and Riechert (2004), who report a 35% approval rate for recombinant protein therapeutics entering clinical trials from 1990 to 1997.
[5]While the 13% R&D-to-sales ratio is the industry average, the top 10 drug manufacturers' average was slightly higher at 16%. The top 10 drug manufacturers include Pfizer, Novartis, Johnson & Johnson, GlaxoSmithKline, AstraZeneca, and Merck.

Changing demands in consumer markets are also driving biopharmaceutical R&D as a result of an increasingly health-conscious population and a shift toward preventive and predictive medical treatments. Traditionally, R&D expenditures targeted primarily cancer and cardiovascular disease markets. However, emerging markets for the treatment of metabolic disease (e.g., obesity and diabetes), treatment of memory disease (e.g., Alzheimer's), and wellness (i.e., preventative or predictive) therapies are expanding the targeted areas for drug development. In addition, R&D is targeting new approaches to the traditional cancer and cardiovascular diseases, such as the development of therapeutic vaccines to treat heart disease and certain types of cancer. Other diseases targeted include neurological, pulmonological, orthopedic, metabolic, renal, and infectious diseases.

Medicare's recent adoption of prescription drug coverage combined with the "baby boom" generation approaching retirement will result in the Centers for Medicare & Medicaid Services (CMS) becoming a dominant consumer (representing 40% of the market) of prescription drug purchases (Burrill & Company, 2006). Demand for generic drugs is expected to increase as CMS becomes a larger consumer. As a result, new R&D programs are focused on reducing the time to market of new drugs and increasing years of peak sales over the life of a drug's patent.

1.2.4 The Biopharmaceutical Product Development Cycle

The exact path a biopharmaceutical firm takes in developing a human therapeutic varies greatly depending on the disease application and the type of product.[6] For example, the R&D process for monoclonal antibodies and recombinant proteins may look quite different from the R&D process for vaccines, DNA (gene-based) therapies, or cell-based therapies. However, all these drugs start with a foundation in basic research and pass through the following generic stages:[7]

- drug discovery,
- preclinical development and testing,
- clinical trials,

---

[6]A biotechnology firm may not even be the sole organization involved in developing such products. Often a firm will license intellectual property from universities or other research organizations, or the firm may license discoveries to large pharmaceutical organizations for further development.

[7]These five stages are presented in chronological order, with the exception that scale-up to commercial manufacturing is undertaken simultaneously with preclinical development and clinical trials.

- scale-up and commercial manufacturing, and
- postmarket surveillance.

The following discussion is intentionally oversimplified to provide readers unfamiliar with the drug development process a brief overview. The drug development process is revisited in Chapter 2, where the technology infrastructure supporting drug development stages is discussed.

Drug Discovery

This stage in the development cycle involves four primary phases: target identification, target validation, lead identification, and lead optimization. Target identification marks the beginning of the discovery process. Researchers use their understanding of molecular biology, genomics, and proteomics to identify pathways to disease or disease "targets."

Researchers identify and validate drug targets and identify compounds that have the ability to affect the activity of a target (e.g., a protein). Those leads demonstrating the highest probability of success—typically 250 optimized leads for every one approved drug, according to PhRMA (2005)—then enter the preclinical stage.

Preclinical Development and Testing

Once researchers have identified a lead compound, they must conduct a number of preclinical studies before testing the drug therapy in humans. Researchers perform experiments to determine the biochemical and physiological effects and properties of the compound. At the same time that researchers are investigating the pharmacological properties of the drug candidate, they are beginning to develop the process for manufacturing the compound.

Clinical Trials

After developing a promising drug, a developer will submit an Investigational New Drug (IND) application to FDA. Drug companies file applications for gene therapies, cellular products, vaccines, and plasma-derived products with the Center for Biologics Evaluation and Research (CBER), and protein therapies, such as monoclonal antibodies, cytokines, enzymes, and growth factors with the Center for Drug Education and Research (CDER). FDA also has an Office of Combination Products to oversee approvals for those products.

Clinical trials typically include three phases:

- Phase I trials are primarily meant to test the safety, tolerance, and behavior of the IND in human subjects. These trials have the smallest patient group. If the therapy is found to be safe, it progresses to the next trial phase.

- Phase II trials expand the patient group significantly. Researchers continue to monitor the safety of the therapy because of the increase in genetic diversity present in a larger patient group, but the focus of this stage is on learning how the IND behaves and interacts within the human organism at different dosage levels.

- Phase III trials monitor safety and drug behavior but are meant to serve as the definitive assessment of the IND's efficacy. These trials have the largest patient group and are generally the last round of trials before the IND is submitted for FDA approval.

Scale-Up and Commercial Manufacturing

Production can be divided into two major stages: scale-up and commercial manufacturing. Scale-up includes manufacturing-process design, product specification, and optimizing production activities. Bulk manufacturing includes commercial-scale upstream and downstream processing. These two stages are discussed together in this section for ease of presentation. However, in practice, scale-up manufacturing activities occur in tandem with the preclinical and clinical development stages. Commercial manufacturing occurs following FDA approval.

The commercial production of biopharmaceuticals includes upstream and downstream processing. Upstream processing is defined as the fermentation process that results in the initial core product material. Downstream processing refers to the actual purification of the product and formulation, bottling, and labeling activities.

Postmarket Surveillance

Postmarket surveillance refers to the on-going, indefinite monitoring of the efficacy and safety of an FDA-approved product. Once clinical trials are complete, the biopharmaceutical drug or therapy sponsor submits an NDA containing all scientific information collected during preclinical and clinical studies. Among other documents, the sponsor must submit an integrated summary of efficacy (ISE) and an integrated summary of safety (ISS), which in combination document the benefit-risk trade-off for the candidate drug.

For some therapies, FDA will approve the drug provisionally and will require additional "Phase IV" studies to evaluate long-term effects. Even

if approval is not conditional, sponsors must produce a Periodic Safety Update Report (PSUR) quarterly for the first 3 years and annually in years that follow. These updates report adverse drug reaction (ADR) cases and alert FDA to potential safety issues that may prompt a reevaluation of therapy labeling. A particular concern related to the postapproval performance of a biopharmaceutical involves stability of the therapy and the conditions under which the therapy must be handled, stored, and dispensed to remain effective.

Postmarket surveillance for safety and efficacy, also known as *pharmacovigilance*, will always be necessary despite even the most rigorous preclinical and clinical testing phase evaluations. These earlier phases usually will have involved only a few thousand human patients, and certain therapy responses or interactions are so statistically rare that they only manifest once the therapy reaches the large market population. The costs of collecting such data and producing these reports can be sizeable, but the private and social costs of an unsafe therapy that must be recalled or remarketed under restricted criteria are much more substantial.

1.2.5 Biopharmaceutical Technology Infrastructure Complexity

Looking toward the future, biotechnology is enabling a new era of human therapeutics that builds on advances in nanosciences and systems biology. The biopharmaceutical industry's ability to create these innovative therapies will largely be determined by several factors, including the productivity and efficiency of R&D, manufacturing quality control and assurance programs, and postmarket surveillance activities. Large variations in productivity and efficiency currently exist at each step in the biopharmaceutical development life cycle. For complex science-based technologies like biopharmaceuticals, major gains in these performance attributes require a complex and ubiquitous technical infrastructure.

While biopharmaceuticals capitalize on the specific and predictable attributes of subcellular organic molecules, like DNA and proteins, inherent random variation associated with producing drugs from living organisms demands an even more complex technical infrastructure than is needed for inorganic complex science-based technologies like semiconductors.

1.3 PROJECT SCOPE AND GOALS

This report characterizes the infratechnologies used to develop new biopharmaceuticals and estimates the biopharmaceutical industry's current annual spending on the technology infrastructure. The report also identifies technological inadequacies or gaps that may exist and estimates the potential efficiency gains an improved infrastructure would hold for developing and producing new biopharmaceutical products. Accordingly, subsequent chapters

- define and characterize the range and roles of technology infrastructure in the biopharmaceutical industry,
- estimate the annual expenditures for this infrastructure incurred by biopharmaceutical firms,
- estimate the "excessive" costs incurred because of inadequacies of existing infratechnologies, and
- assess research priorities and industry's awareness of NIST's capabilities.

Although benefits ultimately accrue to patients, the project methodology concentrated on quantitatively assessing the economic benefits to firms. The inherent difficulty in linking investments in infrastructure technologies to health outcomes precludes providing quantitative estimates of consumer benefit.

1.4 REPORT ORGANIZATION

The report is organized as follows:

- Chapter 2 discusses the biopharmaceutical technology infrastructure in detail and includes stakeholders' comments on technical and market barriers to an improved infrastructure.
- Chapter 3 presents the approach to quantifying annual infrastructure-related spending and estimating the potential an improved infrastructure holds for the costs of developing and manufacturing biopharmaceuticals.
- Chapter 4 describes interview protocols and provides summary statistics on the sample pool of study participants that contributed to this research.
- Chapter 5 presents quantitative results, including the potential efficiency gains from an improved infrastructure.
- Chapter 6 summarizes research findings and offers concluding remarks on opportunities for NIST.

2 The Technology Infrastructure Supporting the Biopharmaceutical Industry

This chapter presents a broad characterization of the biopharmaceutical industry's technology infrastructure and relates stakeholders' views on the technical and market barriers that impede greater efficiency in biopharmaceutical R&D and production. Study participants' insights provide a context for the cost model and the economic methodology presented in Chapter 3. Also included in this chapter is a discussion of six technology focus areas (TFAs) for which counterfactual technology infrastructures were investigated.

2.1 DEFINING THE BIOPHARMACEUTICAL TECHNOLOGY INFRASTRUCTURE

This chapter begins by defining the technology infrastructure. It includes several illustrative examples of the infrastructure components of the biopharmaceutical industry's development and production processes. Feedback from study participants on the TFAs and efforts to bridge those gaps are discussed later in the sections of this chapter devoted to technical and market barriers.

In addition to defining this technology infrastructure, we offer preliminary comments on technology gaps and inefficiencies informed by the National Institutes of Health's (NIH's) Roadmap for Medical Research and FDA's Critical Path report. These documents represent strategic initiatives in the areas of technology development that experts believe will alleviate bottlenecks and reduce the costs of biotech drugs.

2.1.1   Technology Infrastructure Components

The technology infrastructure has been defined as a suite of infratechnologies used jointly by competing firms within a specific industry, often in the form of standards (Tassey, 1997). In the biopharmaceutical industry, this suite of infratechnologies enables the advancement of scientific knowledge through proof of concept (generic technology) research and then applied R&D, the results of which lead to innovation and ultimately new products and process techniques. Technology infrastructure consists of three interlocked components: generic technologies, proprietary technologies, and infratechnologies.

Infratechnologies are defined as a set of "technical tools and processes" that enable efficiency increases in R&D, production, and market development activities. They can include measurement and test methods, materials characterization, standard reference materials, process models, analytical methods, and standards in data formats and systems interoperability. Infratechnologies are methods employed to conduct testing and analyze results using a combination of various tools, techniques, and procedures.

Infratechnologies have a role in every stage of economic activity. In R&D, infratechnologies support the development of generic and proprietary technologies and improve the efficiency, precision, and repeatability of R&D tests. Infratechnologies also improve the quality and reliability of production processes. These improvements affect the efficiency of market transactions by reducing the performance risk to market participants.

Infratechnologies leverage the productivity of the processes of developing generic technologies (also called "technology platforms") and subsequent proprietary technologies (innovations).

Generic technologies represent the first attempts to construct a laboratory proofs-of-concept derived from basic science. Generic technologies have no specific market applications as stand-alone technologies; rather, they represent a potential trajectory toward a new market application. Tassey (1997) suggests that generic technologies must exist before innovative market applications can be identified. This inherent quality means that generic technologies may reduce the technical risk associated with pursuing additional applied R&D. Generic technologies are considered infrastructural because competing firms

draw on the same generic technologies in pursuit of innovation, ultimately leading to new market applications.

Proprietary technologies consist of methods, processes, and techniques firms develop to achieve their strategic objectives in conducting the development and production processes required to bring new products and services to market. Competing firms differentiate themselves through their implementation of these proprietary technologies. Typically, proprietary technologies are traceable to industrial standards to ensure market acceptance.

2.1.2 The Biopharmaceutical Technology Infrastructure Taxonomy

The biopharmaceutical technology infrastructure comprises the aforementioned infratechnology categories. Many of them are embodied in broader technical infrastructure categories such as analytical systems, bioinformatics, and consumables and reagents used every day.[1] These include the following:

- **Analytical Systems** incorporate the generic and proprietary technologies embodied in the equipment and hardware used in determining biological samples' physical and behavioral characteristics. Analytical systems include
    - sequencing platforms;
    - expression platforms, including microarrays;
    - spectrometry platforms;
    - spectroscopy platforms;
    - imaging technologies (MRI, PET, CT);
    - chromatography platforms;
    - electrophoresis platforms;
    - fluorescence and flow cytometry; and
    - screening platforms, including microfluidics.
- **Bioinformatics** includes all computational-based technologies, such as bioinformatic and chemoinformatic databases, predictive modeling software, statistical analysis software, and data capture systems. Bioinformatics encompasses
    - computing systems,

---

[1]This taxonomy leverages NIH's Roadmap for Medical Research and FDA's Critical Path report and the U.S. Measurement System (USMS) assessment, which is currently conducted by the NIST. In addition, professional associations, such as the Analytical Life Science Systems Association (ALSSA), a trade association representing the scientific and technical instrument industry, offer some guidance on organizing the variety of technologies that are applied in the biopharmaceutical product development life cycle.

- databases,

- work management software,

- analytical software and algorithms,

- data acquisition technologies, and

- statistical methodologies.

- **Consumables and Reagents for Analysis** capture all chemical and biological aids that facilitate laboratory analysis of drug targets and their potential lead compounds (i.e., drugs):

  - molecular biomarkers;

  - nucleic acids;

  - radiochemicals;

  - labeling kits;

  - vectors, such as viral vectors for delivering gene therapies;

  - gas/liquid chromatographic supplies;

  - modifying/restriction enzymes; and

  - transfer membranes.

More generally, several infratechnologies are often combined to provide measurement and test methods, standards, reference materials and other critically evaluated data, models, algorithms, and interface specifications. For example, the use of mass spectrometry in proteomic research clearly involves measurement, but it also requires precise calibration, standards for sample preparation, computer interfaces for reading output, and quality databases for matching that output to known peptide sequences. Researchers often take infratechnologies for granted; the consequences of not having technologies such as calibrants, advanced chemical methods, and standard operating protocols—all of which researchers rely on every day—demonstrate the value they contribute to the overall body of technology enabling biopharmaceutical R&D.

2.1.3 The Economic Role of Biopharmaceutical Technology Infrastructure

The methods, techniques, and data discussed above form a complex technology infrastructure that enables productivity and efficiency in each of the major stages of economic activity related to biopharmaceuticals (such as basic and applied research, clinical trials, commercial manufacturing, and market development). Improvements in this infrastructure can have numerous potential economic impacts, including the following:

- Cost reductions
    - Lower labor and materials costs for discovery and development of a given therapy
    - Lower labor, equipment, and materials costs for production
    - Lower transaction costs promoting market penetration and postmarket tracking/assessments
- Accelerated time to market
    - Shorter time between discovery and FDA approval
    - Shorter time between new market demand and production response
- Quality improvements
    - Detection of drug failures in earlier clinical trial phases
    - Investigational New Drugs (INDs) have greater probability of receiving FDA approval
    - Reduced uncertainty in drug efficacy and safety in the population targeted for prescription, longer shelf lives, and less restrictive storage and handling needs

Improvements in the technology infrastructure can lower total costs by enabling go or no-go decisions with respect to research strategies in earlier stages of the R&D cycle. For example, one way to reduce overall costs might be to conduct more extensive *in vitro* (in test tubes or Petri dishes) and *in silico* studies before testing a therapy in animal models. Additionally, accelerating the time to market by realizing productivity gains results in both a cost reduction (fewer labor hours) and a revenue increase (as firms take better advantage of a limited patent life).

Despite the private and social benefits resulting from improvements in the technology infrastructure, economic theory suggests that the biopharmaceutical industry will underinvest in improvements. Generally speaking, the private return on generic technology or infratechnology investments is too low because of a lack of appropriability of the knowledge produced and the relative high risk inherent in these areas of research. Such market failures are often cited as a motivation for industry initiatives through trade associations, joint ventures, and public funding, which include the R&D investments made by NIST and other government organizations.

Investments to improve the technology infrastructure will result in benefits to firms in the biopharmaceutical and other industries developing and using biotechnology, as well as to consumers. However, there are inherent difficulties in quantifying the benefits to consumers. The uncertain nature of drug R&D makes predicting health outcomes difficult;

moreover, quantifying those health outcomes (i.e., quality-of-life metrics) is beyond the scope of this study. As a result, our approach focuses on estimating the potential benefits accruing directly to biopharmaceutical firms from reductions in R&D, manufacturing, and postmarket surveillance costs (in large part from greater success in getting approval for new drugs and getting them into the marketplace).

2.2 TECHNOLOGY FOCUS AREAS IN THIS STUDY

Discussions with RTI's internal experts, experts at the FDA and the nonprofit Critical Path Institute, and other industry participants and a literature review identified four compelling infrastructure technologies that could serve as TFAs for this study. The potential an improved technology infrastructure holds for the biopharmaceutical R&D is greatest with these TFAs. Thus, the following areas helped define what exactly is meant by an improved infrastructure by presenting study participants with a defined baseline against which to estimate the impact potential improvements may have.

For discovery, preclinical, and clinical (Phase I through Phase III) activities, four TFAs were selected:

1. Enhanced bioimaging techniques
2. Standards and metrology in gene and protein expression analysis
3. Improved bioinformatics and *in silico* predictive modeling
4. Identification and validation of molecular biomarkers

These infratechnologies have applications throughout the drug development process and cut across a number of different therapeutic categories.

In addition to the four drug-development-specific TFAs, two process-specific TFAs were included:

5. Infratechnologies to enhance scale-up and commercial manufacturing, including improvements in upstream and downstream processing and process monitoring/quality assurance
6. Infratechnologies to support postmarket surveillance activities, including product surveillance, tracking, and Phase IV clinical trials

2.2.1   Enhanced Bioimaging Techniques

Bioimaging refers to using imaging techniques to review biological structures and functions and to evaluate phenomena in a noninvasive manner. Although often used for diagnostic purposes, bioimaging and structural characterization can also play an important role in drug discovery, preclinical testing, and clinical trials.

In the discovery stage, bioimaging can be used to identify and map potential drug targets. In preclinical development, imaging can be used to inform pharmacokinetic and pharmacodynamic (PK/PD) studies, for example, by dosing a brain tumor with a highly fluorescent drug while monitoring accumulation of the drug in the tumor. PK/PD studies refer to the metabolism of a drug (pharmacokinetics) and how the drug behaves in the body (pharmacodynamics).

Furthermore, in preclinical and clinical stages, bioimaging can serve as a biomarker, signaling future response in animal models or actual response patterns in human trials. Improved structural characterization may reduce costs during manufacturing by increasing the homogeneity in populations produced for each therapeutic category. The ability to create a homogenous population could increase manufacturing yields and potentially reduce downstream purification and processing costs.

However, infrastructural advances must be made in the diverse imaging technologies (including CT, MRI, PET, and ultrasound)[2] to ensure the quantitative data produced by these technologies are accurate, reliable, and reproducible. Biopharmaceutical executives and researchers have also indicated that the throughput and resolution of these technologies can be too low (and the expense too high) for use in preclinical studies. Some technologies offering higher throughput and better resolution require labeling compounds that cannot be used in humans in a clinical trial setting (Chapman, 2005).

2.2.2   Gene and Protein Expression Analysis

Gene and protein expression analysis and profiling provide a complex picture of a cellular state by determining whether genes are turned off or on in a cell, or their expression level, for tens of thousands of genes at a specific time for a specific biological sample.

---

[2]These acronyms stand for computed tomography (CT or "CAT scan"), magnetic resonance imaging (MRI), and positron emission tomography (PET).

Expression analysis can be performed on blood, tumors, normal tissue, and many other types of samples. One example of gene expression analysis is comparing the gene expression levels between normal and cancerous tissue to identify genes that are differentially expressed only in the disease state. Not only could this information be used as a potential molecular diagnostic, but it could also direct new research to enlighten our basic understanding of these disease states.

Expression analysis is predominantly performed on DNA or RNA microarrays. Microarrays are small glass or silicon surfaces that are divided into tens of thousands of "features," where each feature is a small area that contains DNA or RNA sequences that are complementary to a specific gene sequence. These features are also known as "probes."

In a microarray experiment, researchers introduce a specially prepared blood or tissue sample to a microarray and track gene expression patterns by analyzing matches between the sample and the microarray's probes. Each microarray contains thousands of probes that are single-stranded DNA reference sequences. Where an RNA or DNA sequence from an experimental sample is complementary to a probe's sequence, the two will bind and emit a fluorescent signal (see Figure 2-1).

DNA microarrays are an important component of medical research because they document cells' responses to diseases and the effects of drug treatments (Gerhold, Jensen, and Gullans, 2002). They also take the vast amount of information from large-scale sequencing efforts and provide a platform from which DNA and mRNA samples can be tested against entire genomes. Until microarrays' introduction in 1994, scientists used $100 blood tests and other laborious diagnostic procedures to identify individual genes. A microarray permits thousands of those experiments to run simultaneously. Experiments that only a decade ago took weeks, months, or perhaps years to run now take only a couple of days (Malik, 2003). This speed has had a profound effect on the way medical research is conducted.

Repeatability in gene and protein expression analysis has been identified as a bottleneck in efficacy and toxicology testing during all drug development stages. A standardized set of reference materials and standard sample collection and management protocols would streamline this process by reducing variability of results across laboratories and multiple platforms.

Figure 2-1. DNA Hybridization in a Microarray Experiment
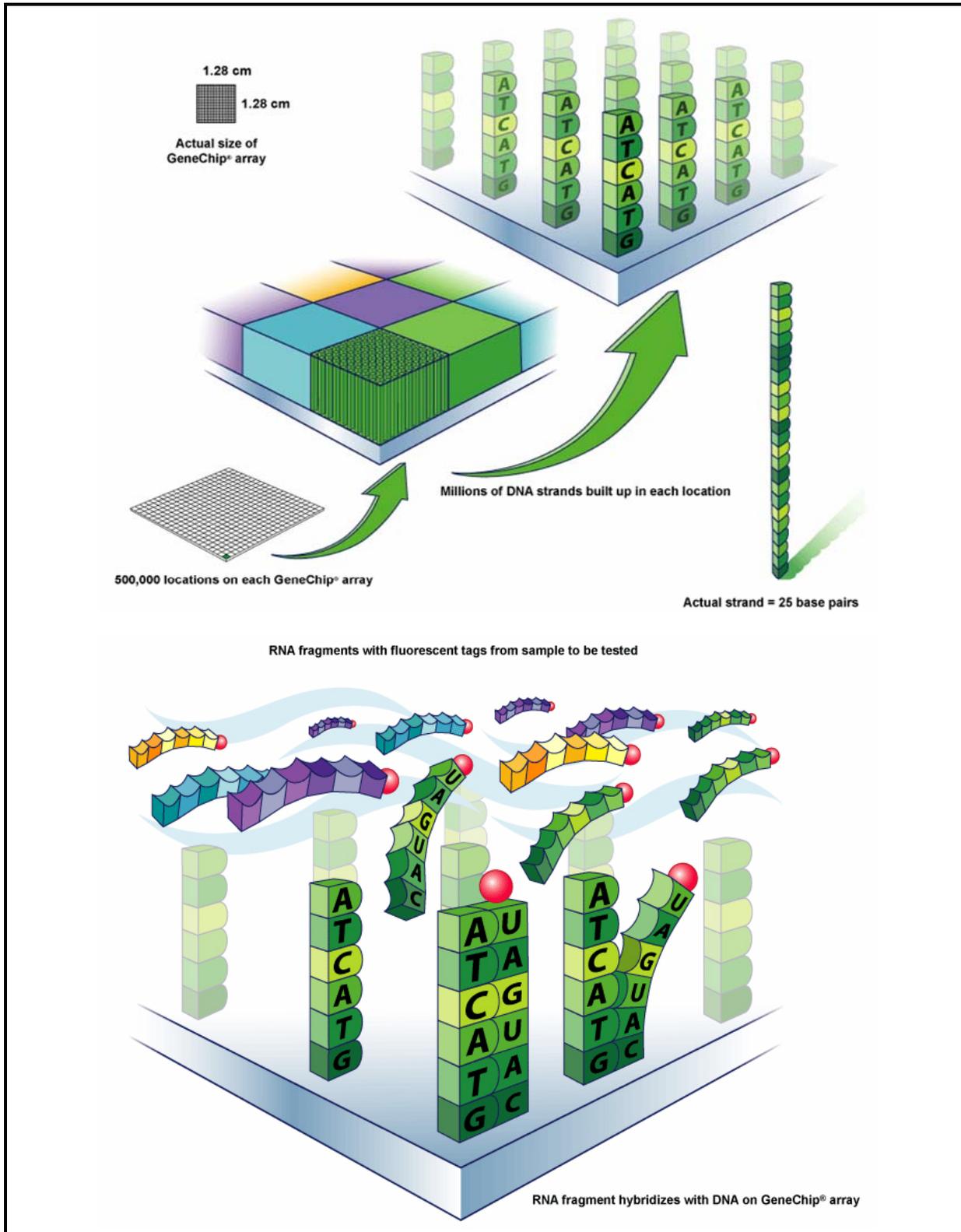
Source: Courtesy of Affymetrix.

Reference materials would ultimately improve robustness of results and reduce the quantity of data submissions. In addition, standard protocols or "best practices" would improve repeatability and allow results from different laboratories to be shared and compared. These would include unifying aspects of experimental design, reducing variability in the quantity and quality of RNA or protein sequence examined in a microarray experiment, using similar statistical and bioinformatic algorithms to analyze the experiment's raw data, and reporting final data in standard formats.

2.2.3 Bioinformatics and *In Silico* Predictive Modeling

Bioinformatics is analogous to what many scientists call computational biology, which means the use of advanced computing techniques to create databases of biological and chemical information, investigate relationships among data, and model biological systems mathematically. This report uses the term bioinformatics.

Bioinformatics has had a turbulent history in the biopharmaceutical sector. The fields of genomics, proteomics, and metabolomics—aided by advances in instrumentation and high-throughput data analysis—have produced a vast amount of data for use in the early stages of drug discovery. Yet the various databases researchers use lack integration—both across content types (i.e., DNA microarray data and screening data) and across stages of drug development. The effort to enhance interoperability is important for moving basic research results more quickly into a clinical setting.

This research also requires analysis servers or computing clusters for computational biology algorithms that require multiple CPUs and large amounts of memory. Data format standards do not currently exist, but they are becoming essential, including consistent ways to record the outputs of microarrays and bioimaging devices and ways to describe experimental protocols, which can greatly increase the value of legacy data (Quackenbush, 2004).

Outside of R&D, manufacturing process control and quality monitoring systems constitute another area where interoperable bioinformatics tools are needed. Reductions in lot-to-lot variation for all therapeutic categories may lower quality assurance costs in meeting good manufacturing practice standards. Additionally, data integration and format standards are needed to provide a link between preclinical, clinical, and postmarket safety and efficacy studies.

In silico (computer-based simulation) modeling may be applied to therapeutic proteins, vaccines, and cell and gene therapies. *In silico* models can be used for target validation and lead identification, predicting the ability of compounds to bind effectively at the appropriate site. By using results from previous *in vitro* or animal tests, computer programs build metabolite models that are virtually screened against known compounds to "test" interactions (McGee, 2005b).

The main advantages of *in silico* modeling are the high throughput and low labor and materials costs. These features make it attractive for screening thousands of lead compounds in a matter of hours, and doing this earlier in the process allows researchers to prioritize compounds for development. Although *in silico* modeling will not replace *in vivo* models (in whole organisms, such as mice or larger animals), it could improve the quality of compounds being tested *in vivo*, reducing failure rates and lowering overall costs.

Costs may be substantial, given that failures due to toxicity, for example, range from 30% to 40% (McGee, 2005b). The need for predictive toxicology is more pressing than the need for predictive absorption, distribution, metabolism, and elimination (ADME) for candidate drug compounds (where *in silico* tools are reportedly regularly in use), and this had been expressed in increased research expenditures and workshops, such as the NIH Summit Workshop on Predictive Drug Toxicology, held in June 2004.

2.2.4   Molecular Biomarkers

Biomarkers are molecular or cellular indicators of susceptibility to a disease or condition. Researchers are actively seeking out and defining biomarkers in the hopes of using them to make drug discovery and testing more efficient. Predictive biomarkers may one day serve as surrogate endpoints to shorten clinical trials or stratify patient populations by identifying those individuals with certain predispositions toward either responsiveness or adverse reactions. In earlier stages, biomarkers may allow drug makers to optimize their selection of drug candidates, avoiding compounds that show early signs of liver toxicity, for example.

Molecular biomarkers are especially cross-cutting in nature, because they have applications throughout discovery, preclinical development, clinical trials, and potentially manufacturing and postapproval market penetration.

Biomarkers are a diverse set of biological measurements that can be tissue based, serum based, or image based and have equally sweeping applications. Types of biomarkers include

- exposure biomarkers, which indicate a genetic predisposition to develop a disease;

- disease biomarkers, which measure a clinical outcome or progression of disease;

- surrogate biomarkers, which are thought to be valid substitutes for clinical outcomes or results;

- efficacy biomarkers, which measure the positive effect of the drug treatment;

- toxicity biomarkers, which measure the toxic intensity of a drug in the body; and

- target biomarkers, which detect the interaction between a drug and its desired target.

Biomarkers' infrastructural needs relate primarily to the lack of validation standards (to verify these markers provide a relevant and reliable signal of future response) and the need for more sensitive detection. At this point, speculating on the exact impact that improvements in biomarker validation and detection would have on the industry is difficult, but the most expensive phases of therapeutic drug development are prime targets improvements in this new area of technical infrastructure.

For example, establishing predictive biomarkers as surrogate endpoints could allow drug sponsors to reduce the lengths of evaluation periods in clinical trials. Moreover, biomarkers could be used to identify subpopulations of trial patients who are predisposed to be responsive to the drug or to have toxicity complications.

This information, if validated, could be used by sponsors to stratify patient populations and better target those who could benefit from the therapy. By excluding certain types of patients from trials, sponsors could lower the incidence of adverse effects and increase their approval rates, thereby getting useful therapies to market, even if they only help niche populations. Even earlier in the development process, improved sensitivity in biomarkers could allow drug makers to optimize their selection of drug candidates, avoiding compounds that show early signs of liver toxicity, for example.

2.2.5 Commercial Manufacturing

The manufacturing process for biopharmaceuticals is more complex than the process for traditional small-molecule pharmaceuticals. This TFA addresses some of the most pressing issues related to the technology infrastructure supporting manufacturing. It includes the need for better models of scale-up to ease the expensive and often unpredictable transition between clinical trial volumes and the larger quantities needed for commercial-scale production. Needs are also associated with new upstream processing platforms and bottlenecks in the downstream processing stage.

This RTI analysis considered integration and interoperability issues related to implementing Process Analytical Technology (PAT) to enable better monitoring and control throughout the production process. The PAT initiative is part of a larger regulatory guidance document released in 2002 entitled "Pharmaceutical CGMPs for the 21st Century: A Risk-Based Approach." This regulatory guidance document was developed through a collaborative effort involving FDA's CDER, the Center for Veterinary Medicine (CVM), and the Office of Regulatory Affairs (ORA). Informing this process were extensive public discussions at the FDA Science Board, the Advisory Committee for Pharmaceutical Science (ACPS), the PAT-Subcommittee of ACPS, and several scientific workshops. Discussions covered a wide range of topics, including opportunities for improving pharmaceutical manufacturing, existing barriers to innovation, and potential approaches for removing both real and perceived barriers (FDA, 2004b).

2.2.6 Postmarket Surveillance

Postmarket surveillance refers to adverse event reporting and monitoring the safety and efficacy of drugs. Concerns about safety and product quality are obviously elevated when the product is a human therapeutic. The technology infrastructure supporting postmarket activities may reduce transaction costs and lower barriers to adoption by helping alleviate those concerns.

Some drugs are given FDA approval, but FDA requires ongoing long-term clinical trials to monitor safety, efficacy, and adverse reactions in patients. Known as Phase IV testing, this postapproval clinical phase may have several objectives, such as

- comparing a drug's performance to others on the market,
- monitoring long-term effectiveness and impact on patient quality of life, and
- determining the cost-effectiveness of a drug therapy relative to alternatives.

This TFA considered potential improvements to quality assurance testing to provide confidence that products are pure and potent. It also addressed the infratechnologies that drug sponsors rely on when monitoring product performance in the marketplace.

2.3 TRENDS IN INFRASTRUCTURE RESEARCH AND DEVELOPMENT

The stages of drug development and commercialization provide a natural framework for understanding the total cost of the drug life cycle and understanding where bottlenecks or excessive costs may hinder efficiency and growth. Figure 2-2 illustrates the process and identifies several of the major technical infrastructure categories that influence each stage. Many of the same infratechnologies are used throughout the five general stages. However, the implementation of the technology is somewhat specific to each stage; hence, the infratechnology needs may be different. This figure highlights the fact that advances in screening capabilities in the preclinical stages can have significant impacts on research costs during the clinical and bulk manufacturing stages. Identifying problems earlier in the process increases the probability of a drug's success and hence reduces costs incurred from subsequent stages.

This section reviews some of the common methods employed in each stage of drug development and discusses technology infrastructure improvements under way today. The diversity of biopharmaceuticals parallels that of the technology infrastructure supporting their development. As mentioned previously, this report simplifies the presentation of the industry's diversity by focusing on therapeutic proteins. Many of the processes, technologies, and trends discussed below may be relevant for some biopharmaceutical products but not for others.

Section 2.4 complements this section by summarizing stakeholders' views on technology gaps, and Section 2.5 addresses obstacles the marketplace sets in the way of developing what otherwise would be socially beneficial advances.

Figure 2-2. Summary R&D Framework for Human Therapeutics

Stages of Biopharmaceutical[a] Development:

Supporting Infrastructure:

Potential Impacts:

5,000-10,000 compounds[b]

Basic R&D and Drug Discovery

- Basic research in biology, genomics, proteomics, and metabolomics to suggest targets
- Target identification
- Target validation
- Lead identification
- Lead optimization

- High-quality protein and gene databases
- Data integration and interoperability
- Mass spectrometry, patch clamp, and high-resolution bioimaging
- High-throughput screening, and *in silico* models
- Standards across gene expression platforms

- Lower cost of discovery
- Increased probability that a compound selected at this stage for preclinical development is ultimately successful

250 compounds[b]

Preclinical Development

- *In vitro, in vivo,* and *in silico* pharmacokinetic (ADME), toxicity studies

- High-throughput and high-content screening
- Bioimaging to monitor ADME and *in vivo* response
- Structural biology data to inform algorithms used in predictive *in silico* models
- Data formats and standards to capture experimental conditions and results

- Lower cost of development
- Increased probability that an IND coming out of this stage is ultimately successful

5 INDs[b]

Clinical Trials

- Phase I: safety
- Phase II: safety, efficacy
- Phase III: long-term safety, efficacy

- Qualification of surrogate endpoints (biomarkers and bioimaging)
- Validation of standard protocols for microarray experiments
- Good Clinical Practice (GCP) standards

- Lower cost of clinical trials
- Reduced probability that an approved drug coming out of this stage will not be recalled later (through identifying subpopulation risks)

Scale-Up

- Bioavailability studies
- Process engineering

Manufacturing

- Upstream processing involving cell culturing or fermentation
- Downstream processing, purification via chromatography
- Quality control using LC-MS, GC-MS, spectroscopy

- Validated databases compliant with FDA guidelines
- Formats and standard reporting statistics, compatible with preclinical and clinical trial data
- Sterility, protein degradation studies to inform standards for storage and handling

- Lower costs of post-market surveillance, holding quality constant
- Reduced waste due to protein denaturation or product expiration

1 approved NDA or BLA[b]

Market Penetration

- Adverse event reporting and post-market surveillance
- Handling, storage, and distribution

[a] As described in Chapter 1, this study focuses on the segment of the biopharmaceutical industry involved in producing products for therapeutic applications, including recombinant proteins, monoclonal antibodies, and gene and cell-based therapies.

[b] These numbers reflect typical ratios of initial compounds screened and Indus filed to NDAs/Biologic License Applications (Blass) approved.

Source: Pharmaceutical Research and Manufacturers of America (Parma) March 2005. *Pharmaceutical Industry Profile 2005.* Washington, DC: Parma.

2.3.1   Drug Discovery

The technical infrastructure supporting drug discovery is diverse and cross-cutting in nature. Infratechnologies used in target identification include gene expression microarrays, bioinformatics databases, and computer algorithms that search for proteins related to known drug targets that are different but perform related functions.

This stage in the development cycle involves four primary phases: target identification, target validation, lead identification, and lead optimization. Target identification marks the beginning of the discovery process. Researchers use their understanding of molecular biology, genomics, and proteomics to identify pathways to disease or disease "targets."

Once identified, a drug target must be characterized and validated through gain-and-loss-of-function studies of the target protein in the disease process. Targets are validated using animal models imitating disease states as an *in vivo* method to demonstrate the relevance of the target protein in the disease process.

In lead identification, compounds are identified that have the ability to affect the activity of a target (e.g., a protein). Various screening techniques are applied to develop a list of potential leads. Screening is followed by further testing to fully characterize each lead's properties. Leads must meet certain criteria specified by the research team, including pharmacodynamic properties (e.g., efficacy, potency, and selectivity) and pharmacokinetic properties (e.g., metabolic stability, toxicological aspects, chemical or physiological optimization potential, and patentability). Those leads demonstrating the highest probability of success—typically 250 optimized leads for every one approved drug, according to PhRMA (2005)—then enter the preclinical stage.

Lead identification relies on large compound libraries used to determine lead compounds using high throughput screening that uses solution-based biochemical assays or cell-based assays. *In silico,* or virtual screening, is an alternative technology that uses three-dimensional modeling to simulate real or predicted reactions between the target protein and lead compounds. Empirical scoring then ranks the lead compounds by how well they bind to the target protein (Giersiefen, Hilgenfeld, and Hillisch, 2003). *In silico* screening is cost-effective and allows for high throughput. However, the impact of this method is constrained by the availability of structural data for both the target protein and the ligand (compound) under investigation.

Mass spectrometry is a leading analytical technique in the drug discovery process, and rigorous calibrations are required to ensure it meets the tolerances required in biopharmaceutical research.

Proteomic techniques are becoming increasingly important in target identification. Proteins expressed at different levels between diseased and healthy tissues suggest that the protein may be a causal agent and

therefore could be a potential drug target (Giersiefen, Hilgenfeld, and Hillisch, 2003).

Although there is considerable interest in using proteomics-based approaches to drive the drug discovery process, the infancy of the infratechnologies for large-scale protein identification and quantitation presents barriers to its adoption. Recently, a large number of start-up organizations were formed with the objective of using proteomic techniques for biomarker-driven drug discovery (GeneProt and Large Scale Biology are two examples). It is telling that both of these organizations have either eliminated or greatly diminished their proteomic efforts.

Research efforts through NIH Roadmap initiatives are addressing some of the technology gaps present in drug discovery through advances in bioinformatics, bioimaging, and *in silico* modeling. To reduce costs and improve the efficiency of drug discovery and development, the pharmaceutical industry has placed a greater emphasis on early ADME and toxicology evaluation. Researchers are attempting to address ADME and toxicity issues earlier by using *in silico* methods in the lead identification and optimization processes (NIH, 2006). These methods employ a collection of infratechnologies, spanning chemical bioinformatics (e.g., chemical libraries), molecular modeling, and structure prediction.

NIST is developing standards in gene expression analysis across multiple gene expression platforms to improve reproducibility and reduce the uncertainty of experiment results across multiple expression platforms. These efforts are directed at enhancing the efficiency of target identification studies through standard reference materials and experiment protocols.

2.3.2 Preclinical Development and Testing

Once a lead compound has been identified, a number of preclinical studies must be conducted before testing the drug therapy in humans. Researchers perform experiments to determine the compound's biochemical and physiological effects and properties.

These tests are primarily performed *in vitro* or *in vivo*. Although *in vitro* studies are usually cheaper and easier to control, *in vivo* studies using animal models are better suited for understanding how the drug or

device will interact with an entire biosystem. Increasingly, testing can be conducted *in silico* using computer models to predict behavior.

At the same time that researchers are investigating the pharmacological properties of the drug candidate, they are beginning to develop the process for manufacturing the compound. For conceptual reasons, we discuss the issues involved with formulation and process design in the section on manufacturing.

Researchers use a range of analytical instruments, molecular biomarkers, assays, and other tools and technologies during the preclinical stage of drug development. In preclinical and clinical studies mass spectrometry and immunoassays are the principal detectors for quantitative analysis of drug levels in biological samples.

Another set of tools used for determining a candidate drug's efficacy and safety during the preclinical testing phase involves biomarkers. For example, alanine aminotransferase (ALT) and aspartate aminotransferase (AST) are two enzymes that have worked well in detecting toxicity in the kidney and liver (McGee, 2005d). Information received on interactions at the preclinical stage can have implications for how a candidate therapeutic will react in the human body.

Bioimaging techniques, which have traditionally played a role as a diagnostic in clinical testing, are moving into the preclinical stage. PET is being used to track the ability of a given therapy to modify tumor growth and magnetic resonance imaging (MRI) is used to study lungs, joints, and brain tissue.

Several initiatives, including those driven by the FDA Critical Path and the NIH Roadmap, have called attention to the infrastructure needs related to preclinical testing technologies; however, these are only a beginning and the technical barriers remain. Industry participants have noted that current imaging technologies may require higher throughput and better resolution to be used effectively for preclinical testing (Chapman, 2005). Thus, the Molecular Libraries and Imaging component of the NIH Roadmap has identified predictive toxicology as an important piece of the technology infrastructure needing further development.

The NIH Roadmap cites structural biology as a priority. Structural biology will make important contributions in the area of *in silico* drug design and lead optimization by providing scientists with accurate structural data of proteins representing potential targets or biomarkers. But researchers

consider the necessary computations expensive and time consuming. Innovation in structural biology could provide foundations for *in silico* drug design and predictive toxicology because they permit scientists to simulate how a given drug molecule will interact (or "dock") with the binding site of a target protein.

Other areas of need include protocols for sample preparation and for recording experimental data. In addition to documenting quantitative results of experiments, researchers must record the qualitative conditions under which the experiments were conducted.

2.3.3 Clinical Trials

After developing a promising therapy, a developer will submit an IND application to FDA. Gene therapies, cellular products, vaccines, and plasma-derived products are filed with the Center for Biologics Evaluation and Research (CBER), while protein therapies, such as monoclonal antibodies, cytokines, enzymes, and growth factors, are filed with the Center for Drug Education and Research (CDER).

Regardless of the product candidate, clinical trials are typically conducted in three phases. First, the therapy enters Phase I trials designed to determine drug safety and dosage. Specifically, investigators monitor how the drug behaves in humans using a small group of healthy individuals. Efficacy, drug interactions, and side effects are evaluated on a larger sample of patients during Phase II. In Phase III, trials are expanded to verify efficacy and monitor long-term effects in 1,000 to 5,000 patients.

A wide array of infrastructure technologies are currently used to monitor volunteers' responses during clinical trials. In addition to standard laboratory tests, the use of bioimaging and biomarkers shows promise as ways to evaluate drug safety and efficacy.

Part of a biomarker's value lies in its ability to identify patient subpopulations. In particular, a biochemical compound or substance may indicate whether an individual is likely to respond positively or adversely to treatment, and this information may be useful in shaping the composition of a clinical trial's patient population. Alternatively, if biomarkers can be linked to clinical outcomes, or endpoints, then their presence can be used as a surrogate and, in theory, allow trial lengths to be shortened. Although laboratories have measured biomarkers as part of clinical trials and used them to do internal assessments, these

markers lack the FDA validation that would be required to formally use them as surrogate endpoints.

Likewise, imaging technologies have the potential to serve as biomarkers or surrogate endpoints, but further qualification is required. The industry recognizes that bioimaging has the potential to provide quality data earlier in the trial phases. For example, various modalities are being used in cancer treatment trials to see if a tumor shrinks in response to the investigational therapy. PET is used to determine if the tumor is using less of the body's resources, while magnetic resonance spectroscopy can detect chemical changes in the tumor.

Another technology that has been discussed with regard to improving clinical trials is the microarray. For example, microarrays are being used in clinical trials to generate gene expression data to be used in evaluating drug safety and efficacy. However, microarrays have yet to achieve the necessary levels of standardization and reproducibility for their data output results to be acceptable to the regulatory community.

Finally, because of the large quantities of data that must be collected, analyzed, and presented to FDA during the approval process, information technology plays an important role in the clinical trials stage. For example, electronic data capture (EDC) technology and tools help streamline the collection and submission process.

Validation and standards are necessary for biomarkers and bioimages to be accepted by FDA as surrogate endpoints. The National Cancer Institute (NCI) calls for improving on this methodology with validated clinical trial-acceptable standards. Janet Woodcock, the FDA's Deputy Commissioner for Operations, suggests that, in general, biomarker qualification for formal use in the drug approval process will require collaborative action. Existing data must be pooled and analyzed, followed by identifying gaps in the understanding of marker performance and executing studies or add-on trials to obtain new data (Woodcock, 2005).

Such a collaborative process is under way to provide validation and standards for microarray expression data, such as the MicroArray Quality Control (MAQC) Project and the External RNA Control Consortium (ERCC). These groups seek to establish an environment for future standardization by analyzing comparability among platforms.

2.3.4   Scale-Up and Commercial Manufacturing

The manufacturing process for biopharmaceuticals varies greatly from the process for traditional small-molecule pharmaceuticals. Generally speaking, production can be divided into two major stages: scale-up and bulk manufacturing. Scale-up includes process design, product specification, and optimization production. Bulk manufacturing includes commercial-scale upstream and downstream processing. We present these two stages together in this section for ease of presentation. In practice, the scale-up manufacturing stages are developed in tandem with the preclinical and clinical development stages and commercial bulk manufacturing occurs following FDA approval.

The commercial production of biopharmaceuticals includes upstream and downstream processing. Upstream processing is defined as the fermentation process that results in the initial core product material. Once a therapeutic protein is harvested from a cell or medium, it enters the downstream processing stage. Downstream processing refers to the actual purification of the product and formulation, bottling, and labeling activities.

The primary tool used in the scale-up stage of biopharmaceutical manufacturing is chromatography. Chromatography on a laboratory scale creates homogenous samples in small amounts that enable full characterization of a lead compound. This information provides insights needed to design the upstream manufacturing process. Mass spectrometry, gel electrophoresis, and immunological assays verify purity and activity.

Process control monitoring is an important component in manufacturing recombinant proteins. The environmental conditions (e.g., temperature, acidity) of cells and surrounding media must be maintained while proteins are being expressed. Additional measurements must be made during each step of downstream processing to ensure proteins retain their activity and structure.

Bioprocess measurements currently include both off-line and on-line measurements. Off–line process measurements include biomass, substrate, and product concentration valuation. On-line measurement values are limited to basic physical quantities that include temperature; culture weight, feed rates, and pH or O_2 concentrations. Because of a lack of adequate online measurement sensors these measurements must be made off line using samples that are sent to an analytical lab.

Biomass concentration values can be generated using a dry-weight measurement, substrate concentration relies on enzymatic analysis techniques, and product concentration uses chromatography or electrophoresis to conduct measurement. Limited sampling frequency is a draw back to off-line measurement, which makes it difficult to detect real-time changes in concentration levels and prevents manufacturers from making in-line adjustments to the process. Process analytics technology and improvements in in-line measurement devices would benefit the commercial manufacturing of biopharmaceuticals such as proteins and monoclonal antibodies.

Chromatography separation is a critical technique used in the production of biopharmaceuticals to ensure homogeneity in successive output batches. The creation of the homogenous samples requires purification that typically involves three or more high-resolution chromatographic steps (Walsh, 2003). The purification protocol developed at this stage defines the procedures for pilot and full-scale commercial purification systems. A variety of chromatographic methods are available, including affinity chromatography, high-performance liquid chromatography (HPLC), and gas chromatography (GC). Purification represents a suite of tools that allow the production of a high-quality, GMP-grade product that is free of bacterial endotoxin or potentially immunogenic proteins derived from the host system.

Maintaining quality through the manufacturing process is at the heart of FDA's recent Process Analytical Technology Initiative, which has been incorporated into current Good Manufacturing Practice (cGMP) (FDA, 2004a). Such guidelines and protocols play an important role in supporting all pharmaceutical manufacturing, including biological products like recombinant proteins.

However, because of inherent differences in proteins, industry participants do not believe it is possible to standardize the production process entirely (Aldridge, 2006). For example, each protein has unique properties like hydrophilic nature, structure, and ability to bind to particular antibodies. However, it is exactly this set of characteristics that determines the best approach to purification, so a broad knowledge base must be developed to optimize production.

2.3.5 Postmarket Surveillance

The important and complex issues associated with biopharmaceutical discovery and development do not end when a product receives FDA

approval and is manufactured on a commercial scale. Postmarket surveillance, in which manufacturers and regulatory agencies monitor the quality, safety, and efficacy of a therapy, is an essential component of the process.

Once clinical trials are complete, the biopharmaceutical drug or therapy sponsor submits a drug application (i.e., NDA) or biologic license application (BLA) containing all scientific information collected during preclinical and clinical studies. Among other documents, the sponsor must submit an integrated summary of efficacy (ISE) and an integrated summary of safety (ISS), which in combination document the benefit-risk trade-off for the candidate. For some therapies, FDA will approve the drug provisionally and will require additional "Phase IV" studies to evaluate long-term effects. Even if approval is not made conditional, sponsors must produce a Periodic Safety Update Reports (PSUR) quarterly for the first 3 years and annually in years that follow. These updates incorporate adverse drug reaction (ADR) cases and alert FDA to potential safety issues that may prompt a reevaluation of therapy labeling.

Such postmarket surveillance for safety and efficacy, also known as *pharmacovigilance*, will always be necessary despite even the most rigorous preclinical and clinical testing phase evaluations. These earlier phases usually will have involved only a few thousand human patients, and certain therapy responses or interactions are so statistically rare that they only manifest once the therapy reaches the large market population. The costs of collecting such data and producing these reports can be sizeable, but the private and social costs of an unsafe therapy that must be recalled or remarketed under restricted criteria are much more substantial.

A concern related to the postapproval performance of a biopharmaceutical involves stability of the therapy and the conditions under which the therapy must be handled, stored, and dispensed to remain effective. Unlike traditional small-molecule pharmaceuticals that are usually dispensed in pill form, biopharmaceuticals are almost exclusively administered via injection.

Proteins by their nature can be unstable and subject to denaturation (i.e., unfolding of the protein structure) when lyophilized or exposed to extreme temperatures or agitation during shipping. Some therapies are

subject to light degradation or adherence to glass or plastic containers, which limits effectiveness.

There is also the issue of shelf life. Proteins stored in liquid form are less stable and have considerably shorter shelf lives than traditional pharmaceuticals. They can show greater stability if stored in a freeze-dried form for later reconstitution, but care must be taken during the freeze-drying process to prevent degradation and ensure proper protein folding upon reconstitution. DNA material, tissues, cell-based products, and other biologics are also sensitive to improper storage and handling.

Many of the technical methods used to evaluate safety and efficacy of a biopharmaceutical during clinical trials can also be used to monitor its safety and efficacy in the marketplace following approval. That is, the same laboratory tests and techniques can be used to determine whether a patient responds favorably to treatment. Drug manufacturers do not intend to document the response of every patient; instead, they rely on reports of ADRs from consumers and health care providers.

In the case of monoclonal antibodies, there will be a significant need for improved immunogenic testing to evaluate patients' immune system response to the foreign substances measured by the production of antibodies specific to the biopharmaceutical. If an immunogenic response is triggered, the biopharmaceuticals efficacy may be reduced or mitigated entirely. Standardized immunogenic testing would provide reliable and comparable information on patients' response to treatment and help mitigate this measurement technology barrier to the use of monoclonal antibodies.

Validated databases, managed either by the drug/therapy sponsor or an outsourced provider, are essential. These must comply with FDA's electronic record requirements outlined in Title 21 CFR Part 11 and the International Conference on Harmonization (ICH) E2B data guidelines on clinical safety data management. Adverse experiences collected into these databases must also be submitted to FDA's Adverse Event Reporting System, an E2B-compliance database that also tracks ADRs reported by consumers and health care professionals via MedWatch.

The primary infrastructure needs associated with postmarket surveillance involve standard data formats, data management protocols, and statistical analysis models. A certain degree of interoperability is required to submit data collected by a biopharmaceutical firm's or subcontracted organization's to FDA's safety surveillance databases. In addition,

conversations with a postmarket drug safety services provider suggest that time and cost could be saved if the data standards used postapproval were compatible with those used in clinical trials and even preclinical stages. Some suggestions are provided by the following FDA's guidelines for industry: "Development and Use of Risk Minimization Action Plans" and "Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment."

Finally, it may be important to provide the infrastructure for data management and reporting today so that these systems will be able to accommodate the increasing volumes of data that may be processed in the future. The concept of personalized medicine, though many years away from implementation in a postmarket setting, may eventually require tracking of detailed genetic information about patients—both in the clinical trial phases and in the marketplace when adverse reactions occur (Keating, 2003).

Health care providers often have developed their own practices for maintaining stability of biopharmaceuticals. For example, they may determine that the maximum time they are willing to store a reconstituted protein is 72 hours after preparation. However, Crommelin and Sindelar (2002) argue that these guidelines are somewhat arbitrary because, in theory, a sterile product reconstituted under asceptic conditions should remain sterile for a longer period of time (depending on other compatibility issues). Thus, they identify the potential need for sterility studies to determine the acceptable storage period for a protein product once reconstituted—studies that would reduce unnecessary waste (Crommelin and Sindelar, 2002).

Aggregation monitoring studies are also needed during the manufacturing process and as part of a stability testing program. Aggregation is one factor that has been commonly associated with the biopharmaceutical immunogenicity. Aggregation occurs when proteins interact to form large clusters of molecules. Several methods exist to characterize and analyze the level of protein or viral vector aggregation with a biopharmaceutical solution. However, no standardized measurement system exists for conducting aggregation studies for existing or future biopharmaceuticals. Improvements to aggregation methods could minimize the risk of immunogenicity as well as improve our understanding of the effects of environmental conditions during biopharmaceutical manufacturing and storage.

2.3.6 Relevant Federal Technology Infrastructure Development Efforts

A number of federal agencies have funded public investments in technology infrastructure, including the National Science Foundation (NSF), NIH, FDA, and NIST. Some of these organizations have primarily contributed to the basic science underlying the industry's technology platforms, while others have been more active in supporting generic technology and infratechnology research.

NSF's and NIH's investments have enriched the basic science and applied research disciplines that underlie the biotechnology industry. In addition to biology and chemistry research that informs drug discovery, NSF's support of engineering, physics, and computing contributes to the technical infrastructure knowledge base (e.g., for mass spectrometry, imaging, and microarray production). Much of this support is through funding of university researchers, who also receive private and state funding for these activities. NIH's research emphasis is more application-specific because it concentrates on understanding disease mechanisms. Some of its institutes, such as the National Institute for Biomedical Imaging and Bioengineering (NIBIB), also conduct generic technology research that the biopharmaceutical industry uses as new technology platforms that lead to innovative drugs through private applied R&D.

In 2002, NIH took a step further in addressing the needs of the medical industry, including biopharmaceuticals, with its Roadmap for Medical Research. The overall goal of the roadmap is the promotion of *translational research,* which focuses on bringing basic research results more quickly to clinical settings. Several of the Roadmap's initiatives address gaps in the technology infrastructure along the path to drug discovery and therapy development. The total budget for the NIH Roadmap's stated research initiatives has grown from $239,716,000 in FY2005 to an appropriation of $329,462,000 for FY2006.

Complementing the NIH Roadmap initiative for translational research is FDA's initiative for *critical-path research,* which emphasizes efforts to improve the product development process itself through better evaluation tools. The most recent budget request for FY2007 calls for $5.94 million (out of a total of $1.95 billion requested by the agency) to support its Critical Path Initiative. These funds are intended to support partnerships and targeted research to "modernize medical product development" (FDA, 2006). Examples of these partnerships include a medical imaging

initiative to validate PET as a surrogate endpoint in cancer drug clinical trials and an electrocardiogram data warehouse to help address safety issues during trials and in postmarket review.

The technology research being conducted through the NIH Roadmap and FDA Critical Path Initiative calls for supporting infrastructure investments, particularly in measurement, data, standards, and models. As discussed in the following two sections, NIST could play an important role in developing these infratechnologies through its Chemical Science and Technology Laboratory (CSTL) and vast expertise in developing standards.

2.4 TECHNICAL BARRIERS TO AN IMPROVED INFRASTRUCTURE

This section synthesizes the comments study participants made during in-depth interviews when asked about technology gaps that inhibit their work currently and technical problems they foresee on the horizon. Chapter 4 discusses primary data collection efforts in detail, but stakeholder comments are presented in this earlier section to articulate views on the infrastructural needs in advance of developing the cost model that quantifies potential efficiency gains.

2.4.1 Bioimaging

Increasingly, FDA is requesting imaging studies to be included in the regulatory submissions for new drug approvals. However, respondents indicated that there is inadequate technical infrastructure in place to support a large-scale adoption of imaging in preclinical and clinical trials industry-wide. Bioimaging experts interviewed during this study identified technical barriers and efforts to overcome them. These included the following:

- **Internal Standardization Efforts.** Maintaining consistency in the processes to capture imaging data is critical to the validity of imaging study results. Firms must develop calibration procedures that are consistent across multiple imaging devices, among images taken by the same device, and images taken across time.

  For example, a Phase III medical imaging study may include 1,000 patients at 100 different clinical sites. A dispersed data collection strategy demands strict adherence to study protocols for calibration or imaging equipment as well as anatomical positioning of research subjects.

- **Internal Evaluation and Validation.** The software used to interpret images has not been able to keep pace with the rate of assay development in high-content screening. Experts stated that imaging device manufacturers have a "box" mentality and they have not provided adequate image interpretation.

  This is particularly important in molecular imaging where no generic software tool can be applied universally to all cell-based assays. Commercially available systems are only suitable for a small number of assay types, leading researchers to develop customized computational algorithms to interpret images and meet their specific needs.

  One scientist interviewed suggested that developing an automated computer algorithm for a specific assay type could take anywhere from 1 week to 1 year to complete depending on the complexity of the assay developed and the cell type.

- **Quality Control Programs and Tools.** Researchers spend time aligning data definitions in interpretation software because of differing taxonomies surrounding medical or anatomical regions of images (e.g., a specific cell, receptor, or lobe of the brain).

- **Reagents.** Imaging assays may require contrast agents to reveal bioactivity at the drug target site or to follow the drug as it moves through the body. These reagents must be purchased or produced internally. Bioimaging reagents—chemical solutions that carry out chemical or enzymatic reactions—require advanced chemistries that may be underdeveloped within an organization.

- **Data Transfer and Archival Systems.** Imaging data are transferred from clinics or labs to several different sites for analysis, interpretation, and validation. Currently, no online digital transfer system exists to securely send imaging data across the Internet. As a result, firms rely on couriers to transfer digital imaging data from the clinical sites to their internal sites for analysis.

- **Data Management in a Regulatory Environment.** Long-term data storage and retrieval systems must be created to archive imaging data. Current image file size is between 500 megabytes (MBs) and 10 gigabytes (GBs). File sizes will trend toward 120 GB over the next few years.

  Storage capacity quickly becomes an issue when dealing with large files in a regulatory environment. Image files that are changed or manipulated are always saved as new files to maintain transparency and documentation for FDA review, thereby multiplying storage requirements.

- **Industry Comparability and Testing Programs and Standardization Efforts.** Researchers follow or participate in efforts that seek to study and evaluate comparability issues or that seek to establish industry standards. For example, Digital Imaging and Communications in Medicine (DIACOM) is an industry committee working to standardize clinical imaging data formats and transfer procedures. Differences in transfer

procedures lead to data comparability and integration during later analysis stages.

2.4.2 Gene and Protein Expression Analysis

Study participants provided consistent accounts of the benefits and challenges of gene expression and proteomics analyses, particularly in an era of production-style genomics studies. Problems often arise from a lack of consistent agreement among technology platforms (e.g., microarrays vs. flow cytometry), manufacturers (e.g., Affymetrix v. Agilent), laboratories, time, and protocols, among other factors. One interviewee characterized the situation as "distressing" because the platforms and laboratory infrastructure are expensive; however, the implications of carrying poor data forward in the discovery process dwarf those costs.

In the emerging genomic era, firms frequently find themselves in a position that is at odds with classical statistical studies: studies used to have few variables but many replicates. Although commoditization is bringing microarray experiment costs down, they remain sufficiently expensive at $500 to $1,000 per assay that few replications are run (O'Connor et al., 2007). Microarrays yield huge volumes of data variables, but cost limitations prevent a large number of data points for each variable.

Internal standardization initiatives at many firms are driven by the need to increase the confidence researchers have in the data output because of the many assumptions that must be made. Some of these assumptions are as follows:

- the concentration of the gene of interest in the sample is sufficient,
- there is little competition between the genes of interest and the probes on the microarray,
- there is a linear relationship between fluorescence and gene expression,
- the microarray is accurate and true, and
- the microarray scanner is calibrated correctly and reading the fluorescence correctly, among many other assumptions.

Differences between assays emerge because laboratories (even within the same organization) may have different equipment, use different calibration and assay protocols, or run different analyses on data output using different analytical methods. Other factors that complicate the reproducibility and reliability of results include the volatile nature of

enzymes, tissue degradation, and the amount of noise inherent in gene expression assays. Gene expression studies are subjective, and human error can be introduced into the experiment during multiple steps.

Firms are aware of the above limitations to gene expression studies and have responded by standardizing internally and investing in tools and programs that increase their confidence and assurance in the accuracy and quality of their experiments.

Interviewees discussed the following efforts:

- **Internal Standardization Efforts.** Most firms have established uniform internal protocols and operating procedures that labs are required to follow that dictate proper sample preparation, handling, assay, and data output procedures. However, these often differ among firms, inhibiting collaboration.

- **Internal Evaluation and Validation.** Formal and informal committees review processes, materials, and protocols to evaluate or validate best practices. For example, one firm identified 15 different analytical methods that were appropriate for evaluating microarray data output. Because each method yielded different results, the firm sponsored a project that evaluated all methods, selecting the most appropriate method. Standards would have greatly facilitated the process by reducing the number of replicated assays needed. Another example includes a firm that analyzes each microarray it receives from its supplier for "dim spots" and other quality defects.

- **Quality Control Programs and Tools.** Firms research, develop, implement, and maintain quality control programs and tools. Tools include algorithms that evaluate data quality and Web-based tools that track real-time data on analyses. This activity category would exist even in the absence of standards; however, respondents indicated that efforts would be more productive and efficient if industry standards were established. One company develops new quality control metrics for each new array type it receives.

- **Reference Materials.** Many firms purchase from a limited number of providers or develop in-house reference materials that are "spiked into" samples. These reference materials have known values, which can be used as benchmarks. However, the materials that are available do not reflect "real-world" sample complexity, even though they provide researchers with some degree of confidence in the accuracy of their instruments.

- **Industry Comparability and Testing Programs and Standardization Efforts.** Researchers follow or participate in efforts that seek to study and evaluate comparability issues or that seek to establish industry standards. For example, the Microarray Quality Control (MAQC) Project is evaluating the comparability of analysis results across microarray platforms (brands), laboratories, and time.

Table 2-1 illustrates the labor effort one large microarray laboratory devotes annually to its technology infrastructure. Inadequate industry standardization requires technical staff to proactively manage data quality and assurance risks on an ongoing basis. In addition, senior research staff attend conferences on standardization and devise methods for managing quality.

The lab foresaw no changes to its quality control staff given that this function would exist regardless of the scope and quality of industry standardization. However, industry standards for protocols, sample management, and analysis methods and tools, such as standard reference materials, would lower resources assigned to positions that contribute to or use gene expression analyses.

2.4.3 Bioinformatics and *In Silico* Predictive Modeling

The industry's massive data production capacity relies on information technologies to manage and analyze data output, transmission and storage. In an ideal world, information flowing from external and internal sources, and across development stages (where possible), would converge in software applications that would evaluate drug candidates on an ongoing basis.

However, interviewees reported that in reality data production capacity in the posthuman genome era outstrips the industry's ability to manage, coordinate, analyze, and communicate the resulting data across software systems and technology platforms. Yet these analyses inform decision making on a daily basis.

Table 2-1. Potential Labor Allocation Involved in Quality Control at One Microarray Laboratory with an Improved Technology Infrastructure

Position	Current Staffing Allocation (FTE)	Allocation with Improved Standardization (FTE)	Employee Description
Quality Control Manager	1	1	Senior Ph.D. scientist
Laboratory Director	0.5	0.25	Ph.D. scientist
Informaticist	1	0.5	M.A. scientist
Laboratory Manager	0.3	0	B.A., with 10 years' experience
Director of Bio Stats	0.5	0	Senior Ph.D. scientist
Director of R&D	0.5	0	Senior Ph.D. scientist

Visualizing and combining high-throughput data streams from multiple sources, including legacy data, is the standardization dilemma most frequently mentioned during interviews. Problems often arise when combining multiple data streams from different sources, even within the same biopharmaceutical company. Legacy data issues frustrate organizations with large data repositories. In many respects, start-up firms are better off because they can build systems from the ground up that use the latest in information architecture theories.

Current efforts are fundamentally about the ability to exchange, interpret, and analyze data. In the words of one interviewee, a lack of standards "creates a constant struggle." Nearly all interviewees voiced a need for "middle layer" software tools to communicate data between systems, and standardized ontologies do not simply describe the data but make statements about them and their relationship to other data. Stakeholders offered the following insights:

- **Internal Standardization Efforts.** Expenditures on internal standardization efforts include establishing standard analytical processes, evaluating and validating existing and proposed systems, and troubleshooting inconsistencies in data design and analysis across organizational boundaries.

- **Data Translation and Communication Tools.** At least two of the firms interviewed cited the use of customized software tools to communicate and exchange data between database systems. One company is working with its software vendor, while the other is developing its own in-house tools.

  Data become locked into definitions and formats when they are transferred from industrial-scale data production tools like sequencers and flow cytometers into software applications. New instrumentation, competing software platforms, and differences in how researchers prefer to define data elements combine to impede information sharing.

- **Industry Comparability and Testing Programs and Standardization Efforts.** Firms fund senior research staff to participate or follow industry standardization efforts. The majority of interorganizational interaction is through ad-hoc meetings and semiformal working groups comprising academics and industry researchers. One interviewee characterized these endeavors as "grass roots efforts" undertaken by groups "on the cutting edge of research."

Integrating data is difficult because the processes that capture that data may have defined the sample parameters differently or may have used different terms to describe variables. One senior scientist in bioinformatics stated that it could take 1 to 2 weeks to capture data from gene expression, 3 to 4 weeks from protein expression, a few weeks for

text mining, and then an additional 3 to 4 weeks to integrate the information.

2.4.4 Molecular Biomarkers

When discussing activities and investments that comprise this category of infrastructure spending, stakeholders offered the following insights:

- **Internal Evaluation and Validation.** Firms must apply gene or protein expression profiling techniques to validate potential biomarkers that may be associated with disease or drug targets. Following biomarker discovery, firms develop assays to test biomarkers' association with a disease or drug target's response. The outcomes of the association studies must then be externally validated.

- **Quality Control Programs and Tools.** In the biomarker assay development process, firms must develop internal standard operating procedures (SOPs) and good laboratory practices (GLPs) associated with conducting biomarker assays.

  Similar to gene expression, firms must validate the technology platform and the reagents used in the biomarker assay. The large data outputs generated from these assays require firms to identify potentially new analytical techniques capable of eliciting information from large data sets. Some examples include cluster profiling, self mapping, or other relationship or pattern recognition analytical methods.

- **Industry Comparability and Testing Programs and Standardization Efforts.** Firms participate in industry consortia that work to resolve interplatform differences, comparability of data sets, sample quality, user variability, and data analysis methodologies. These standards are often adopted by peer-reviewed journals as requirements for publication of biomarker discovery experiments.

  Currently there is no standard or accepted process for validating biomarker association studies. Validation methods include replicating association tests with an independent patient and control samples or employing a second technology platform to confirm the initial association findings (Ginsburg and Haga, 2006).

2.4.5 Commercial Manufacturing

Participants identified that excess technology infrastructure expenditures in commercial manufacturing arise from data verification, validation, and QA/QC processes. Related to these technology infrastructures, problems may arise from data transfer from the development lab bench to production, data assembly, and interpretation at manufacturing facilities. Thus, an inadequate data infrastructure impedes the efficient transfer of scientific, process, and regulatory compliance data from R&D to

manufacturing. These data transfer complications can result in delays in production and additional labor costs associated with tracking down or regenerating lost information.

- **Data Transfer from Lab to Production.** Batch testing during preproduction scale-up is costly. Processes that do not transfer from the lab to the production floor result in multiple trial batches produced before the process is perfected. During process scale-up, batches are tested to ensure they meet the development specifications. Batches failing during production can cost as much as $1 million each.

  One participant provided the example of cell culture testing. Current cell culture development may take 5 to 10 days to attain sufficient yields. Process engineers then need to know what factors or variables to adjust in the process to improve the yields. It then takes another 5 to 10 days for samples taken from the cell culture to be sent to an analytical department for analysis. Once results of the analyses are returned, process engineers repeat the previous steps until the process is perfected.

- **Internal Data Integration.** The labor and time associated with data transfer, integration, validation, and continued process monitoring represent significant cost and productivity losses during the manufacturing stages of the biopharmaceutical life cycle.

  Much of the industry relies on unstructured and disparate data systems. Such inadequate information infrastructure creates inefficiencies and costs due to labor and time associated with tracking down information, conducting redundant tests, and replicating analysis.

- **Supporting Multiple Proprietary Data Formats.** One line manager interviewed said that systems and equipment original equipment manufacturers (OEMs) are at the root of a significant share of their plants' data integration costs. Manufacturing plants include numerous measurement instruments and types of processing equipment. OEMs develop their own proprietary data formats for both process monitoring and diagnostic data on the instrument itself.

  OEMs will only support the maintenance of purchased equipment if OEM proprietary data formats are supported continuously by the biopharmaceutical manufacturing plants. As a result, the plant often has redundant data systems, one to support the proprietary data formats of the OEM and a second system to integrate the multiple OEM data formats across the manufacturing processes.

Manufacturing operations operate in a regulatory environment that requires validation of manufacturing processes. The data infrastructure inadequacies increase the cost of demonstrating regulatory compliance as well as increasing manufacturing costs by limiting firms'

understanding of their manufacturing processes and constraining their ability to efficiently conduct process monitoring.

2.4.6 Postmarket Surveillance

Industry participants spend a considerable amount of time conducting surveillance, tracking adverse events, and filing Periodic Safety Update Reports for approved biopharmaceuticals. Advanced techniques for managing and integrating the data from these activities have the potential to reduce the costs of these postmarket surveillance activities, as well as provide the health care community with assurances that any unanticipated risks are detected as quickly as possible.

Thus, technical barriers related to postmarket surveillance were primarily associated with data collection, data entry and analysis, quality control, data systems design and validation activities, training on reporting adverse events, and product relabeling activities motivated by safety issues. Specific examples include the following:

- **Data Collection and Integration.** Data on adverse event reporting and postmarket safety studies must be collected across regional sites that may have different data formats, have data collection forms, and potentially include multiple languages. Firms dedicate labor and resources to developing reporting protocols and training clinicians on data entry. In addition to multiple sites reporting adverse events, a variety of data collection mechanisms may be used to submit data. Collection methods may include paper forms, automated call centers, and electronic data capture devices. Lack of a common data dictionary across sites may result in errors or delays due to misreporting of data.

  These factors require a staff of data entry and quality control technicians to manually enter paper reports into centralized databases and perform quality checks to ensure full integration of reported data with consistency across data elements.

- **Analysis.** FDA requires periodic safety update reporting. Adverse events are investigated and the adverse event data must be compiled and analyzed by statisticians and a scientific panel.

- **Validation of systems.** Data systems infrastructure must be constructed and validated. This can cost in the millions of dollars and requires 6 to 12 months to develop. Once in place, the core database and system must be validated through rigorous initial testing; testing is continued periodically to ensure functionality.

2.5 MARKET BARRIERS TO AN IMPROVED INFRASTRUCTURE

Constraints familiar to researchers in the private sector hamper the development of the technology infrastructure, particularly the development of infratechnologies. Industry stakeholders discussed that, although overcoming technical barriers holds significant economic potential for the industry, resource and time constraints combine with competing investment priorities and poor pay-off ratios to inhibit private-sector investment.

Interviewees stated that industry has been interested in improving infratechnologies for some time, and although improvements have been made in the last decade, much work remains to be done. Several semiformal groups have formed to strategize these issues, including ones with some measure of representation from NIST, but significant market barriers persist. These barriers are rooted in the following:

- **Intellectual property rights:** Biopharmaceutical companies guard their intellectual property positions fiercely. As a result, when these companies participate in research consortia that address infrastructure issues, they can be reluctant to share technical knowledge, even when it has an infrastructure character. Companies also may question the motives of coalitions and industry groups that advocate industry-wide adoption of a standard set of procedures, protocols, ontologies, or data formats.

- **Financial and time commitment:** Most efforts are currently "grass roots" efforts with inconsistent funds and schedules. Business and technology life cycle effects erode progress on topics that require dedication and long-term investment. If it takes 5 to 10 years, at one group's current pace, to develop a standard, will that standard be relevant when it is ready? More funding and commitment of personnel by participating companies would allow acceleration of standards development, but the public-good nature of the needed infratechnologies reduces investment incentives.

- **Coordination Issues:** Multiple experts interviewed told of instances in which one semiformal group was working on issues concurrently with another, without either being aware of the other's progress. The extent of duplication of effort between the two groups was not fully known, but such occurrences illustrate the importance of coordination.

In summary, "ad-hoc" standardization groups often duplicate each other's efforts, "bumping into each other" in the research community. These groups also face hurdles posed by the business cycle, competing R&D priorities, and the length of commitment (of both time and

resources) required to enact real progress on standardization. Also, because participating firms tend to rotate staff, these ad hoc groups lack continuity and institutional memory.

Because of the public-good nature of technical infrastructure (its economic value comes from common and uniform use), economic theory suggests that individual firms will underinvest in such infrastructure because private returns to needed investments in infratechnologies are either low or negligible for individual companies. Lacking complete and equitable participation by all stakeholders means that some portion of the industry will "free ride" on the investments of others. This is the case despite collective (industry-level) private benefits resulting from investments in the technology infrastructure. In general, if innovators do not believe they will gainfully appropriate economic returns and control the generation of economic value from the knowledge they developed, they will likely avoid the risk inherent in infrastructure research. The only way for individual companies to capture an adequate return on investment in infratechnologies is to use them as proprietary protocols. Not only is this inefficient from an investment perspective but the demand (buyer) side of the market objects strenuously to the lack of uniform industry standards (Tassey, 2000, 2005).

Industry stakeholders interviewed over the course of this project saw a natural role for NIST to accelerate the development and further the sophistication of the biopharmaceutical industry's technology infrastructure, given NIST's reputation and NIST's expertise in standardization and chemical and biological sciences.

3 Economic Analysis Methodology

The economic analysis has two objectives: to estimate current private-sector expenditures on biopharmaceutical technology infrastructure and to estimate the efficiency gains an improved technology infrastructure could yield the industry.

The first objective entailed surveying the industry to capture spending and then extrapolating the results to estimate industry-level expenditures. To meet the second objective, RTI leveraged recent drug development cost studies and built an economic model that recalculated the cost of bringing a new drug to market. The model included both the cost of failed INDs and the time value of money. Expert opinions on feasible improvements in costs, stage length, and the probability of success were used to estimate counterfactual cost-model scenarios.

The cost model offers two output categories:

- the change in actual R&D expenditures, per IND and per FDA-approved drug and
- the change in the present value of R&D expenditures. Including and excluding the cost of failures.

The time value of money concept—where $1 today is worth less in real terms than the same nominal $1 was 10 years ago—takes into account investment options, inflation, and other time-based factors affecting the value of money. The discount rate used to calculate the present value can also be used to adjust for risk and uncertainty. Compressing the schedule of the drug development cycle greatly affects the current, or present, value of the cost of drug development. Thus, the model was developed to show both actual and present-value savings.

This chapter presents the theory and technical approach to meeting these two objectives. It begins by reviewing existing cost studies, including a 2007 DiMasi and Grabowski paper on biopharmaceutical drug development whose analytical results were used as the baseline parameters in the cost model.

3.1 EXISTING BIOPHARMACEUTICAL R&D COST STUDIES

Few published studies focus exclusively on the R&D cost of biopharmaceutical drugs. Consequently, the model employed in this study referenced previous models constructed for traditional pharmaceuticals, particularly a series of studies conducted by the Tufts Center for the Study of Drug Development (CSDD) (DiMasi et al., 1991; 2003).[1] These studies were updated in 2007 to account for differences in biotechnology drug R&D (DiMasi and Grabowski, 2007).

3.1.1 Small-Molecule Pharmaceutical Development Cost Studies

Several research groups have attempted to estimate the cost of drug development over the years. The three groups whose work was most relevant to this analysis were the CSDD at Tufts University, the Boston Consulting Group (BCG), and the Global Alliance for TB Drug Development. Each study estimated the total cost of bringing a single traditional (small-molecule) pharmaceutical drug to market.

Cost estimates ranged between $240 million and $880 million (cumulative in 2000 dollars) per drug developed over a period spanning 8 to 15 years. Discovery research and clinical trials were consistently the two most expensive stages of R&D, while preclinical represented the smallest share of total cost. Table 3-1 summarizes the cost estimates for each study across the R&D stages.

The total cost estimates varied in absolute terms, and the different definitions of cost categories make it difficult to compare estimates across studies. However, the issue of failure rates and the R&D stage at which failure occurs has a significant impact on both the total and relative costs measured. The three studies have significantly different motivations and therefore inherent structural characteristics that must be considered when comparing their results presented in Table 3-1.

---

[1]See also DiMasi (2002) and Adams and Brantner (2006).

Table 3-1. Summary of Cost Studies: Estimated Absolute Costs and Relative Share by R&D Stage

R&D Stage	Tufts[a] (millions)	BCG[b]	TB Alliance[c]
Discovery	N/A	$530 (60%)	$60 (25%)
Preclinical[d]	$121 (30%)	$90 (10%)	$10 (4%)
Clinical	$282 (70%)	$260 (30%)	$26 (11%)
Failure[e]	N/A	N/A	$144 (60%)
Total Cost Per Drug	**$403 (100%)**	**$880 (100%)**	**$240 (100%)**

N/A = Not reported as separate costs

[a]Source: DiMasi, Joseph A., Ronald W. Hansen, and Henry G. Grabowski. 2003. "The Price of Innovation: New Estimates of Drug Development Costs." *Journal of Health Economics* 22:151-185.

[b]Source: Tollman, Peter, Philipe Guy, Jill Altshuler, Alastair Flanagan, and Michael Steiner. 2001. "A Revolution in R&D: How Genomics and Genetics are Transforming the Biopharmaceutical Industry." Boston: The Boston Consulting Group. Available at: http://www.bcg.com/publications/files/eng_genomicsgenetics_rep_11_01.pdf. Obtained on January 12, 2006.

[c]Source: Global Alliance for TB Drug Development. 2001. "The Economics of TB Drug Development." Nancy Pekar (ed.) Available at: http://www.tballiance.org/3_2_C_BalancingIncentivesandAccess.asp. As obtained on January 10, 2006.

[d]Tufts preclinical cost includes the costs of drug discovery research.

[e]Tufts and BCG incorporated failure costs into the costs by R&D stages.

The CSDD at Tufts, a research group funded by members of the pharmaceutical industry, focused on the cost of late-stage clinical trials. The study is based on a dataset of drug candidates in clinical trials. For this reason, the clinical-stage costs are highly detailed with low uncertainty, while a more approximate approach was employed to quantify the cost estimates in the discovery and preclinical stages using annual aggregated R&D expenditure data over a 20-year period.

The BCG report is primarily focused on new genomic approaches and demonstrating the potential cost savings that these technologies might bring to drug development. As discussed earlier, most benefits from genomics and proteomics research trajectories would occur in the earlier stages of the drug development process. Given the focus of the BCG report, costs may be overreported in the earlier stages of development.

Finally, the Global Alliance report (the most detailed in cost categories) was focused on estimating the development cost of an orally ingested TB drug using a conventional small-molecule approach. The costs related to small-molecule drug development can be significantly different from the costs of biopharmaceuticals using protein-based approaches to development.

3.1.2 Large-Molecule Biopharmaceutical Development Cost Studies

It is not the intent of this study to estimate the cost of developing a new biopharmaceutical product, but rather to estimate the potential an improved infrastructure holds for the industry. Thus, this study relies heavily on a recently completed study by DiMasi and Grabowski at CSDD (2007). Given the importance of this study to results presented in Chapter 5, the DiMasi and Grabowski study is discussed in detail below.

Recent studies have suggested that biopharmaceuticals may cost more in preclinical stages and may be more difficult to manufacture, but that their clinical trials require fewer patients and have higher success rates (see Reichert [2004]). Compared with a traditional success rate around 20%, the Tufts CSDD reports 30% to 35% success rates for rDNA and mAbs that entered clinical trials between 1990 and 1997 (Tufts, 2005). Reichert has also published summary data on stage lengths for these classes of biopharmaceuticals.

Biopharmaceutical trials have fewer subjects and fewer studies compared with small-molecule drugs. Some biopharmaceutical trials are relatively small because they are targeting diseases with small patient populations. In addition, biotechs target hard-to-treat illnesses, which means that even moderate efficacy may induce FDA to approve such drugs based on smaller trials. Finally, because major classes of disease exhibit distinctly different levels of difficulty in proving efficacy, the size and duration of clinical trials vary accordingly.

The baseline drug development cost estimate was $1,241 million (DiMasi and Grabowski, 2007). The results were presented at the BIO2007 international convention in Boston, Massachusetts. This study in *Managerial and Decisions Economics* represents the first published study that estimates the R&D costs related to biopharmaceuticals.

DiMasi and Grabowski computed their total capitalized R&D cost estimate using data for 17 compounds[2] tested in humans between 1990 and 2003. The project cost data on these 17 compounds come from two sources: data on four compounds are from a previous study (DiMasi et

---

[2]The authors do not provide a distribution by recombinant proteins and monoclonal antibodies.

al., 2003),[3] while project data for the remaining 13 compounds were provided by a single biotech firm.[4]

In addition to detailed expenditures for the 17 compounds, the authors' analysis used a much larger data set to estimate average development times, clinical success rates (receiving FDA approval), and clinical stage transition probabilities. This information was pulled from the Tufts CSDD database[5] for 522 compounds that included 278 therapeutic recombinant proteins and 244 monoclonal antibodies tested in humans between 1990 and 2003.

Components of the Total Capitalized R&D Cost Estimate

The full cost of drug development includes not only the direct costs for successful INDs, but also the net costs associated with failures (attrition rates) and the opportunity cost of engaging in long-term development cycles. To characterize the full costs of drug development, the following information is needed:

1. actual (i.e., out-of-pocket or direct expenditures) costs per IND for the preclinical period,

2. actual costs per IND for the clinical period,

3. clinical success and stage attrition rates, and

4. the discount rate used to capitalize the time series of costs over the duration of the development process up to the date of market approval.

Preclinical development costs are all R&D expenditures incurred before submitting an IND application and the entry of a candidate therapy into clinical trials. These include expenditures on disease mechanism research, screening, animal models, and PK/PD tests, as well as process engineering studies and the production of small volumes of biological product to use in preclinical testing.

Subsequently, the clinical period covers the time starting with first-in-human testing in Phase I up to the market approval date. Also included in the cost estimates are costs incurred during the clinical period for long-

---

[3]Four recombinant proteins and two monoclonal antibodies part of a larger sample of 68 compounds from 10 multinational pharmaceutical firms.

[4]Biotech firms provided detailed project data for 13 compounds as part of a consulting project. The premise of the project was to test the firms' hypotheses that their R&D costs were lower than the estimated $802 million for traditional pharmaceutical firms from DiMasi et al. (2003)

[5]The Tufts CSDD database aggregates information from a collection of business intelligence databases, trade press news, company annual reports, and company Web sites.

term animal model testing conducted during that period; regulatory submission costs; and chemistry, manufacturing, and control costs. Table 3-2 summarizes baseline R&D costs used in this study, which are reported in constant 2005 dollars.

Stage Costs per IND. Costs for the preclinical stage were developed using times series data on aggregate annual R&D costs per drug. Clinical stage costs for clinical periods were calculated using the actual expenditure data from the project dataset for the 17 compounds for Phase I, Phase II, and Phase III (see DiMasi et al. [2007]).

Transition Probabilities and Clinical Success Rates. Transition probabilities reflect the probability of any one IND progressing from one clinical trial phase to the next, which is conditional on the IND's success in the current phase. Table 3-2 illustrates how biopharmaceutical companies apply a portfolio approach to derive the estimated cost per approved drug when it is known that not all INDs will be successful and it is not known in which stages unsuccessful INDs will fail.

As shown in Table 3-2, (A) is the average actual (sometimes called "out of pocket") cost a firm incurs for an IND that progresses through the entire stage. Some INDs will progress through all stages and, therefore, incur all the stage costs, while others will fail after proceeding through perhaps one stage or a portion of a stage. Transition probabilities are calculated by multiplying (B) the percentage entering each phase by (C) the probability of success in that same phase. Transition probabilities have a cumulative effect on the share of INDs that progress to the next clinical stage.

The computation is as follows: the study's authors multiply the actual stage cost (A) by the probability of an IND entering the stage (B), which represents the expected outlay per IND in that stage. The sum of the values in (D) represents the cumulative expected stage costs for the IND. Knowing some INDs will fail and being able to project failure rates based on historical information, a biopharmaceutical firm would expect the outlay for any one IND to complete Phase III trials to be $169.0 million.

A clinical success rate of 30.2% (E) reflects the overall probability that an IND that enters clinical trials will eventually receive FDA approval. Dividing $169.0 million by 30.2% yields $559.6 million; this is the estimated cost per approved drug. The successful drug carries its costs

Table 3-2. Actual and Expected Costs per IND by Stage

| Development Stage | Actual Stage Cost per Successful IND[a] (millions) | Transition Probability | | Expected Stage Cost per IND[a] | Expected IND[a] Success Rate | Estimated Stage Cost per Approved Drug |
		Share of INDs entering Each Stage	Probability of INDs Success in Each Stage			
Computational Reference	*A*	*B*	*C*	*D = A*B*	*E*	*F = D /E*
Preclinical	$59.9	N/A	100%	$59.9	30.2%	$198.3
Phase I	$32.3	100%	83.7%	$32.3	30.2%	$106.9
Phase II	$37.7	83.7%	56.3%	$31.5	30.2%	$104.5
Phase III	$96.1	47.1%	64.1%	$45.3	30.2%	$149.9
Total Out-of-Pocket				**$169.0**		**$559.6**

[a]IND = Investigational New Drug
Source: Adapted from DiMasi et al. (2007).

as well as the costs of those that fail. Column (F), therefore, represents the expected stage cost per approved drug, taking into account the probability of success for a portfolio of INDs when it is not known in advance whether any particular IND in a portfolio will be successful.

The baseline estimates of expected R&D costs per approved drug reported in Table 3-2 are spread over many years. Preclinical costs may accrue over a period of 4 to 5 years on average, and it may take another 8 years to move through clinical trials and regulatory review. Investments made 12 years in advance of a return clearly have a different implicit cost than investments made immediately prior to a return. To take this "time value of money" into account, RTI followed the literature in reverse-discounting up to the date of marketing approval.

DiMasi and Grabowski estimated the cost of capital for a sample of biotechnology companies at 5-year intervals beginning in 1989 using the capital asset pricing model (CAPM). The CAPM framework is used to estimate the rate of return investors require and, thus, the minimum expected rate of return a company's financial managers require when deciding whether to approve a candidate project at the company. The authors' capitalized cost estimates were derived using a benchmark of 11.5% cost of capital from the late 1990s to mid-2000s. This discount rate took into consideration fluctuations in prevailing interest rates and equity-market premiums. They also confirmed the appropriateness of this rate through conversations with financial managers active in the industry.

R&D costs were spread evenly over each development stage's average length (estimated in months) and then compounded to the point of marketing approval to estimate total capitalized costs. Monthly costs in Table 3-3 were calculated by dividing the expected stage costs per IND from Table 3-2 by the mean stage length. The annual costs from Table 3-2 were broken down further into monthly costs to better apportion stage costs over time.

DiMasi and Grabowski capitalized the expected monthly costs by first converting the discrete annual discount rate to an equivalent continuous rate, changing the period to monthly, and then compounding continuously up to FDA approval. The capitalized cost per stage is calculated and presented in Table 3-3. Dividing this capitalized expected cost by the mean success rate of 30.2% results in the expected capitalized cost per approved drug.

3.2 ESTIMATING CURRENT EXPENDITURES ON THE TECHNOLOGY INFRASTRUCTURE

While business analysts and economists have investigated the costs of developing traditional pharmaceutical products, this study is the first to estimate annual industry spending on the technology infrastructure and among the first to narrow the research scope to encompass only the biopharmaceutical industry.

Table 3-3. Capitalized R&D Costs per IND by Stage

Development Stage	Expected Stage Cost per IND[a] (millions)	Stage Length (months)	Monthly Cost (millions)	Start of Stage to Approval (months)	End of Stage to Approval (months)	Expected Capitalized Stage Cost per IND (millions)	Mean IND[a] Success Rate	Expected Capitalized Stage Cost per Approved Drug (millions)
Preclinical	$59.9	52.0	$1.15	149.7	97.7	$185.6	30.2%	$614.7
Phase I	$32.3	19.5	$1.66	97.7	78.2	$71.8	30.2%	$237.7
Phase II	$31.5	29.3	$1.08	78.2	48.9	$56.3	30.2%	$186.5
Phase III	$45.3	32.9	$1.38	48.9	16.0[b]	$61.0	30.2%	$202.0
Total Present Value of Costs	$169.0					$347.7		$1,240.9

[a]IND = Investigational New Drug
[b]Average number of months for FDA to complete its regulatory review for biopharmaceuticals.
Source: Adapted from DiMasi and Grabowski (2007).

Current private-sector expenditures on technology infrastructure were estimated by interviewing and surveying biopharmaceutical drug developers and the technology suppliers, service vendors, consultants, and academics who support them. Reviewing corporate financial statements and industry R&D expenditure surveys offers little in the way of estimates of technology infrastructure spending. Technology infrastructure expenditures span R&D, capital investments, and many other line items in corporate financial reports. Thus, this study relied on the input collected from in-depth interviews and surveys of industry researchers, experts, and others to estimate spending. Chapter 4 provides more detail on this study's data collection activities.

Research directors and managers at biopharmaceutical companies identified their title, business unit, and technical background and provided estimates of their firms' (or business units') total annual spending for RTI-defined categories. They estimated the relative proportions of labor, capital, and materials expenses comprising the estimate and the distribution of those costs among the TFAs RTI identified. They also discussed the timing of all of those costs within the product development period.

Aggregated responses by stage were scaled to annual national expenditures using activity measures relevant for each drug development stage (see Table 3-4). Study participants provided several measures for comparing to and aggregating their data with other participants. Measures used to aggregate responses included

- the number of scientists and engineers comprising the business unit's research staff,
- the number of FDA-approved biopharmaceutical products, and
- the percentage of sales corresponding to the unit for which they are responding.

3.2.1 Technology Infrastructure Supporting Drug Discovery, Development, and Clinical Trials

Annual technology infrastructure spending supporting drug discovery, development, and clinical trials was estimated using data supplied by participants engaged in these areas.

Many businesses have specialty groups that strategize infrastructural issues, and these groups often span drug development stages. To capture spending appropriately, RTI asked groups to provide total

Table 3-4. Estimating Current Expenditures on the Technology Infrastructure

Activity Measures	Factor Used for Scaling	Available Breakdowns
Drug discovery, development, and clinical trial activities: Average annual expenditures per scientist	Number of scientists involved in discovery, preclinical, and clinical trial activities (~34,000)[a]	By type of expenditure (capital, materials, labor) By technology areas (bioimaging, bioinformatics, gene/protein expression platforms, molecular biomarkers, other)
Manufacturing activities: Average annual expenditures per drug	Number of approved biopharmaceutical drugs (~264)[b]	By type of expenditure (capital, materials, labor) By infratechnologies related to activities (preproduction, upstream and downstream processing, process monitoring, transaction costs)
Postmarket activities: Average annual expenditures per drug	Number of approved biopharmaceutical drugs (~264)[b]	By type of expenditure (capital, materials, labor) By infratechnologies related to activities (adverse event monitoring, database management, drug safety reporting, labeling, other)

[a]The number of scientists involved was calculated by taking the 2001 estimate of the number of scientists involved in biotechnology activities for human-health applications and applying a 1.9% annual growth rate to obtain a 2006 estimate. Sources: Department of Commerce (2003), Battelle (2006), and PhRMA (2006b).

[b]Source: BIO (2006) and PhRMA (2006a), based on list of FDA approvals as of June 2006.

spending rather than attempt to parse spending out too finely. To properly weight the responses, respondents reported the number of scientists[6] on their unit's research staff.

It was assumed that the average annual spending per scientist captured by the survey was representative of national spending. The average infrastructure spending per scientist was multiplied by the total estimated population of scientists to derive national expenditures.

The biopharmaceutical industry's scientific research staff totaled about 34,000 scientists in 2006. RTI derived this estimate by adjusting the Department of Commerce's 2001 estimate of the number of scientists active in biotechnology activities for human-health applications by the industry's average annual employment growth rate. PhRMA reported average annual employment growth of 1.9% between 2001 and 2004, and this rate was used to grow the 2001 estimate to 2006 (Battelle, 2006; PhRMA, 2006b).

---

[6]Specifically, respondents indicated the number of scientists, engineers, clinical laboratory technicians, and computer scientists involved in biotechnology activities—a definition that matches the one used by our secondary data source.

3.2.2 Technology Infrastructure Supporting Commercial Manufacturing Activities

Interviews with and completed surveys submitted from manufacturing operations managers were used to estimate technology infrastructure spending supporting commercial-scale manufacturing activities.

Rather than aggregate by the number of scientists, RTI used the number of FDA-approved product lines currently manufactured by the respondent's facility and the total national number of product lines currently in production.

Values from respondents were analyzed to obtain the "mean manufacturing technology infrastructure expenditures per approved biopharmaceutical," which was then multiplied by the number of approved biopharmaceuticals available from Biotechnology Industry Organization (BIO) publications.

Recognizing that each individual drug requires specific technology infrastructure investments that must be made to support manufacturing, the number of approved biopharmaceuticals was nevertheless deemed the best available factor by which to scale estimates.[7]

To this end, survey data were reviewed to assess if any particular respondents' drug types were unlikely to be representative of the biopharmaceutical drug population as a whole. Individual production processes are unique even for similar types of therapies (e.g., two types of therapeutic proteins may require vastly different upstream processing and the downstream purification process will be similarly tailored). No significant inconsistencies in the data reported warranted adjusting the extrapolation approach.

3.2.3 Technology Infrastructure Supporting Postmarket Surveillance Activities

The approach to estimating private-sector expenditures on the infratechnologies that support efficiency in these activities was similar to that for commercial manufacturing. The mean technology infrastructure expenditure per approved biopharmaceutical was multiplied by the total approved number of biopharmaceuticals to obtain total technology infrastructure expenditures supporting postmarket activities.

---

[7]The number of scientists working in commercial-scale manufacturing also was collected in RTI's survey, but the Department of Commerce (2003) study does not include these individuals in the total number of scientists. Instead, they are classified as "production workers" and reported in combination with managers and other administrative positions.

3.3 ESTIMATING POTENTIAL REDUCTIONS IN BIOPHARMACEUTICAL DEVELOPMENT AND PRODUCTION COSTS

The second objective of this study was to measure the "excessive" costs incurred by the industry because of an inadequate technology infrastructure.

Excessive costs are those that could be avoided if the industry had access to an improved, and therefore "more adequate," infrastructure. Industry experts helped define hypothetical improvements to the technology infrastructure supporting each TFA that are feasible within the next 10 years. These improvements were included in the survey, and respondents provided their assessments based on the scenarios RTI presented. (Section 3.3.7 concludes this chapter with an accounting of these improvements.) This section outlines the economic model that estimated potential reductions in R&D and production costs by recalculating the average cost to develop and manufacture a drug.

The metric most commonly used for assessing industry's spending is the "R&D cost per approved drug."[8] The economic model recalculates average R&D cost per approved drug using the baseline costs from the DiMasi and Grabowski study. Baseline estimates included biopharmaceutical drug development costs, times, and the probability a candidate drug moved from one R&D stage to the next.

Respondents' estimates of the impact the improvements would have on costs, time, and quality were inputted into the model. Respondents also offered their estimates of changes in the probability that a drug candidate moves from one stage to the next and related measures, relative to industry averages.[9] The ***percentage change*** in costs possible given the

---

[8]RTI followed this approach, rather than looking at annual R&D costs at the firm level, for multiple reasons. First, and most importantly, the models for R&D cost per approved drug build in the cost of failure—a very important component of the industry's total R&D spending and a major focus of efforts to streamline drug development. Second, private firms may respond to reductions in drug failure rates, not by lowering their R&D spending but by producing more drugs. If this is the case, total industry R&D spending would remain unchanged, but the R&D cost per approved drug would decrease.

[9]There are likely substantial benefits to patients' quality of life from receiving better treatments earlier. However, it is beyond the scope of this study to assess those quality-of-life metrics quantitatively. It is also beyond the scope of this study to quantify the additional profits that firms may earn by reaching the market earlier and thus being able to take better advantage of their limited patent lives. Finally, improvements to the technology infrastructure supporting the biopharmaceutical industry may affect bioagriculture, bioenergy, or biodefense. Quantifying those impacts is also outside the scope of this study.

hypothetical improvements for each stage of drug development and production was used to estimate excessive costs.

3.3.1 Baseline Preclinical Cost per Approved Drug

Preclinical costs are the R&D expenditures incurred before submitting an IND and entering a candidate therapy into clinical trials. The total costs of preclinical development, $TC_{preclinical}$, at a firm for a specific year or set of years is divided by the number of biopharmaceuticals that were eventually approved out of that cohort:

$$\text{Discovery/Preclin.} \atop \text{Cost per Approved Drug} = \frac{1}{D} \left[TC_{preclinical} \right] \qquad (3.1)$$

where D is the number of approved drugs that eventually result from these discovery and preclinical research efforts.

During the survey, respondents were asked how baseline costs, stage length, and the probability of success could change with an improved technology infrastructure. The data in Table 3-5 both summarize and present an alternative view of the DiMasi and Grabowski results discussed in Section 3.1.2.

3.3.2 Baseline Clinical Trials Cost per Approved Drug

Clinical trial costs are easier to quantify than the costs of earlier R&D stages, because drug sponsors often track them individually for each compound filed as an IND. However, not all INDs become approved drugs. The clinical trial cost per approved drug is given by

$$\begin{aligned}
\text{Clinical Trial Cost} \atop \text{per Approved Drug} &= \frac{1}{D} \left[\text{Total Clinical Trial Costs Incurred} \right] \\
&= \frac{1}{D} \left[C_I \cdot IND + C_{II} \cdot IND(1 - p_1) + \right. \\
&\qquad \left. C_{III} IND(1 - p_1)(1 - p_2) \right] \qquad (3.2)
\end{aligned}$$

where C_I, C_{II}, and C_{III} represent the average cost of taking a candidate through Phase I, II, and III trials, respectively, regardless of whether those compounds are successful.

These costs include expenses related to conducting clinical procedures, as well as the costs associated with manufacturing small quantities of drugs for trial use, managing the trials, and analyzing the resulting data.

The remaining terms in Eq. (3.2) track the expected number of individual compounds incurring those phase costs. *IND* represents the total number of drugs entering Phase I trials (the number of INDs filed). All of

Table 3-5. Baseline Parameters for the Biopharmaceutical R&D Cost Model

Stage	Stage Costs per IND (millions)	Distribution of Failed INDs, by Clinical Trial Phase	Stage Length (Time to Next Stage in Months)
Discovery/preclinical	$59.89[a]		52.0[a]
Phase I	$32.28[a]	23.4%	19.5[a]
Phase II	$37.69[a]	52.4%	29.3[a]
Phase III	$96.09 [a]	24.2%	32.9[a]

[a]Adapted from DiMasi et al. (2007).

these compounds would incur costs for Phase I. The probabilities that a drug entering Phase I trials fails in Phase I or in Phase II are given by p_1 and p_2, respectively. $IND(1 - p_1)$ represents the expected number of compounds incurring Phase II costs, and $IND(1 - p_1)(1 - p_2)$ represents the expected number incurring Phase III costs.

Noting that (D/IND) reflects the probability that an IND entering Phase I clinical trials is successfully approved, $p(approval)$, the above expression can be rewritten as follows:

$$\text{Clinical Trial Cost per Approved Drug} = \frac{CI + IND(1 - p_1) + IND(1 - p)(1 - p_2)}{p(approval)}. \quad (3.3)$$

This value represents the average actual cost of clinical trials for one approved drug, taking into account the cost of failures that firms experience along the way.

As with the discovery and preclinical development costs, this analysis relied on the DiMasi and Grabowski paper to determine a baseline value for the cost of conducting clinical trials. In this way, the baseline estimates are a function of within-stage costs (C_I, C_{II}, and C_{III}) and success and failure probabilities (p_1, p_2, and $p(approval)$).

The timing of investments is important to calculating the total capitalized cost per approved drug. The periods over which costs were reverse discounted using the average biopharmaceutical cost of capital were adjusted using participants' expected stage length reductions. These adjustments affected the capitalized cost by shortening firms' investment horizons.

3.3.3   Baseline Commercial Manufacturing Costs

The baseline model of manufacturing costs is less complex than the baseline model for preclinical/clinical R&D costs because it does not account for the probability of failure. Manufacturing costs include any expenses incurred to produce commercial-scale volumes of an approved biopharmaceutical drug or therapy:[10]

Annual Manufacturing Costs =
>   Annual Preproduction Costs +
>   Annual Upstream Processing Costs +
>   Annual Downstream Processing Costs +
>   Annual Process Monitoring Costs   (3.4)

Preproduction costs refer to all costs incurred to scale up from clinical trial volumes to the volumes needed in the marketplace. Upstream and downstream processing costs refer to the costs of producing the raw product material (usually through fermentation or cell culturing) and the costs of extracting, filtering, and purifying that raw product, respectively. Process monitoring costs include the costs of tracking and modifying environmental parameters, as well as testing batch characteristics. Survey respondents determined the percentage of total annual manufacturing costs accounted for by each of these activities.

Total annual manufacturing costs for biopharmaceuticals are calculable from the Department of Commerce (2003) report detailing estimates of biopharmaceutical sales. Survey respondents were asked to estimate manufacturing costs as a percentage of annual sales. The average of their responses was applied to total sales to estimate potential reductions. RTI verified the appropriateness of the average using company annual reports and additional secondary resources.

3.3.4   Baseline Postmarket Surveillance Costs

Postmarket surveillance costs include commercializing a drug and verifying that it behaves in the marketplace as designed. Surveillance includes testing product quality and stability, as well as continuing to monitor the product for adverse events in the wider population. The annual cost of postmarket activities is as follows:

---

[10]Recall that costs to produce small batches for use in preclinical studies or clinical trials, though very important, are counted by the industry as R&D expenditures.

Annual Postmarket Transactions Costs =
Annual Quality Assurance Costs +
Annual Adverse Event (AE) Monitoring Costs +
Annual AE and Safety Data Management Costs +
Annual Regulatory Drug Safety Reporting Costs +
Annual Product Label Updating Costs   (3.5)

As with manufacturing costs, the baseline postmarket costs were determined through a combination of surveys, interviews, and secondary sources. These costs were expressed as a percentage of total annual sales of biopharmaceuticals.[11] The shares reported in the online survey informed the baseline value for each type of transaction cost. In practice, some of these transaction costs are incurred by manufacturers, who must verify quality, while other transaction costs are incurred by individuals performing postmarket safety surveillance.

3.3.5   Hypothesized Economic Impact of an Improved Technology Infrastructure

RTI considered the cost savings that would be associated with an improved (more adequate) infrastructure. The definition of "improved infrastructure" was described for survey respondents and interview participants in terms of a suite of infratechnologies that could feasibly be made available within the next 10 years.

In some instances, new infratechnologies or processes can be described for which it is reasonably clear what the improved attributes will be. However, for others the advancement might be a relative improvement in accuracy, reliability, and stability of existing infratechnologies. In the latter case, respondents were asked to consider advancements of 50% over the current state of the art in those areas.

When considering the impact of this improved technology infrastructure on biopharmaceutical development and production costs, the model anticipated three possible effects:

1. *The advancements may change the within-stage costs (e.g., $TC_{preclinical}$, C_l, upstream processing costs).*

   For example, standardization may reduce the need for redundant tests to match and verify results. Better approaches to scale up and enhanced process control may reduce manufacturing costs.

---

[11]A recent economic study estimated the postmarket safety surveillance costs at approximately 0.3% of sales, though this estimate is for all pharmaceuticals, not just biopharmaceuticals (Ridley et al., 2006).

Other infrastructure improvements may lower transaction costs experienced postproduction.

2. *The advancements may change the transition probabilities or probability of overall approval.*

An improvement in infrastructure used in the preclinical stage may allow researchers to better characterize the properties of their investigational compounds, which may affect the success probabilities in later stages (e.g., by increasing the probability of passing Phase I and, therefore, the overall approval rate).

3. *The advancements may change the duration of any of the stages, possibly bringing products to the point of marketing approval more quickly.*

This does not affect the expenses incurred by biopharmaceutical development firms, but it does affect the timing of those expenses. A shorter time to approval will be reflected by a lower present value cost.

A single improvement to the infrastructure may affect any or all of these dimensions of the total cost of biopharmaceutical development and production. Standardization or calibration techniques that increase consistency of results may lower stage costs or shorten the time spent in one stage by eliminating redundant tests, for example. However, improved consistency alone might not directly improve the probability of success. Alternatively, improvements in the technical capabilities of a particular infratechnology may increase the quantity of known information about a compound and thereby improve the probability of success, yet this benefit might be offset by an *increase* in the mean stage costs if it means new tests or equipment. With any change there may be one-time costs of adoption; therefore, the total cost of development could increase during the transition.[12]

3.3.6 Model Implementation Using Impact Metrics

In a recent study, DiMasi (2002) explored the effect that the above impacts could have on pharmaceutical drug R&D costs. His results suggest that hypothetical reductions in all stage lengths of 25% would lower the future-value total cost per approved drug by 16%. Additionally, a hypothetical increase in the overall success rate for compounds entering clinical trials from 21.5% to 33.3% would lower the present value cost per approved drug by up to 30.2% (DiMasi, 2002).

---

[12]We will not estimate these transition costs, because we are not trying to describe exactly what it will take or cost to achieve hypothetical infratechnology improvements.

A similar approach was used in this analysis to determine potential changes in the total cost of biopharmaceutical development—starting with a baseline estimate and re-estimating that cost in response to changes in the underlying parameters. However, the model's estimates rely on expert opinions about potential infrastructural improvements to determine the appropriate changes to those parameters. The survey asked respondents to estimate impact values that were then used to reestimate the cost of biopharmaceutical drug development (see Table 3-6).

Potential Impact on Discovery/Preclinical Development and Clinical Trial Costs

The impact metrics are relatively straightforward when the effect of an infratechnology improvement is on within-stage costs or durations. For example, if an improved infrastructure allows the same task to be performed at a lower cost, this is captured by estimating the percentage change in the cost of processing a compound (e.g., $C_{preclinical}$, C_I) attributable to the supporting infratechnologies. Likewise, if validation procedures allow for the use of biomarkers as surrogate endpoints that shorten clinical trials, this is captured by estimating the percentage change in the length of the stage. These metrics are shown for the R&D stages in Table 3-7.

These percentage reductions were applied to the baseline stage costs and stage lengths. To verify that estimates from survey respondents were reasonable, we used in-depth interviews to collect information on some intermediate metrics (e.g., number of redundant tests eliminated, cost per test, and length of time required for each test).

The assessment of the impact on success probabilities was more complicated, because the improvements in one stage can have effects that filter into later stages. For example, the likelihood that a drug fails in Phase I determines the quality of drugs entering Phase II and is itself determined by the quality of compounds coming out of preclinical development.

First, the approach assumed that the probability that an IND succeeds in clinical trials is essentially determined by the end of the preclinical stage. By this point, the drug sponsor has established a target, optimized the lead, and identified the patient populations. What remains is a screening process. If a drug has short-term safety problems, it assumed those are

Table 3-6. Estimated Potential Impact of an Improved Technology Infrastructure

Final Estimates	Based on
Potential Reduction in Annual R&D Costs Potential reduction in R&D cost per drug = a function of Potential reductions related to bioimaging, Potential reductions related to bioinformatics, Potential reductions related to gene/protein expression, and Potential reductions related to molecular biomarkers	A baseline figure of $1.3 billion for the R&D cost per approved drug and associated baseline stage costs and probabilities[a] Total annual R&D for the industry Impact metrics: percentage reductions in stage costs and lengths, revised success and failure rates
Potential Reduction in Annual Manufacturing Costs	Total annual sales for the industry Estimated manufacturing costs as a percentage of sales[b] Impact metrics: percentage reductions in stage costs
Potential Reduction in Annual Postmarket Costs	Total annual sales for the industry Estimated postmarket costs as a percentage of sales[b] Impact metrics: percentage reductions in stage costs

[a]Adapted from DiMasi et al. (2007) in *Managerial and Decision Economics*.

[b]Survey respondents provided information about their current manufacturing and postmarket surveillance costs as a percentage of their total sales. These percentages, along with industry estimates of annual sales, were used to estimate a baseline figure for manufacturing costs and for postmarket surveillance costs.

Table 3-7. Metrics for Assessing the Impact on R&D Costs[a]

Stage	Percentage Reduction in Stage Length (Time to Next Stage)	Percentage Reduction in Stage Cost
Discovery/preclinical	_____%	_____%
Phase I	_____%	_____%
Phase II	_____%	_____%
Phase III	_____%	_____%

[a]This table appeared in the online survey for respondents to complete. The survey allowed both preclinical and clinical trial scientists to fill out all rows of the table, so that they could account for some of the interactions between these R&D stages. For example, *in silico* modeling used by preclinical scientists to assess toxicity may allow for the reduction of certain clinical trial procedures.

detected in Phase I. If a drug is ineffective or has long-term safety problems, the model assumed those are detected in Phases II and III.

Yet, regardless of where failures occur, the inherent probability of approval conditional on entry in clinical trials, p(approval), was assumed to be determined during discovery and preclinical trials. Therefore, the

value of enhancements to the technology infrastructure supporting these two initial stages was measured by asking respondents to estimate a new p(approval). They also estimated a new p(recall | approval), the probability that an FDA-approved biopharmaceutical is recalled, withdrawn, or relabeled in response to adverse effects.

Holding p(approval) constant, the distribution of the remaining failures across the three trial stages was estimated. The industry average p(approval) is estimated at 35%, meaning that 35 of 100 candidates entering Phase I clinical trials ultimately receive FDA approval. The remaining 65 candidates are expected to fail during one of the three trial phases. The survey provided firms with the industry average distribution of failures within clinical trials and asked them how this distribution would change under the improved infrastructure scenario.

A matrix for these types of responses appears in Table 3-8. The estimates about the distribution of failures given in the middle of Table 3-8 were converted to the estimates about p_1 and p_2 needed for the theoretical model.

Using this information on improved probabilities of success and failure and the improved stage costs and lengths (calculated using the baseline figures and percentage reductions reported in the surveys), the model calculated an improved discovery/preclinical cost per approved drug and an improved clinical trial cost per approved drug. By comparing the difference between the baseline cost and the improved cost estimates, the model identified the potential impact of an improved (adequate) infrastructure:

$$\text{Potential Impact on Disc/Preclin. Costs} = \frac{\text{Improved Disc/Preclin. Cost per Approved Drug} - \text{Baseline Disc/Preclin. Cost per Approved Drug}}{\text{Baseline Disc/Preclin. Cost per Approved Drug}}.$$

Applying this percentage to the current annual R&D expenditures on discovery/preclinical biopharmaceutical development yields a dollar value for the potential industry-wide impact. The same calculation and logic apply to the potential impact on clinical trial costs.

Table 3-8. Metrics for Assessing the Impact on Success and Failure Rates

Change in IND Overall Success Rates[a]	Typical Percentages[a]	Estimated Percentages in World with Improved Infrastructure[b]
p(approval)		
Probability that an IND ultimately receives FDA approval	35%	_____%
Change in Clinical Trial Failure Rates		
Of the remaining INDs that **fail** in clinical trials, what percentage **fail** in each stage?		
Percentage of failures occurring in Phase I	15%	_____%
Percentage of failures occurring in Phase II	46%	_____%
Percentage of failures occurring in Phase III	39%	_____%
Total	**100%**	**100%**
Change in Postmarket Failure Rates		
p(recall | approval)		
Probability that an approved drug is recalled	0.4%	_____%

[a]Based on estimates from Pavlou and Reichert (2004) and PAREXEL's Pharmaceutical R&D Sourcebook (Mathieu, 2005).

[b]The definition of "improved infrastructure" was detailed for respondents in terms of a suite of infratechnologies that could feasibly be made available in the next 5 to 10 years. For infratechnologies that already exist to some degree, respondents considered advancements of 50% over the current state of the art.

Comparing Impact Assessments across the R&D-Related Technology Focus Areas

In some cases, the proposed potential improvements may be complementary, but in other cases, they may be substitutes that researchers will choose between when it comes to adoption. If the TFAs involve infratechnologies that are complementary, then the cost reduction from implementing different technologies may be more than the sum of the individual impacts.[13]

For example, consider the bioimaging and bioinformatics TFAs. Suppose standards and improved change measurements that allow for enhanced bioimaging techniques reduce the cost of conducting Phase III clinical trials by an estimated 10%. Suppose also that advancements in image formatting and search capabilities alone would improve Phase III costs by 5%. Should both advancements be available simultaneously the net

---

[13]Likewise, if the TFAs involve infratechnologies that are substitutes, the net impact may be less than the sum of the individual cost reductions.

impact may be slightly lower than 15% because some bioimaging and bioinformatics infratechnologies may be addressing the same inefficiency.

Thus, the greatest potential improvements in cost and time per stage, irrespective of TFA, were used to determine the extent to which the hypothetical improved infrastructure would affect total R&D costs. Shortening the investment horizon while simultaneously reducing cash outlays significantly reduces capitalized costs. Model results thus illustrate how R&D costs per IND and the expected capitalized costs per approved drug were affected by anticipated lower costs from an improved infrastructure as well as shorter development times.

It is important to note that these calculations hold all other costs and production factors constant. Outside of this model, real world shifts and reassignment of resources is possible if companies identify opportunities. Thus, any total potential cost reduction is based on historical, observed costs that do not reflect future costs.

Potential Impact on Manufacturing Costs

A similar approach was used to estimate the potential impacts on manufacturing costs, although the model did not have to account for changing probabilities or stage lengths and did not have to combine TFAs. Table 3-9 reports the impact estimates collected from the survey and interviews.

These reductions were applied to the baseline estimates for manufacturing costs, and these elements were summed to estimate a total potential reduction in manufacturing costs due to an improved (adequate) infrastructure.

Table 3-9. Metrics for Assessing the Impact on Manufacturing Costs

Phase/Activity Cost	Percentage Reduction in Cost by Phase/Activity
Preproduction (scale-up from clinical trial volumes) costs	_____%
Upstream processing costs	_____%
Downstream processing costs	_____%
Process monitoring and quality assurance testing costs	_____%

Potential Impact on Postmarket Costs

Estimated reductions in postmarket costs were applied to the baseline estimates and summed to estimate a total potential reduction in postmarket costs due to an improved infrastructure (see Table 3-10).

3.3.7 Hypothetical Improvements by TFA Posed to Survey Respondents and Interviewees

This section discusses the hypothetical improvements posed to respondents for each TFA. Respondents chose one TFA area to explore in the survey and were requested to explore the one with which they were most familiar.

Bioimaging

This TFA focused on the inadequacies in bioimaging capabilities, including the need for novel reagents to produce high-quality images and better software algorithms to resolve these images.

For bioimaging, RTI asked respondents about the potential R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of bioimaging in biopharmaceutical discovery, development, and testing. The advancements listed come from the NIH Technology Road Map, the FDA Critical Path Opportunities List, and other industry sources.

Participants were asked to consider the following advancements:

A. Qualified imaging biomarkers in a disease or therapeutic category of interest.

B. Availability of labeling and contrast agents that are proven to be stable *in vivo*, nondestructive to the molecule being tested, and nontoxic in humans.

C. Reconstruction algorithms that improve image quality (e.g., reduce fuzziness, define clear borders).

Table 3-10. Metrics for Assessing the Impact on Postmarket Costs

Phase/Activity Cost	Percentage Reduction in Cost by Phase/Activity
Quality assurance testing	_____%
Adverse event (AE) monitoring	_____%
AE and safety data management	_____%
Regulatory drug safety reporting	_____%
Product label updating	_____%

D. Automated or computer-assisted interpretation of images, including edge detection and size measurements.

E. National image test bed and benchmarking methodology to assess the quality of the image analysis algorithms and software.

F. Standard protocols for patient/specimen positioning, instrument calibration, and settings to reduce variability and allow images to be compared across tests (in discovery) or trials (in the clinical testing phases).

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include the following:

- R&D labor and materials cost savings associated with shorter clinical trials or smaller trial populations.

- Elimination of alternative tests required to identify and validate a target or response.

- Reduction in the number of animal subjects needed (since noninvasive imaging would allow the same animal to be followed for longer periods of time).

- Reduction in the number and skill level of technicians required to verify the analysis of an image.

- Elimination of additional (redundant) tests needed to verify results, because of poor image quality.

- Reduction in R&D labor costs for time spent assessing performance capabilities of image analysis software.

- Reduction in R&D labor costs associated with calibrating equipment because of the availability of standardized procedures and reference phantoms.

Biomarkers

Respondents considered the R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of biomarkers in biopharmaceutical discovery, development, and testing.

Survey respondents and interviewees were presented with the following hypothetical improvements:

A. Standardized and accepted validation process for biomarker discovery.

B. Safety biomarkers that predict human toxicity in preclinical stages of drug development.

C. Validated efficacy biomarkers for diseases or therapeutic categories in an area of interest that could be used as surrogate endpoints in clinical trial activities.

D. Enhanced ability to predict biological response to treatment (e.g., drug efficacy measurement).

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include the following:

- Reduction in the number of tests conducted in preclinical studies due to the reduced uncertainty of drug efficacy.

- Avoidance of future R&D costs due to early detection of toxicity that eliminates drug candidates before entering clinical trials.

- Shorter clinical trial periods due to faster detection of biologic response to drugs.

- Reduced number of patients in clinical studies due to more accurate selection criteria, resulting in improved stratified patient populations.

- Reduction in labor and material costs during clinical trials due to the ability to set dosage levels on a per-patient basis.

Bioinformatics

The study addressed a host of issues related to using databases and information technology in laboratories, clinical settings, and production facilities. These issues ranged from the need for reliable databases in drug discovery to data management techniques that link preclinical and clinical trial results and new ways to store high-content data such as images.

In addition, the study addressed the need for refined *in silico* models based on such data. Researchers need improved algorithms capable of simulating cell and protein behavior. These algorithms would ultimately improve the predictions of drug efficacy at earlier stages of development such as target identification and lead optimization studies. Models of predictive ADME and toxicity are also important needs for improved candidate selection.

Respondents reflected on the potential R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of bioinformatics in biopharmaceutical discovery, development, and testing.

Participants were asked to consider:

A. Shared data standards for formatting and content. In addition to formatting standards for storing different types of data (such as image data), this advancement would include standards for the metadata that record information on the conditions of an experiment, similar to the content standards used as Minimum Information About a Microarray Experiment (MIAME).

B. Greater availability of currently gathered data through publicly accessible, curated databases. These databases would include

(but not be limited to) data concerning the natural history of rare diseases, adverse events, toxicology properties of drug candidates, and ADME properties of drug candidates.

C. Improved data mining applications and algorithms. Potential improvements include the incorporation of natural language processing into a text search.

D. Improved accuracy of *in silico* model predictions, including (but not limited to) models of protein structure and binding, cellular localization, PK/PD properties, clinical trial simulation, and disease modeling.

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include the following:

•   Reduction in time and labor costs spent on finding and interpreting data from previous experiments.

•   Reduction in labor and material costs related to redundant experiments.

•   Reduction in the number of animal subjects needed in preclinical trials (since *in silico* models can be used to predict how drug candidates will interact with the subject).

•   R&D labor and material cost savings associated with fewer product redesign-synthesize-test cycles (since more accurate *in silico* models could be used to predict the effect of a proposed structural modification of a therapeutic product).

•   Increased clinical trial success rates due to early detection of toxicity using improved *in silico* models of information obtained from expanded adverse event databases; these models may eliminate drug candidates before entering clinical trials.

•   Reduction in the number of trials and patients through the simulation of clinical trials using *in silico* modeling.

Gene Expression

The TFA on standards and metrology in gene and protein expression examined costs incurred during discovery stages and preclinical and clinical investigation stages that result from the absence of a standardized set of reference materials.

In addition, this TFA considered the costs incurred during the manufacturing stage as a result of inadequate methods and standards to optimize expression systems at the upstream processing stages of manufacturing for therapeutic proteins and vaccines.

The gene expression component queried respondents about potential R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of gene and/or protein

expression analysis in biopharmaceutical discovery, development, and testing.

Survey respondents and interviewees were presented with the following hypothetical improvements:

For gene expression:

- Publicly available synthetic mRNA reference materials for microarray performance assurance.

- Standard technical protocols for microarray experiments.

- Standard protocols for RNA and DNA extraction.

- Data and analysis to benchmark microarray performance.

- Microarray scanning equipment calibration tools.

For protein expression:

- Techniques for measuring the presence of low-abundance proteins.

- Improved sensitivity and lower coefficients of variation (≤10%) in mass spectrometry analysis to allow for its use in later stages of drug development.

- Improved availability of antibodies with high affinity, specificity, and selectivity for use with protein microarrays.

- Standards for protein microarray experiments.

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include the following:

- Reduction in labor, microarray, and consumables costs for redundant experiments and avoided downstream data capture and analysis costs for those experiments.

- Avoided downstream R&D costs for investigating genes and proteins mistakenly identified as potential targets.

- Greater confidence in and comparability among results from microarray experiments allowing for the elimination of redundant tests.

- Reduction in labor costs associated with calibrating equipment due to the availability of standardized procedures and reference phantoms.

- Reduction in or elimination of microarray scanning errors due to poor equipment alignment and calibration and associated downstream impacts of avoidable data capture errors.

- Elimination of additional (redundant) tests needed to verify results due to poor image quality.

Commercial-Scale Manufacturing

The NIH Technology Road Map, the FDA Critical Path Opportunities List, and industry sources identified a number of areas that could benefit from additional research. The survey asked individuals involved in manufacturing biopharmaceuticals to consider advancements in the following areas related to the technology infrastructure supporting different stages of manufacturing.

Preproduction:
- Predictive modeling and underlying data for determining batch yield given a specified level of inputs and production parameters (which would reduce the number of test batches required during process scale-up).

Upstream Processing:
- Robust expression systems that produce raw proteins in higher yields with fewer impurities (for example, the use of transgenic plants and animals).

- Technologies, such as disposable bioreactors and mixing systems, which can accommodate rapid changes in manufacturing processes and can reduce the risk of biological product contamination.

Downstream Processing:
- Purification technology that can improve flow rates and increase capacity over current methods. Examples of new technology might include membrane chromatography or improved, high-pressure affinity chromatography among other methods.

Process Monitoring (crossing over all phases of production):
- Implementation of PAT to better understand, monitor, and control production processes in real time. For example, this might involve the following:

  – Improved detection of contamination in biological products (e.g., viruses, bacteria, and other organisms) through the use of microarrays or proteomics infratechnologies.

  – Uniform standards for spectroscopic instruments; for example, standards for appropriate instrument qualification and calibration standards for techniques such as Raman and Terahertz spectroscopy.

- Improved methods for product characterization, including enhanced potency assays and appropriate statistical and sampling techniques. For example, this might involve the following:

  – More reliable and quantitative nonanimal-based tests of vaccine potency.

  – Potency measurements that provide reliable information about the quality of cells or tissues to be used in therapies.

Postproduction:

- Improved certainty about product characteristics, including potency, sterility, purity, and handling requirements.

Enhancements in these areas could lead to a number of impacts on manufacturing productivity, such as the following:

- Reduced time to market and higher product quality as a result of better scale-up procedures, shorter downstream processing purification periods, and faster and more reliable batch testing procedures.

- Improved production yields for a given amount of raw material.

- Reduced labor and material costs with fewer trial and error iterations and fewer contaminated batches.

- Labor saved in sterilization activities between batches.

- Lower transaction costs to customers as a result of greater reliability of data and fewer instances of inactive product released.

- Smaller amount of product required for testing.

Postmarket Surveillance

Respondents considered advancements in the following areas related to the technology infrastructure supporting postmarket surveillance activities:

- Improved statistical methods for signal detection in adverse event monitoring.

- Development of standards for adverse event reporting systems.

- A standardized process and protocol for ensuring data consistency across international adverse event databases.

- Improved interoperability between safety and clinical database management systems.

- Automated monitoring software tools to provide notification of required product label changes based on changes in adverse event frequencies.

- Improved integration linking data entry, clinical databases, randomization databases, and regulatory systems.

Enhancements in these areas could lead to a number of impacts on the productivity of postmarket surveillance activities, such as the following:

- Reduced labor costs of physicians required to conduct adverse event monitoring and data analysis.

- Reduced labor costs in reconciling differences across international adverse event databases.

- Reduced transaction costs in updating product labels based on adverse event frequency changes.

- Reduced labor costs to generate electronic submissions to regulator agencies.

- Reduced risk due to the mitigation of potential errors in analyzing and interpreting adverse event data.

4 Primary Data Collection

This study's data collection activities focused on individual researchers and organizations that apply biotechnology to developing human therapeutic drug products. Primary data collection methods included on-site and telephone interviews with technical experts from the biopharmaceutical industry and an Internet survey of biopharmaceutical companies. Secondary data sources included the professional literature and economic surveys conducted by the federal government and research organizations. This chapter describes the data collection process and the key stakeholders surveyed for this study.

Table 4-1 summarizes the impact metrics respondents were asked to quantify during interviews and in the Internet survey. The metrics include percentage cost reductions for preclinical, clinical, and manufacturing activities; changes in the percentage distribution of where drugs fail; and changes in the length of time for preclinical and clinical phases. Interaction effects between these impact metrics were captured using a spreadsheet model employing the methodology described in Chapter 3.

Table 4-2 summarizes the number of responses received from our interviews and survey by TFA. We conducted 44 in-depth interviews with technical experts, senior scientists, and research directors from the biopharmaceutical industry. These interviews provided a wealth of information relating to current technology infrastructure expenditures and the potential impact of improvements on the technology infrastructure. Technical experts who participated in the study represented firms whose

Table 4-1. Impact Metrics Respondents Quantified

Impact Category	Impact Metric	Comment
Cost Reductions in Discovery/Preclinical Stage and Clinical Trials	Percentage change in R&D costs per successful drug (see Table 3-7)	Standardization may reduce the need for redundant tests to match and verify results
Cost Reductions in Manufacturing	Percentage change in manufacturing costs (see Table 3-9)	Better approaches to scale-up and enhanced process control may reduce manufacturing costs
Cost Reductions in Postmarket Surveillance	Percentage change in postmarket surveillance costs (see Table 3-10)	Infrastructure improvements may lower transaction costs experienced in postproduction
Change in the Probability of Success of INDs	Change in the success rate (%) of INDs (see Table 3-8)	Screening out more drugs in the preclinical stage reduces the need for clinical trials
Changes in When Drugs Fail in the Clinical Stages	Change in the percentage distribution of where drugs fail (see Table 3-8)	Identifying failed drugs sooner rather than later in the clinical stage reduces the use of expensive Phase II and Phase III trials
Shortening of Drug Development Time	Percentage time reduction of preclinical and clinical phases (see Table 3-7)	Decreasing drug development time reduces present value costs (holding actual costs constant)

Table 4-2. Number of Responses by Technology Focus Area

Technology Focus Area	Interview	Internet Survey	Total
R&D Sector			
Bioimaging	7	7	14
Biomarkers	9	9	18
Bioinformatics	10	6	16
Gene expression	7	19	26
Commercial Sector			
Commercial manufacturing	8	12	20
Postmarket surveillance	3	5	8
Total	44	58	102

combined annual R&D spending accounted for as much as 42% of total industry R&D spending[1] and 49% of annual R&D sales.[2]

4.1 TECHNICAL EXPERT INTERVIEWS

RTI conducted detailed on-site and telephone interviews with 44 technical experts knowledgeable in at least one of the six TFAs. We used three information sources to identify prospective participants for the telephone interviews: conference proceedings, professional and technical literature, and referrals. Several professional societies sponsor workshops and conferences on technology state of the art and application to one or more components of the biopharmaceutical product life cycle.

4.1.1 Technical Expert Interview Methodology

RTI obtained the attendance lists and conference proceedings from several recent biopharmaceutical technology conferences and workshops. We also searched the professional and technical literature for authors of recent articles on each TFA. Interviewees often referred us to other individuals knowledgeable about activities in their realm of expertise or in complementary research areas. In some cases, RTI interviewed several representatives from the same organization to acquire as broad a perspective as possible on the organization's infratechnology investments.

The technical experts RTI contacted represented academia, government, and public and private enterprises. The experts we formally interviewed are among the most active in the field, as indicated by their conference attendance, number of articles published in the professional and technical literature, number of corporate white papers authored, and participation in formal societies. Interviewees represented professors, directors of R&D programs, and product line managers, as well as both executive and technical staff (see Table 4-3).

---

[1]RTI calculated market share for the technologies applied in the R&D phases by dividing company-reported R&D expenditures in 2004 by the total R&D expenditures for all publicly traded biotechnology firms. Market share for the manufacturing and postmarket segment was calculated by dividing net sales of biopharmaceuticals of the individual firm by the total sales for all commercialized biopharmaceuticals.

[2]Responses to the Internet survey were anonymous and hence could not be linked to R&D expenditures or sales. Thus, Internet respondents are not represented in the market share figures. As a result, the market share of firms participating in this study is greater than reported here.

Table 4-3. Examples of Interviewed Technical Experts′ Job Titles

Background	
Research Director	Vice President, R&D
Senior Scientist	Chief Medical Officer
Executive	Facility Manager
Program Director	Quality Manager
Consultant	Global Unit Lead

Each interview was conducted over the telephone and lasted between 30 and 90 minutes. Typical interviews consisted of a brief overview of the study and its objectives, followed by a series of questions aimed at assessing the impacts that potential improvements to the technical infrastructure would have on different stages of a biopharmaceutical product's life cycle. If further clarification was needed after the initial interview about particular comments or issues raised, the interviewee was recontacted.

RTI worked diligently to corroborate anecdotes about the challenges that inadequate infrastructures pose, areas of particular need for investment, and views on the current state of the industry. Some experts were unavailable for lengthy discussions but did discuss general industry trends and their perspectives on infrastructural needs. These short conversations confirmed comments and insights gathered from other interviews but are not included in the responses in Table 4-2.

4.1.2 Topics Covered in Technical Interviews

RTI asked experts to reflect on the current spending, the current distribution of costs across different TFAs or processes, and any potential role NIST or industry associations might play in infratechnology development and adoption.

The questionnaire consisted of three sections, each containing a series of table-format questions. Appendix B contains a copy of the survey.

The first section asked for background information on the firm, such as the firm's primary line of business, size, and the respondent's business unit's life-cycle stage. The remaining sections asked about the following:

- Current Infratechnology Costs: Respondents were asked to estimate their annual expenditure on infratechnology, including all labor, materials, and equipment expenditures for in-house

activities as well as contributions to consortia or partnership initiatives. This section also asked for information on

- – distribution of costs by expenditure category (i.e., labor, capital, and materials) and enabling technology category,
- – historical trends in infratechnology expenditures, and
- – use of and experience with technologies included in the TFAs.

- Impacts of Technology Improvements: This section asked respondents to consider how future improvements in one specific TFA might affect the key drivers of drug development and commercialization costs. Questions asked about the following:

  - – percentage changes in stage length and cost by R&D stage, manufacturing stage, or postmarket stage depending on which TFA the respondent selected;
  - – changes in the overall success rates (percentage of all drugs entering clinical trials that receive FDA approval);
  - – changes in the distribution of attrition rates (percentage for all drugs failing in clinical trials that fail during a specific clinical phase); and
  - – changes in the probability of recall once a drug receives approval.

- Potential Role for NIST: Interviewees were asked to what extent, if any, NIST might play a role in facilitating the technology improvements presented in the survey.

The survey instrument served as the general structure for the interviews. However, in many cases the unanticipated information, comments, and anecdotal examples obtained during the interviews proved most insightful.

4.2 INTERNET SURVEY OF BIOPHARMACEUTICAL FIRMS

RTI conducted an Internet survey to reach a broad audience of industry participants. The information obtained from this survey was intended to complement the more detailed qualitative information obtained through the expert interviews. The Internet survey's objectives were to

- obtain a national industry average of current infratechnology costs,

- identify the technology areas where the costs of inadequate infratechnology are greatest, and

- identify which components of the technology infrastructure could best be served by NIST's involvement.

The information gathered from the Internet survey helped quantify the magnitude of the current costs associated with the technology

infrastructure and confirm the impact estimates gathered from our technical expert interviews.

The primary channel used to distribute the online survey was BIO, a national trade association representing the biotechnology industry. BIO represents companies that apply biotechnology in a number of different industries, including alternative fuels, agriculture, and human therapeutics. BIO agreed to support our research efforts by distributing an invitational e-mail to a subset of the organization's members requesting their participation by completing the Internet survey.

BIO sent an e-mail message to a select list of its membership to inform them of the study and request their participation. The e-mail message was targeted to biotechnology firms affiliated with the application of biotechnology to human health applications. The invitational e-mail directed recipients to an RTI Web page that housed the survey.

Because BIO's membership is not limited to any particular type of biotechnology firm, we anticipated that companies that responded to the survey might represent an array of firms, including biopharmaceutical manufacturing, small biotechnology start-ups, and contract research organizations (CROs), as well as service firms providing support to larger biopharmaceutical firms in analytical, raw materials, or managerial support services.

RTI attempted to increase the number of Internet survey responses by sending similar invitational e-mails to various Internet-based professional networking communities and technology user groups such as a molecular profiling (gene/protein expression) technology users group. Many individuals responded to the invitational e-mail and completed at least some portion of the Internet survey; however, only 58 responses were sufficiently complete to be included in the impact analysis.

The Internet survey was anonymous (except if the respondent chose to self-identify), so we do not know what portion of industry sales or R&D expenditures was represented by these completed surveys.

Seventy-one percent of respondents to the Internet survey were involved with one or more of the R&D stages of the biopharmaceutical life cycle. Twenty-one percent were involved in commercial manufacturing, and the remaining 9% of respondents were associated with postmarket surveillance activities (see Table 4-2). Chapter 5 provides a discussion of the combined results of the interviews and survey.

5 Analysis Results

This chapter presents the analytical results from the infrastructure expenditure analysis and the counterfactual cost model described in Chapter 3. Study participants provided insights and data that permitted the calculation of three sets of results:

- estimated annual expenditures on the technology infrastructure and related activities by the biopharmaceutical sector;

- potential impacts an improved technology infrastructure holds for each TFA individually, including cost reductions; changes in stage length; and the probability of FDA approval from a cohort of INDs; and

- lower bound and upper bound cost model scenarios that account for the uncertainty surrounding participants' views.

The quantified potential impact calculations hold all other costs and production factors constant. Outside of this project's model, real-world shifts and reassignment of resources are possible and likely, as are unrelated process improvements that may lead to overall reductions in industry costs.

5.1 ANNUAL BIOPHARMACEUTICAL INDUSTRY EXPENDITURES ON THE TECHNOLOGY INFRASTRUCTURE

RTI estimates that the biopharmaceutical industry currently spends $1,219 million annually on technology infrastructure–related issues, encompassing $884 million in spending on drug R&D-related activities and $335 million on commercial manufacturing and postmarket surveillance activities. The empirical results presented represent a snapshot of the biopharmaceutical industry. Estimates were developed

using cross-sectional data informed by companies and experts working in the biopharmaceutical industry and therefore represent only a single point in time.

5.1.1 Annual Expenditures by Technology Focus Area

Study participants provided their companies' infrastructure-related expenditures by TFA. Assuming that their aggregated spending parallels the balance of the industry's, total annual expenditures by TFA can be estimated and are presented in Table 5-1:

- Gene expression systems and biomarkers accounted for over half of total technology infrastructure spending in the R&D segment at 30% and 24%, respectively.

- Bioimaging accounts for 15% and informatics for 22%.

- Study participants allocated the remaining 8% of infrastructure expenditures to other R&D activities and technology areas.

Table 5-1 also includes estimates of annual R&D spending and sales that participants attributed to the TFAs. Companies do not necessarily categorize infrastructure-related expenditures as R&D expenses, and a direct comparison ignores the fact that spending by TFA falls into multiple corporate spending accounts. However, the relative shares do provide a measure of the intensity with which firms invest in these TFAs. Survey data indicate that infrastructure spending accounted for 8.4% of R&D.

The commercial segment had spending of $335 million (approximately 8% of sales), broken out as follows:

- Commercial manufacturing accounted for 48% of total infrastructure expenditures in the commercial segment.

- Postmarket surveillance accounted for 52% of total infrastructure expenditures in the commercial segment.

The R&D segment clearly incurs the majority of expenditures; however, this distribution does not reflect any relative level of inadequacy among TFAs. Many environmental factors must be considered when comparing the magnitude of costs across these areas. For example, regulatory requirements may be significantly different in terms of testing frequency, number of measurements, tests, or analyses placed on the R&D effort than on commercial activities.

Table 5-1. Technology Infrastructure Expenditures by Technology Focus Area, 2005

Technology Focus Area	Technology Infrastructure Spending (millions)	Percentage Distribution	Relative Spending	Intensity
Bioimaging	$136	15%	$4,011 per scientist	1.3% of R&D
Biomarkers	$212	24%	$6,240 per scientist	2.0% of R&D
Bioinformatics	$198	22%	$5,813 per scientist	1.9% of R&D
Gene expression analysis	$265	30%	$7,800 per scientist	2.5% of R&D
Other	$73	8%	$2,136 per scientist	0.7% of R&D
Subtotal: R&D Activities	**$884**	100%		
Commercial manufacturing	$162	48%	$613,000 per approved drug	6.4% of sales
Postmarket surveillance	$173	52%	$656,000 per approved drug	2.0% of sales
Subtotal: Commercial Activities	**$335**	**100%**		
Industry Total	**$1,219**			

Source: RTI estimates.

5.1.2   Annual Expenditures by Cost Component

Table 5-2 disaggregates the total estimate of $1,219 million into capital, labor, and materials expenditures categories:

- $218 million or 18% in capital costs (includes all equipment, software, and licensing fees related to technology infrastructure),

- $817 million or 67% in labor costs (includes all research, implementation support activities, and participation in consortia related to technology infrastructure), and

- $184 million or 15% in materials costs (includes materials, reagents, and any other materials consumed in QA/QC processes).

Labor is the dominant cost component, accounting for two-thirds of the industry's total technology infrastructure expenditures, while the combined capital and materials represent the remaining 34%. Labor intensity is high in the R&D and postmarket segments, which rely heavily on human capital, and is lower in the capital-intensive manufacturing sector.

The cost per scientist (for R&D expenditures) and the cost per approved drug (for manufacturing and postmarket surveillance) were scaled to total industry spending using the total number of scientists and the total number of approved drugs, as discussed in the methodology chapter. The scaling factors are included in the table for reference.

Table 5-2. Annual Technology Infrastructure Expenditures by Cost Category, 2005

Expenditure Category	Percentage Category Distribution	Cost per Scientist[a]	Scaling Factor (Scientists)	Total Costs (millions)
Capital	17%	$4,389	34,000	$149
Labor	68%	$17,558	34,000	$597
Materials	16%	$4,053	34,000	$138
Subtotal of R&D Activities	**100%**	$26,000		**$884**

Expenditure Category	Percentage Distribution	Cost per Approved Drug[a]	Scaling Factor (Approved Drugs)	Total Costs (millions)
Capital	28%	$173,513	264	$46
Labor	47%	$286,407	264	$76
Materials	25%	$153,080	264	$40
Commercial Manufacturing	100%	$613,000		$162
Capital	13%	$87,467	264	$23
Labor	83%	$546,667	264	$144
Materials	3%	$21,867	264	$6
Postmarket Surveillance	100%	$656,000		$173
Commercial Activities		**$1,295,000**		**$335**
Industry Total				**$1,219**

[a]R&D activities = $/scientist; commercial activities = $/approved FDA drug

Note: Totals may not add because of rounding.

5.1.3   Annual Expenditures on R&D-Related Activities

The biopharmaceutical industry currently spends approximately **$136 million annually on the technology infrastructure supporting bioimaging.** Bioimaging is playing an increasingly important role in drug development partly because of increased technical capabilities and a focus by the industry on chronic diseases (Wang, 2005). Increasingly, FDA is requesting imaging studies to be included in the regulatory submissions for new drug approvals.

Spending on biomarkers totals $212 million annually for technology infrastructure to support biomarker discovery and validation. Researchers are using genomic or proteomic technologies such as SNP screening, RNA profiling, protein profiling, and pathway analysis to discover new biomarkers in DNA, RNA, and proteins. As stated earlier, the intent is for biomarkers to be used as substitutes or surrogates for clinical outcomes. They can provide early drug safety or efficacy

information to researchers on the biological response to treatments, potentially reducing the duration and cost of clinical trials (Gingsburg and Haga, 2006). The technology infrastructure supporting biomarkers is still in the early stages of development.

Bioinformatics spending was estimated at $198 million annually. Current efforts are fundamentally about the ability to exchange, interpret, and analyze data. In the words of one interviewee, a lack of standards "creates a constant struggle."

The industry's massive data production capacity relies on information technologies to manage and analyze data output. The other TFAs, particularly gene expression, rely on bioinformatics to investigate hypotheses, identify gene targets, and guide research strategies. Analysis results indicate that **gene expression was the largest infrastructural spending area at $265 million annually**.

In an ideal world, information flowing from external and internal sources, and across development stages (where possible), would converge in software applications that would evaluate drug candidates on an ongoing basis. However, interviewees reported that in reality data production capacity in the posthuman genome era outstrips the industry's ability to manage, coordinate, analyze, and communicate the resulting data across software systems and technology platforms. Yet these analyses inform decision making on a daily basis.

Many interviewees stated that their companies continually invest in their bioinformatics systems—software, hardware, and procedures—to make as effective use as possible of the genomic data produced internally or acquired from external databanks.

5.1.4 Annual Expenditures on Commercial Activities

Analysis results indicate that **commercial manufacturing infrastructure spending was $162 million annually**. Participants identified that technology infrastructure expenditures arise from data verification, validation, and QA/QC processes. Related to these technology infrastructures, problems may arise from data transfer from the development lab bench to production, data assembly, and interpretation at manufacturing facilities.

Manufacturing operations operate in a regulatory environment that requires validation of manufacturing processes. The data infrastructure inadequacies increase the cost of demonstrating regulatory compliance

and manufacturing costs by limiting firms' understanding of their manufacturing processes and constraining their ability to efficiently conduct process monitoring.

Results indicate that technology infrastructure spending associated with **postmarket surveillance was $173 million annually in 2005.** Participants in the study indicated that infrastructure expenditures accounted for 0.5% of annual sales. This result is somewhat higher than a recent empirical estimate of $56 million (0.3% of annual sales) for the average pharmaceutical company's expenditures on postmarket surveillance in 2003, though their average market size is larger (Ridley et al., 2006).

Previous studies calculated postmarket surveillance spending as a percentage of global sales, whereas the current study calculates it as a percentage of domestic sales in the United States. The denominator is smaller, yet because most biopharmaceutical companies are American, the postmarket surveillance infratechnology expenses accrue at R&D facilities located in the United States. Thus, overall results from this study are consistent with recent published estimates given the 3-year difference between estimates and the use of domestic rather than international sales.

Technology infrastructure costs related to postmarket surveillance were primarily associated with data collection, data entry and analysis, quality control, data systems design and validation activities, training on reporting adverse events, and product relabeling activities motivated by safety issues.

Participants we interviewed confirmed that labor expenses accounted for 83% of postmarket surveillance infratechnology spending. Capital spending was only 13% and consisted primarily of database design and information collection technologies. The remaining 3% was associated with materials expenses.

Postmarket surveillance costs have been increasing because of greater complexity in regulatory compliance, divergence between global regulatory bodies, and lower levels of data element harmonization through increased complexity in safety study design and reporting requirements. Participants suggested that, although overall infrastructure costs have risen over time, postmarket activities have been underfunded in previous years and are only recently reaching adequate levels as a result of increased attention to risk management.

5.2 POTENTIAL EFFICIENCY GAINS FROM AN IMPROVED TECHNOLOGY INFRASTRUCTURE

This section presents the potential efficiency gains from the counterfactual scenario detailing an improved technology infrastructure for each TFA. RTI presents results both as actual expenses and as present-value costs, which takes into account the time value of money and the opportunity cost savings to firms as the length of the drug development process is shortened due to an improve technology infrastructure.

Reducing R&D expenses while simultaneously shortening development time and improving research effectiveness generates savings on actual expenses. Given that R&D expenses for an approved drug can accrue over a period of 12 years, changing both the amount and timing of expenses can have a significant impact on benefits calculations (see Figure 5-1).

Figure 5-1. Illustration of How Expected Costs per Approved Drug Change with an Improved Technology Infrastructure

C_0 Baseline expected R&D costs per approved drug

C_1 Expected R&D costs per approved drug with an improved technology infrastructure

Present-value efficiency gains were calculated using the same 11.5% hurdle rate employed by the 2007 DiMasi and Grabowski study. Although the discount rate in this study is higher than the 7% social discount rate recommended by the Office of Management and Budget (OMB) for public benefits, the 11.5% rate better approximates the return-on-investment requirements faced by biopharmaceutical companies. Refer to Section 3.1 for a discussion of the 11.5% discount rate.

This section begins with a comparison of survey results on counterfactual drug approval rates, transition probabilities, and potential stage length and time reductions for each TFA except bioimaging. Survey results for bioimaging were inconclusive as respondents were not willing to speculate on the potential impacts for several key model parameters, although all were in agreement that improvements would generate significant benefits. In the absence of sufficient survey data, anecdotal evidence from study participants in this TFA was compiled to illustrate potential efficiency gains.

Study participants assessed the potential improvements with respect to a single TFA and improvements are presented on that basis, except in the summary section in which improvements are compiled to assess the potential simultaneous improvements to all TFAs' technology infrastructures.

5.2.1 Counterfactual Success Rates and Cost and Development Time Reductions for Biopharmaceutical INDs

Survey results indicate that an improved infrastructure would increase the probability that FDA would approve an IND. They also indicate that a greater share of INDs that fail would do so earlier in clinical trials. Although each clinical trial has the intent of studying a different aspect of an IND, each subsequent phase has a larger patient group, and the analytical studies conducted are cumulative. Greater confidence and assurance, conveyed through improved statistical power in comparatively smaller patient groups, translate into lower clinical trial costs under the counterfactual scenario.

Table 5-3 presents the overall changes in success and failure rates across clinical trial phases predicted by survey respondents by TFA. Participants were asked to assess how the typical distribution of failures within clinical trials would change under the improved infrastructure

Table 5-3. Comparison of Estimated Potential Success and Failure Rates by TFA

Technology Focus Area	IND Approval Probability	For INDs Failing in Clinical Trials, Percentage Failing by Phase[a]				Probability of Recall
		Phase I	Phase II	Phase III	Total	
Baseline	*30.2%*	*23.4%*	*52.4%*	*24.2%*	*100.0%*	*0.40%*
Biomarkers	41.0%	39.2%	37.0%	23.8%	100.0%	0.30%
Bioinformatics	40.0%	30.0%	40.5%	29.5%	100.0%	0.30%
Gene expression	45.0%	37.5%	35.5%	27.0%	100.0%	0.10%

[a]Using the baseline case as an example, if 100 INDs enter clinical trials, the IND approval probability suggests that ~30 INDs will eventually receive FDA approval. The remaining 70 INDs will then fail in one of the three clinical trial phases. During Phase I, ~16 INDs or 23% of the 70 INDs would fail. In Phase II, an additional 37 INDs or 53% of 70 are likely to fail. The remaining 17 or 24% of the 70 IND failures would then fail during Phase III.

Note: Comparable data for bioimaging could not be calculated; thus, bioimaging was excluded from this table. RTI estimates based on DiMasi and Grabowski (2007).

scenario. Holding the probability of FDA approval constant, the distribution of the failures across the three trial stages shifted toward earlier failures.

The DiMasi and Grabowski study forms the baseline for this analysis, which is that FDA ultimately approves 30.2% of biopharmaceutical INDs that enter clinical trials. Of the INDs that fail, 52.4% fail during Phase II and 24.2% fail in Phase III.

Several interviewees stated that improvements to each counterfactual technology infrastructure would improve researchers' knowledge of a drug's mechanism and its likely reception by the human body or the drug target. Overall, improved confidence in and assurance of candidates would result in better INDs entering clinical trials, suggesting that the drugs in clinical trials would be more likely to gain FDA approval.

An improved gene expression infrastructure holds the greatest potential (45.0% versus 30.2%), in terms of an anticipated improvement in FDA approval rates. Improved infrastructures for bioinformatics and biomarkers could boost IND approval probabilities to 40.0% and 41.0%, respectively. It is important to note, however, that no statements were made on how the absolute number of INDs submitted to the FDA for clinical trials may change.

A small number of larger biopharmaceutical firms reported that they would not adopt new industry standards after investing in and developing systems, protocols, and procedures that they view as effective. Yet, an

improved infrastructure may improve accuracy within the same protocol. Such firms may not adopt a new series of protocols, but if the measurement systems supporting their proprietary methods were to improve, the methods would also improve. Other respondents reported that improvements would have little to no impact on success rates. These individuals believed that improvements in technical infrastructure would not address the fact that it is not always possible to predict exactly how a novel compound would react until it is inside the body. However, an improved infrastructure would yield significant improvements in time and cost.

The results in Table 5-3 are in percentage terms because it is not possible to predict the absolute number of INDs that would enter clinical trials under the improved infrastructure scenarios. In terms of the distribution of INDs that fail during clinical trials, the percentage of failures shifted toward earlier failures. That is, a greater proportion of failures would occur in Phase I, which in turn leads to a lower rate of failures in Phase II where efficacy is evaluated for the first time. The proportion that fails in Phase III remained relatively constant or increased slightly in the improved infrastructure scenarios relative to the baseline. Such muted variation in Phase III failures between scenarios may reflect the limitations to our current scientific knowledge base of the interactions between chemical compounds and the human body. An increased understanding of biological systems and processes would likely have a greater impact on Phase III failures than an improved technology infrastructure.

Finally, respondents believed that the overall probability of recall once a drug had received approval was likely to decrease. Overall, most participants believed that the improved infrastructure scenario presented would have a positive effect, thus reducing the likelihood that a drug, once approved, would be recalled.

The magnitudes of the reductions in cost and stage length were typically associated with the stage where the technology was applied or used most often. For example, biomarkers were predicted to have little net impact on early discovery/preclinical and Phase I clinical trials because improvements could increase upfront costs, even though such expenditures could achieve costs and time savings in later Phase II and Phase III clinical trials, when researchers would have more efficient surrogates for clinical end points. Biomarkers have the ability to shorten the length of clinical trials by providing researchers with clinical

information on drug safety and efficacy more quickly. For example, surrogate biomarkers may provide information on the progression of a disease following treatment more quickly than without the biomarker. Biomarkers may also improve patient stratification when selecting volunteers for clinical trials. This may lead to more targeted clinical trials with higher success rates.

Conversely, gene expression impacts were concentrated in earlier stages of development. Improvements in comparability standards for gene expression and other infratechnologies could potentially improve the efficiency of discovery screening and preclinical PD/PK and ADME toxicity analysis. Study participants were consistent in their belief that an improved gene expression technology infrastructure would not have the same magnitude of impact on Phase I and II clinical trials as on earlier stage activities. Whereas survey results indicate the counterfactual technology infrastructure could yield discovery/preclinical phase cost reductions of 19% and time reductions of 26%, those same benefits are comparatively lower in Phase III clinical trials (8% reduction in both time and cost).

The trend for the counterfactual bioinformatics infrastructure parallels that for gene expression, but the data for bioimaging were inconclusive; therefore, bioimaging was not included in Table 5-4.

5.2.2 Potential Gains from Improvements in Bioinformatics

Table 5-5 presents the baseline data from the 2007 DiMasi and Grabowski study, and Table 5-6 illustrates how the counterfactual infrastructure for bioinformatics could affect the cost, development time, and probability of success for a new drug. According to the data supplied by experts, the counterfactual technology infrastructure could increase the probability of FDA approval for the average IND from 30.2% to 40% (see Section 3.3.7 for an accounting of the counterfactual bioinformatics infrastructure).

Participants stated that improved standardization in measurements, data formats, procedures, and protocols could reduce costs associated with redundant testing and thereby reduce stage length, while improvements in automation for data acquisition, interpretation, and feedback could also reduce costs. The reduction in current costs would result from efficiency gains and productivity improvements via improved access to more timely and accurate information, better understanding of biological processes, and more insightful analysis enabled by higher-quality data.

Table 5-4. Estimated Reductions in R&D Cost and Stage Length

Technology Focus Area	Stage	Expected Stage Cost per Investigational Compound (IND)	Stage Length (in months)
		Percentage Change from Baseline	
Bioinformatics			
	Discovery/preclinical	–20%	–17%
	Phase I	–15%	–12%
	Phase II	–13%	–11%
	Phase III	–13%	–8%
Biomarkers			
	Discovery/preclinical	–1%	–7%
	Phase I	–6%	–14%
	Phase II	–16%	–16%
	Phase III	–46%	–44%
Gene expression analysis			
	Discovery/preclinical	–19%	–26%
	Phase I	–15%	–15%
	Phase II	–11%	–10%
	Phase III	–8%	–8%

Note: Comparable data for bioimaging could not be calculated; thus, bioimaging was excluded from this table.

Table 5-5. Baseline Expected Cost per IND and per Approved Drug

Stage	Baseline Stage Cost, per IND (millions)	Probability of IND Entering Stage	Expected Baseline Stage Cost, per IND (millions)	Baseline Probability of IND Approval	Estimated Baseline Stage Cost, per Approved Drug (millions)	Baseline Stage Length (months)
Discovery/ preclinical	$59.9		$59.9	30.2%	$198.3	52.0
Phase I	$32.3	100%	$32.3	30.2%	$106.9	19.5
Phase II	$37.7	83.7%	$31.5	30.2%	$104.5	29.3
Phase III	$96.1	47.1%	$45.3	30.2%	$149.9	32.9
Total			$169.0		$559.6	133.7

Note: See Table 5-4 for estimated changes in FDA approval, distribution of IND failures within clinical trials, and probability of a recall of an FDA-approved drug. The period between completion of Phase III clinical trials and FDA approval is assumed to be 16 months.

Source: DiMasi and Grabowski (2007).

Table 5-6. Potential Efficiency Gains from an Improved Bioinformatics Infrastructure

Stage	Improved Stage Cost, per IND (millions)	Improved Probability of IND Entering Stage	Expected Stage Cost, per IND (millions)	Improved Probability of IND Approval	Estimated Stage Cost, per Approved Drug (millions)	Improved Stage Length (months)
Discovery/ preclinical	$47.7		$47.7	40.0%	$119.3	43.2
Phase I	$27.4	100%	$27.4	40.0%	$68.6	17.2
Phase II	$32.6	79.4%	$26.8	40.0%	$67.0	26.1
Phase III	$83.3	59.9%	$48.1	40.0%	$120.1	30.2
Total			$150.0		$375.0	116.6

Note: See Table 5-4 for estimated changes in FDA approval, distribution of IND failures within clinical trials, and probability of a recall of an FDA-approved drug. Table 5-5 includes percent changes in cost and phase length relative to the baseline. The period between completion of Phase III clinical trials and FDA approval is assumed to be 16 months.

Source: RTI estimates based on DiMasi and Grabowski (2007).

Bioinformatics activities are closely linked to microarray experiments, flow cytometry, and other data acquisition systems because data production systems feed databases. The quality and accuracy of the data and the confidence researchers have in it directly affect the confidence they have in the subsequent statistical analyses. Thus, infratechnology activities and investments discussed in the gene expression TFA would have a symbiotic relationship with the bioinformatics TFA.

The expected reductions in costs and stage lengths change the stage cost per IND and the stage length per IND, which are represented in the first and last data columns in Table 5-7. But the improved infrastructure is estimated to also increase the probability of FDA approval and cause INDs that fail to do so earlier in clinical trials.

Companies take a portfolio approach to drug development, meaning that they expect that some INDs to fail. Each IND that becomes an FDA-approved drug carries the costs of failed INDs from its cohort, thereby representing the total actual cost of developing a new drug. Mathematically, the expected cost per IND is the sum of each stage cost multiplied by the probability that the IND reaches that stage.

Once the clinical trial data and final application are submitted to FDA, there is still the probability that FDA will not approve the IND.

Table 5-7. Potential Cost Reductions per IND and per Approved Drug: Bioinformatics

	Baseline (millions)	Improved Infrastructure (millions)	Percentage Change
Expected actual R&D cost per IND	$169.0	$150.0	–11%
Expected actual R&D cost per approved drug	$559.6	$375.0	–33%
Expected present-value R&D cost per IND	$374.7	$298.5	–20%
Expected present-value R&D cost per approved drug	$1,240.9	$746.3	–40%

Source: RTI estimates based on DiMasi and Grabowski (2007).

Participants believe that an improved gene expression infrastructure would improve R&D effectiveness to such an extent that it would be possible to have an average approval probability of 45%. Given that companies expect each IND to have an average cost of $150 million, but also expect that only 45% of INDs that complete clinical trials would be approved, they expect the average cost of actually receiving an approval to be $375 million. This figure is the expected cost per IND divided by the probability of FDA approval.

Note that, although the cost per IND for all stages has decreased from the baseline in the improved scenario, the probability of entering Phase III trials has increased because the probability of the IND being a success has increased. As a result, more INDs may enter Phase III trials and incur additional costs. A larger number of INDs from a portfolio would be commercialized. Total costs for the portfolio would increase but so would the return on the total portfolio.

Thus, when taking into consideration improvements to the probability of success, cost, and development times, the estimated actual cost per approved drug drops from $559.6 million to $375.0 million, a 33% reduction (see Table 5-7). In comparison, the expected actual R&D expense per IND is $19.0 million lower.

Accounting for firms' investment horizons and their opportunity costs with the weighted average 11.5% cost of capital provides a more dynamic view of these savings. Development times are expected to decrease from just over 11 years to just under 10 years, on average. Baseline and counterfactual present-value costs were adjusted to account for difference in timing and relative magnitude of cash flows. The efficiency gains result in a present-value cost per IND of $298.5 million under the counterfactual scenario, which is a reduction of 20% from the baseline.

The estimated present-value cost per approved drug changes from
$1,241 million to $746 million, a reduction of 40%, all else held equal.
The FDA Critical Path reports that "…some believe extensive use of *in
silico* technologies could reduce the overall cost of drug development by
as much as 50%" (FDA, 2004b). The results from this analysis, while
lower than suggested in the FDA report, are reasonably close to the
FDA's rough estimate.

5.2.3   Potential Gains from Improvements in Biomarkers

Biomarkers have the potential to lower the cost of developing new drugs
by providing potential safety and efficacy information to researchers with
greater precision and in less time. Biomarkers could have the largest
impact on Phase II and III clinical trials, as respondents estimated that
biomarkers could reduce these costs by as much as 46% and shorten
stage length by 44%. During Phase III, researchers must validate the
efficacy of drug treatment in the largest patient group. Improved
biomarkers would allow trial design to better stratify the patient cohorts
and reduce uncertainty with respect to clinical endpoints.

The improved infrastructure scenario could also increase the success
rate for new drug candidates entering clinical trials. However, the largest
impact would come from the improvement in researchers' ability to
predict biological responses in humans earlier in the drug development
process. While the improved infrastructure scenario could increase
productivity in discovery and preclinical testing phases, the largest
impact would come from the ability to shorten the clinical trials. If
researchers are able to obtain critical drug efficacy information faster
during clinical trials, it would greatly reduce the number of months and
potentially the number of patients required to conduct a clinical trial.

One participant cited that with improved biomarkers clinicians and drug
researchers could better understand the immune system response, as
well as body tolerances for drug dosages. However, another participant
commented that, although safety biomarkers could potentially lower
clinical trial costs, the complexity of the human body and disease
pathways make it highly unlikely that such biomarkers could exist within
the next 10 years.

Table 5-8 presents survey results on how biomarkers would affect
development times and costs. While impacts on the discovery/preclinical
stage are comparatively low (1% of cost and 7% of lead time), the
benefits of biomarkers grow as clinical trial patient groups become larger

Table 5-8. Potential Efficiency Gains from an Improved Biomarkers Infrastructure

Stage	Improved Stage Cost, per IND (millions)	Improved Probability of IND Entering Stage	Expected Stage Cost, per IND (millions)	Improved Probability of IND Approval	Estimated Stage Cost, per Approved Drug (millions)	Improved Stage Length (months)
Discovery/ preclinical	$59.4		$59.4	41.0%	$144.9	48.4
Phase I	$30.3	100%	$30.3	41.0%	$74.0	16.8
Phase II	$31.7	76.9%	$24.3	41.0%	$59.4	24.6
Phase III	$51.9	55.0%	$28.6	41.0%	$69.7	18.4
Total			$142.7		$347.9	108.2

Note: See Table 5-4 for estimated changes in FDA approval, distribution of IND failures within clinical trials, and probability of a recall of an FDA-approved drug. Table 5-5 includes percentage changes in cost and phase length relative to the baseline. The period between completion of Phase III clinical trials and FDA approval is assumed to be 16 months. Source: RTI estimates based on DiMasi and Grabowski (2007).

in each stage. Study participants believed that improved biomarker technology infrastructure could lead to Phase III clinical trial cost reductions of 46% and lead time reductions of 44%, as shown in Table 5-4.

When taking into consideration improvements to the probability of success, cost, and development times, the estimated actual cost per approved drug drops from $559.6 million to $347.9 million, a reduction of 38% due in large part to reduced clinical trial costs (see Table 5-9).

As with bioinformatics, RTI also accounted for firms' investment horizons and their opportunity costs using the weighted average 11.5% biopharmaceutical cost of capital. Average time from discovery through the completion of Phase III clinical trials is expected to decline from over 11 years to 9 years. Baseline and counterfactual present-value costs were adjusted to account for differences in timing and relative magnitude of cash flows. The efficiency gains results in a present-value cost per IND of $278 million under the counterfactual scenario, which is a reduction of 26% from the baseline. The estimated present-value cost per approved drug changes from $1,241 million to $677 million, a reduction of 45%.

5.2.4 Potential Gains from Improvements in Gene and Protein Expression

As discussed previously, gene expression and bioinformatics processes are tightly integrated. For example, microarray analyses generate data

Table 5-9. Potential Cost Reductions per IND and per Approved Drug: Biomarkers

	Baseline (millions)	Improved Infrastructure (millions)	Percentage Change
Expected actual R&D cost per IND	$169.0	$142.7	–16%
Expected actual R&D cost per approved drug	$559.6	$347.9	–38%
Expected present-value R&D cost per IND	$374.7	$277.5	–26%
Expected present-value R&D cost per approved drug	$1,240.9	$676.9	–45%

Source: RTI estimates based on DiMasi and Grabowski (2007).

that informaticists use to develop lists and identify clusters of genes expressed in tissue samples and the genes' functions. Researchers evaluate the data output and determine and refine research strategies. Gene expression data may be developed internally or be acquired from outside the organization. Large genomic data sets are publicly available, but interviewees in this TFA stated that most gene expression data are produced internally. Indeed, external databases may not be usable if the data they contain cannot be integrated with information generated internally.

Standardization has the potential to make acquiring and evaluating gene expression data less resource intensive. One interviewee provided the following example: if a lab runs three samples per patient during a clinical trial, and there are 300 patients, it would expect costs of $670,000. Expenses would be $130,000 for labor, $360,000 for microarrays, and $180,000 for reagents. A suite of standards and measurement tools could enable a process in which only 200 patients were needed to get the same statistical power as 300 patients, thereby reducing the lab's clinical trial costs.

Ultimately, firms want drug targets to fail as early in the process as possible. Failing early prevents resources from being assigned to the project in the first place, which is important because the further an erroneous target moves forward in the discovery and preclinical phase, the greater is the cost that will need to be distributed over successful projects.

Table 5-10 shows how study participants believed the improved technology infrastructure for gene expression presented in Section 3.3.7 would affect costs, lead times, and the probability of having an IND

Table 5-10. Potential Efficiency Gains from an Improved Gene Expression Infrastructure

Stage	Improved Stage Cost, per IND (millions)	Improved Probability of IND Entering Stage	Expected Stage Cost, per IND (millions)	Improved Probability of IND Approval	Estimated Stage Cost, per Approved Drug (millions)	Improved Stage Length (months)
Discovery/ preclinical	$48.4		$48.4	45.0%	$107.6	38.5
Phase I	$27.4	100%	$27.4	45.0%	$61.0	16.6
Phase II	$33.5	79.4%	$26.6	45.0%	$59.0	26.4
Phase III	$88.9	59.9%	$53.2	45.0%	$118.2	30.4
Total			$155.6		$345.8	111.9

Note: See Table 5-4 for estimated changes in FDA approval, distribution of IND failures within clinical trials, and probability of a recall of an FDA-approved drug. Table 5-5 includes percentage changes in cost and phase length relative to the baseline. The period between completion of Phase III clinical trials and FDA approval is assumed to be 16 months. Source: RTI estimates based on DiMasi and Grabowski (2007).

become an FDA-approved drug. One company thought that an ideal world would allow one lab of four technicians and six Ph.D. analysts to be five times as productive. That company further added that a 50% improvement would have the same magnitude of impact because of the relationship between information flows in the discovery process.

When taking into consideration improvements to the probability of success, cost, and development times, the estimated actual cost per approved drug drops from $559.6 million to $346 million, a reduction of 38% (see Table 5-11). The expected actual R&D expenses per IND are more than $13 million lower. Once again, these estimates are adjusted for the fact that close to 60% of INDs enter Phase II trials under the improved scenario.

Accounting for firms' investment horizons and their opportunity costs with the weighted average 11.5% cost of capital provides a more dynamic view of these savings. Average time from discovery through the completion of Phase III trials is expected to decline from over 11 years to approximately 10 years under this scenario. Baseline and counterfactual costs were adjusted to account for difference in timing and relative magnitude of cash flows. The efficiency gains result in a present-value cost per IND of $304 million under the counterfactual scenario, which is a reduction of 19% from the baseline. The expected present-value cost per approved drug changes from $1,241 million to $676 million, a reduction of 45%, all else held equal.

Table 5-11. Potential Cost Reductions per IND and per Approved Drug: Gene Expression

	Baseline (millions)	Improved Infrastructure (millions)	Percentage Change
Expected actual R&D cost per IND	$169.0	$155.6	–8%
Expected actual R&D cost per approved drug	$559.6	$345.8	–38%
Expected present-value R&D cost per IND	$374.7	$304.2	–19%
Expected present-value R&D cost per approved drug	$1,240.9	$676.0	–45%

Source: RTI estimates based on DiMasi and Grabowski (2007).

5.2.5 Potential Gains from Improvements in Bioimaging

Respondents were unable to quantify metrics needed to populate the improved infrastructure model for bioimaging. However, they did provide quantitative examples of how improvements in infrastructure might reduce current costs, identifying potential opportunities for reductions in R&D stage costs and duration through improving access to image capture and analysis technology, data standardization, and data management.

For example, typical costs of a Phase III imaging study could run as high as $100 million. Such studies can include as many as a 1,000 patients at 100 different clinical sites. Interviews indicated that a cost savings of 15% could result under the improved infrastructure scenario. The impacts would include the following:

- **Cost reduction** of 10% due to
  - reduction in number of subjects for each clinical study
  - compressed timeline to conduct study
  - reduction of internal resources and staff (i.e., CRO radiologist costs)
- **Cost reduction** of 5% due to
  - mitigation of the courier costs to transport images to an analysis lab for interpretation

Inadequate ability to acquire and effectively use bioimaging technology has resulted in increased errors in imaging data and time delays in conducting imaging analysis. To date, the industry has relied on major research hospitals such as MD Anderson, the Mayo Clinic, Sloan-Kettering, Duke, and Johns Hopkins to conduct clinical imaging studies,

and image analysis and interpretation has been largely outsourced to CROs.

Increased demand for clinical imaging studies is putting additional burden on the imaging hospitals conducting these studies. Currently it takes between 3 and 4 months from image creation to image interpretation. This time lag makes it impossible to look at a single subject over time. One scientist reported that by the time a baseline scan is interpreted too many things have changed to isolate any positive treatment effects. As a result, clinical trials have to increase the number of subjects to obtain statistically valid results.

The variance in imaging protocols and data requirements has caused costly time delays and increased errors in image capture. Errors have been increasing as hospitals and imaging technicians are asked to follow a diverse set of imaging study protocols and data requirements defined by individual biopharmaceutical firms. Time delays associated with correcting errors can be significant because analysis of imaging data is conducted by a CRO in a different location. The transfer of data can sometimes take weeks. As a result, errors are discovered long after the data are originally captured.

Drug companies must validate images from large numbers of globally dispersed clinical trial sites. Current levels of data standardization and image labeling and formatting are inadequate and require additional effort to ensure the comparability of images across trial sites. Any standards also have to meet the privacy guidelines of the Health Insurance Portability and Accountability Act (HIPAA).

The American College of Radiology and the National Electrical Manufacturers Association formed a joint committee to develop a Standard for Digital Imaging and Communications in Medicine known as DICOM. Study participants reported that this effort was a step in the right direction, but that it was not sufficient to meet current or future needs.

In addition, the biopharmaceutical industry lacks the digital data storage capacity or bioinformatics database infrastructure to archive, store, and retrieve image data effectively. This problem will increase in the near future. Scientists predict that, as imaging technology improves, the size of the typical imaging file will also increase.

5.2.6 Potential Impacts on Commercial Manufacturing Costs

Whereas gains from an improved infrastructure were measured relative to the average cost per IND and per FDA-approved drug on the R&D side, the analysis estimated efficiency gains relative to *total* industry manufacturing expenses. The estimated efficiency gains in this report for manufacturing and R&D are therefore not additive. This is because information on average or "typical" products was not available publicly, requiring RTI to measure potential gains relative to estimated total industry manufacturing costs.

Study participants reported that manufacturing costs amounted to 17% of their companies' total annual sales. The impact results reported in Table 5-12 were based on the total annual domestic biopharmaceutical sales of $37.3 billion in 2005 as reported in company annual 10-K reports. Thus, baseline total biopharmaceutical production costs approximated $6.3 billion in 2005. RTI applied Frost and Sullivan's 2004 estimates to distribute total production costs across individual manufacturing phases.

An inadequate data infrastructure impedes the efficient transfer of scientific, process, and regulatory compliance data from R&D to manufacturing. These data transfer complications can result in delays in production and additional labor costs associated with tracking down or regenerating lost information.

Improvement to the technology infrastructure to support manufacturing could reduce manufacturing costs by 23%, according to data supplied by study participants. This 23% estimate represents a reduction of $1.5 billion from estimated 2005 annual manufacturing costs. Table 5-12 reports the percentage reductions in cost by phase from the estimated baseline.

Study participants provided their estimates of how an improved infrastructure would reduce manufacturing costs by phase. They estimated a 29% reduction in preproduction costs and a 22% reduction in downstream processing costs, among other impacts, that would result from the improved infrastructure described in Section 3.3.7. That counterfactual technology infrastructure was the implementation of the FDA's PAT initiative.

Table 5-12. Impact on Commercial Manufacturing Costs

Stage/Activity Cost	Baseline Production Costs		Potential Change in Cost by Stage/Activity		
	Percentage of Total[a]	Baseline Total (millions)	Percentage Change	Change in Cost (millions)	Costs under an Improved Infrastructure (millions)
Preproduction	30%	$1,900	−29%	−$551	$1,349
Upstream processing	20%	$1,267	−18%	−$228	$1,035
Downstream processing	40%	$2,533	−22%	−$557	$1,976
Process monitoring and quality assurance testing	10%	$633	−23%	−$146	$491
Total commercial manufacturing costs		**$6,333**		**−$1,482**	**$4,851**

[a]From Frost and Sullivan (2004).
Source: RTI estimates.

Industry experts expect that PAT would increase production yields, reduce time delays, and reduce product quality variability through continuous real-time process monitoring and sample measurement. For example, cell culture development may take 12 to 14 days. Once sufficient yields are produced, a sample is taken and sent to an analytics lab. The analysis requires an additional 5 to 10 days to run the sample and generate results.

Improved technology infrastructure through PAT would allow for real-time analysis, which may

- allow process engineers to react to information as it is generated,
- potentially eliminate as many as 14 days of off-line analysis,
- eliminate a number of bioreactor runs, and
- generate labor savings.

One interviewee stated that the cost of batches lost during production because of an equipment failure could average $1 million. PAT aims to reduce variability in product profiles between batch runs by

- providing a better understanding of product tolerances,
- lowering the standard deviation and coefficient of variation in processes, and
- optimizing buffer volumes required during purification stages.

The net impact would be fewer variations in end-product quality. Cost savings would accrue through such impacts as using fewer raw materials, incorporating fewer manufacturing processes, and lowering inventory and facility expenditures.

5.2.7 Potential Impacts on Postmarket Surveillance Costs

Analysis of the potential that an improved infrastructure holds for postmarket surveillance costs did not provide sufficient insight with which to estimate cost reductions. In most cases, study participants agreed that gains were possible, but they also stated that those gains were too intangible for them to assign percentage reductions to. Several experts suggested that the variation in costs per postmarket surveillance study was simply too high to estimate potential across the board reductions.

Major cost components in postmarket studies relate to the design of a core reporting system, data management, clinical site management, and analysis. Interviewees hypothesize that the improved infrastructure proposed in Section 3.3.7 would at least offer the following impacts:

- data management activities could be reduced by as much as 35% globally,
- clinical site management activities could be reduced by as much as 20%, and
- data analysis could be reduced by as much as 25%, given standardized reporting and consistent data formats.

Data on adverse events and postmarket safety must be collected across regional sites that may employ differing, and often competing, tracking systems maintained in multiple languages. The situation is made more complicated by reporting obligations that vary by country. Firms dedicate labor and resources to developing reporting protocols and training clinicians on data entry.

5.3 COMBINED POTENTIAL IMPACT OF AN IMPROVED TECHNOLOGY INFRASTRUCTURE

The preceding section reviewed the efficiency gains study participants believed the counterfactual technology infrastructure would have for each TFA individually. This section presents estimates on the gains the *combined* impacts from the four R&D TFAs would have on stage length and costs.

There exists some potential for varying degrees of substitution among the TFAs that the study was not able to capture. Some counterfactual infrastructural improvements—like gains in gene expression and bioinformatics, which were evaluated independently—exhibited at least some influence across all phases of R&D. The efficiency gains presented in Section 5.2, therefore, are not mutually exclusive and cannot be summed across the four TFAs because summing impacts would overestimate potential benefits.

Therefore, RTI developed lower- and upper-bound estimates where the lower bound was composed of the smallest estimated gains in each stage across all the TFAs and the upper bound was composed of the greatest estimated gains. The methodology used to develop the lower-bound (conservative) and upper-bound (optimistic) estimates entailed looking across all TFAs and selecting the least and most optimistic impact estimates for each of the four major cost parameters (i.e., direct stage costs, stage durations, IND approval rates, and transition probabilities) that drive the economic model results. The resulting ranges of estimates for our four key cost parameters are individual TFA estimates rather than an average taken across TFAs.

Table 5-13 contains the counterfactual approval rates and distribution of failures for INDs during clinical trials that are implicit in the results depicted in Tables 5-14 and 5-15. The lower-bound scenario included the smallest gain in terms of time and cost by stage (e.g., biomarkers in preclinical and Phase III clinical trials, bioinformatics in Phase I, gene expression in Phase II). The inverse was true for the upper-bound scenario.

The lower-bound scenario posited that (1) the probability an IND would be approved by FDA increased from 30.2% to 40.0%, (2) INDs failed slightly earlier in clinical trial phases on average, and (3) the probability an approved drug would be recalled declined from 0.4% to 0.3%. Study participants whose expertise was predominantly in gene expression and bioinformatics had consistent views on the possible magnitude of the impact. The upper-bound scenario data are strongly influenced by gains from biomarkers.

Table 5-13. Combined Scenarios' Potential Improvement in IND Success and Failure Rates

Technology Focus Area	IND Approval Probability	For INDs Failing in Clinical Trials, Percentage Failing by Phase				Probability of Recall
		Phase I	Phase II	Phase III	Total	
Baseline	*30.2%*	*23.4%*	*52.4%*	*24.2%*	*100%*	*0.40%*
Individual Scenarios						
Biomarkers	41.0%	**39.2%**	**37.0%**	**23.8%**	100%	**0.30%**
Bioinformatics	**40.0%**	**30.0%**	**40.5%**	**29.5%**	100%	0.30%
Gene expression	**45.5%**	37.5%	35.5%	27.0%	100%	**0.10%**
Combined Scenarios						
Lower bound	40%	30%	41%	30%	100%	0.30%
Upper bound	45%	39%	37%	24%	100%	0.10%

%	= Lower bound
%	= Upper bound

Note: Comparable data for bioimaging could not be calculated; thus, bioimaging was excluded from this table. RTI estimates based on DiMasi and Grabowski (2007).

Table 5-14. Combined Scenarios' Potential Efficiency Gains

Stage	Baseline		Improved Infrastructure Lower-Bound Scenario		Improved Infrastructure Upper-Bound Scenario	
	Expected Stage Cost per Approved Drug (millions)	Stage Length (months)	Estimated Stage Cost per Approved Drug (millions)	Stage Length (months)	Estimated Stage Cost per Approved Drug (millions)	Stage Length (months)
Discovery/preclinical	$198.3	52.0	$148.5	48.4	$106.0	38.5
Phase I	$106.9	19.5	$75.9	17.2	$61.0	16.6
Phase II	$104.5	29.3	$68.6	26.4	$55.2	24.6
Phase III	$149.9	32.9	$128.2	30.4	$67.0	18.4
Total	$559.6	133.7	$421.2	122.4	$289.2	98.1

Source: RTI estimates based on DiMasi and Grabowski (2007).

Table 5-15. Combined Scenarios' Potential Cost Reductions

	Baseline (millions)	Lower-Bound Scenario: Improved Infrastructure (millions)	Upper-Bound Scenario: Upper Bound Improved Infrastructure (millions)	Percentage Change
Expected actual R&D cost per IND	$169.0	$168.5	$130.1	–0 to –23%
Expected actual R&D cost per approved drug	$559.6	$421.2	$289.2	–25 to –48%
Expected present-value R&D cost per IND	$374.7	$347.8	$239.9	–7 to –36%
Expected present-value R&D cost per approved drug	$1,240.9	$869.6	$533.1	–30 to –57%

Source: RTI estimates based on DiMasi and Grabowski (2007).

Following this methodology, the combined technology infrastructure scenario could reduce the expected actual R&D cost for a new FDA-approved drug to $421 million under the lower-bound scenario and $289 million under the upper-bound scenario (see Tables 5-14 and 5-15). These estimates are 48% and 25% less, respectively, than the baseline cost estimated by DiMasi and Grabowski (2007).

Whereas the average time to take a candidate from discovery through Phase II clinical trials is currently around 11 years, the least optimistic results suggest that development time would be reduced to around 10 years, and the most optimistic scenario suggests that the time could be reduced to slightly more than 8 years.

The combined scenario results are highly sensitive to the impact an improved biomarker infrastructure could have on clinical trials, and the results in Tables 5-13 through 5-15 should be interpreted cautiously. Note, in particular, the reduction in time for Phase III. Whereas the baseline length is 32.9 months, the lower bound from the survey results is 30.4, and the upper bound is 18.4, which was a consensus view from biomarkers experts. These data illustrate the magnitude of impact that study participants believed an improved biomarker infrastructure could have. However, they also illustrate the sensitivity of upper-bound estimates to stakeholders' views on how a nascent technology may mature. Their views with respect to timing and ultimate impact introduce relatively greater uncertainty into the results than is the case for estimates obtained for bioinformatics and gene expression TFAs.

6 Conclusion

This study estimates the biopharmaceutical industry's annual spending on technology infrastructure–related investments and activities. It also identifies key areas within the technology infrastructure in which industry has underinvested because of market and technical barriers or because of unattractive risk-reward ratios at the firm level. The potential efficiency gains an improved technology infrastructure holds for the industry are estimated relative to the current costs of developing, manufacturing, and monitoring successful FDA-approved biopharmaceutical drugs. This chapter offers some summary remarks regarding the study's findings.

6.1 CURRENT STATE OF THE BIOPHARMACEUTICAL TECHNOLOGY INFRASTRUCTURE

The biopharmaceutical industry's technology infrastructure is not only composed of advanced laboratory instrumentation, chemistries, and information technologies but also of standard operating protocols, test methods, data, and institutional memory that have accrued over time.

Implicit in this study's findings is that biopharmaceutical infrastructure varies across companies. The relative sophistication of any one firm's infrastructure is a function of the amount of intellectual capital it has acquired as much as its investments in tangible assets such as instruments and software. These differences result from firms reacting differently to and overcoming individual technical barriers differently in the absence of well-coordinated industry standardization. In addition, these differences go beyond R&D stages and are manifest in varying quality assurance and control programs (QC/QA) as well as adverse event reporting.

One company RTI interviewed mandated strict adherence to its defined corporate standards and protocols and maintained program areas that developed tools in-house for communicating and representing information. External data not meeting defined screening criteria were routinely rejected to avoid introducing additional uncertainties into the company's R&D. The screening standards go beyond those for data formats. This organization recognizes gaps in the broader industry technology infrastructure inclusive of laboratory and QA/QC procedures and exercises extreme caution as a consequence.

Most experts interviewed over the course of this study conceptualized technology infrastructure expenditures into two general categories. The first is that these expenditures are, in part, an investment in current and future R&D efficiency. The labor effort, systems, and instruments expended are essential, unavoidable, and integral to a firm's primary economic activity. The second is that these expenditures represent costs incurred to develop work-arounds and overcome technical barriers stemming from a pervasive lack of standardization.

As a result, organizations view the current level of expenditures on infrastructure technologies to be above the level required with a more efficient infrastructure. Whereas the industry expended $884 million on the R&D technology infrastructure and $335 million on the commercial manufacturing and postmarket surveillance infrastructure (see Figure 6-1), some portion of that total $1,219 in spending could have been avoided with an improved infrastructure.

Challenges in the biopharmaceutical technology infrastructure and the differences in how organizations respond to these challenges are rooted in

- the inherent trial and error of the drug discovery and development process;
- development times averaging 12 years, which is compounded by changes in regulatory requirements, information systems, and procedures;
- variability in methodologies and protocols that acquire information and variations in how that information is described and characterized;
- few industry standards for ontologies, data formats, and data communications systems;

Figure 6-1. Annual Biopharmaceutical Industry Technology Infrastructure Spending

Annual Biopharmaceutical Technology Infrastructure Expenditures: $1,219 million
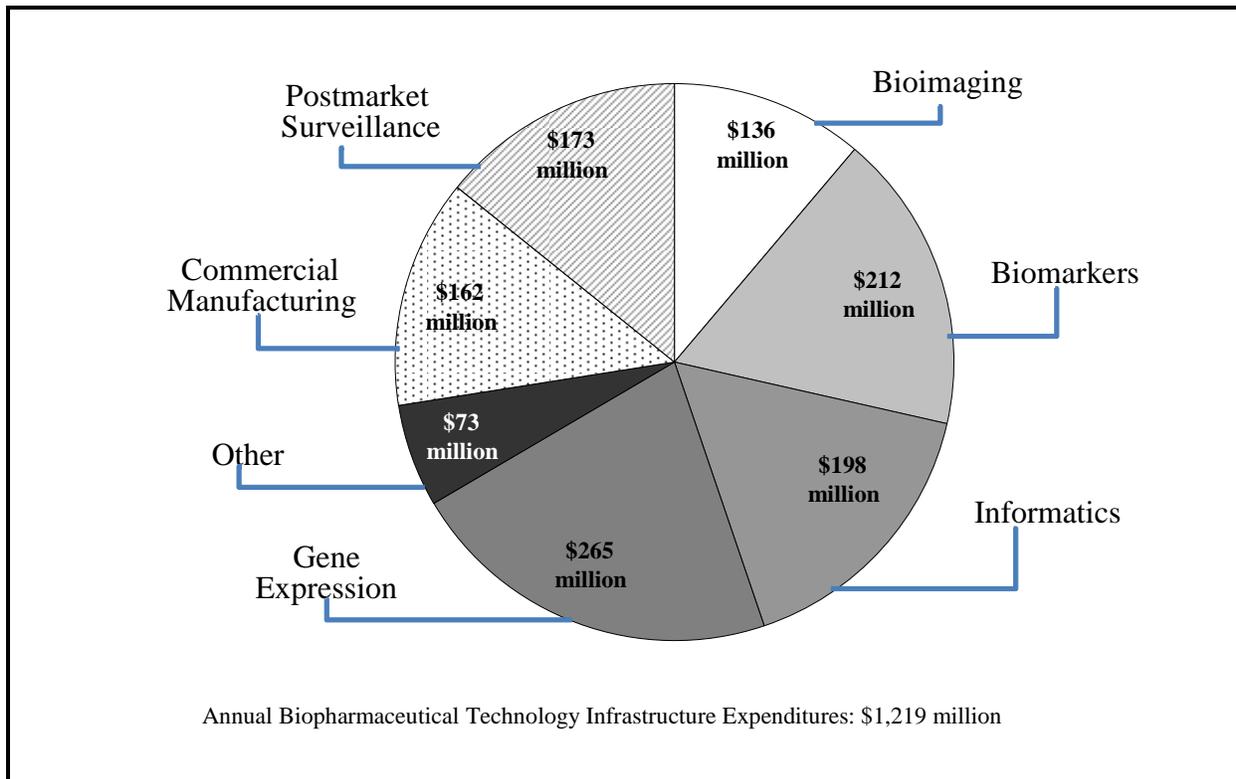
- the rapid introduction and adoption of data acquisition technologies (which far outpaces the development of industry's ability to manage, communicate, analyze, and synthesize data); and

- changes in the regulatory environment in a diverse set of countries and foreign languages;

Semiformal industry groups have organized to study how data are acquired, analyzed, and interpreted differently among technology platforms in the hope of guiding the industry toward greater standardization. However, the absence of financial and technical resources, competing research priorities, and the length of commitment required impedes their progress.

Before 2000, research methods were largely manual and time based—large volumes of data were not being produced, stored, exchanged, or analyzed. Analyses had many replicates, but few variables—meaning that processes were iterative, time consuming, and data management needs were predictable. Then, there was a technology leap-frogging event in the late 1990s when high-throughput laboratory instrumentation

went online and enabled rapid, inexpensive acquisition of genomic data, including the human genome (O'Connor et al., 2007). Laboratories found themselves generating in only 1 day 100 times or more data than they previously generated in 1 month.

Many organizations have realized that standardization is needed to limit uncertainty, increase confidence, and give information users some assurance of the information's accuracy. In some instances, companies and their vendors have researched, developed, and adopted ad-hoc standards and protocols, and these processes can work well internally.

However, company or laboratory standards are not an efficient replacement for industry standards, particularly in the current environment in which massive volumes of data are exchanged between laboratories, are compared over time, and used to analyze dynamic biological specimens. The pragmatism of encountering an infratechnology problem and responding on an ad-hoc basis translates to multiple responses to similar problems, which in turn impedes comparability and information exchange. Given the complexity of genomic information, conceiving or characterizing data element descriptors differently or not imparting how data interrelate can preclude any meaningful comparison or integration.

The experts RTI interviewed said that problems are growing, but that many companies have yet to see the extent of the problem because "they haven't gotten there technically yet."

6.2 CHARACTERIZING POTENTIAL EFFICIENCY GAINS FROM AN IMPROVED TECHNOLOGY INFRASTRUCTURE

Rather than ask the industry how expenditures could have been different in 2005, this study took the approach of evaluating how a specific set of improvements could increase efficiency going forward. In-depth interviews with people in industry, academia, and government agencies and an Internet survey quantified the impact a series of feasible infrastructure improvements would have on R&D and post-FDA approval activities. Stakeholders offered their views on how specific improvements to the technology infrastructure could

- lower the development cost of the average biopharmaceutical drug,

- increase the probability the drug would be approved by FDA by enhancing data quality and analytical methods,

- shorten the drug's time to market,

- lower the ongoing costs for manufacturing that drug and improve manufacturing tolerances, and

- make the postmarket surveillance infrastructure more efficient and responsive.

Even though industry has made significant investments in some areas of technical infrastructure over the past several years, the results of this study indicate that the expected actual expense for a new approved drug could be reduced by 25% under a lower bound scenario and by 48% under an upper bound scenario (see Table 6-1). This study also found that commercial manufacturing costs could be reduced by up to 22%.

A 25% to 48% improvement is significant and would require both substantial and broad-based advances in a range of technical infrastructures. However, industry respondents to the surveys conducted for this study believe such advances are feasible within a reasonable timeframe, perhaps 5 to 10 years. These results are consistent with a 2004 FDA report in which one expert suggested that biomarkers could reduce the cost of developing a new drug by 50% (FDA, 2004b).

Moving forward, drug development—for traditional products as well as biopharmaceuticals—will experience a global shift away from the lab toward computational systems that analyze drug candidates, predict cellular responses, and more effectively select good candidates for clinical trials. The systems biology approach to the development of new medicines discussed in this report's introduction requires a technology infrastructure that supports and integrates each of this study's TFAs. Databases of accurate, well-defined information comparable across both time and technology platforms rely as much on how data are captured as on the sophistication of the algorithms, ontologies, and formats used to integrate and understand the information. Protocols and measurement systems that govern how and at what quality level information is integrated into databases are needed. These elements are essential if the potential advantages of personalized medicine are to be realized in which patients are prescribed drug therapies that are most effective for their individual genetic make-up.

Table 6-1. Potential Cost Reductions in Biopharmaceutical Development with an Improved Technology Infrastructure

Technology Focus Area (TFA)	Expected Actual Cost per Approved Drug (millions)	Percent Change from Baseline	Expected Capitalized Cost per Approved Drug (millions)	Percent Change from Baseline	Development Time (months)
Baseline	*$559.6*	*—*	*$1,240.9*	*—*	*133.7*
Individual TFA Scenarios					
Bioimaging	—	—	—	—	—
Biomarkers	$347.9	–38%	$676.9	–45%	108.2
Bioinformatics	$375.0	–33%	$746.3	–40%	116.6
Gene Expression	$345.8	–38%	$676.0	–45%	111.9
Combined Scenarios					
Lower bound	$421.2	–25%	$869.6	–30%	122.4
Upper bound	$289.2	–48%	$533.1	–57%	98.1

Note: See Table 5-4 for estimated changes in FDA approval, distribution of IND failures within clinical trials, and probability of a recall of an FDA-approved drug.

Table 6-2 highlights key technology infrastructure priorities that emerged during interviews and that informed the scenarios against which potential efficiency gains were quantified. Most notable among participants' comments were (1) the need to establish a basic foundation of standardization within and across each TFA and (2) that gains in one TFA spill over and increase gains from others. The process of aligning research strategies and investments in the following improvements to the technology infrastructure would be best approached in a coordinated, interdisciplinary manner.

6.2.1 Advanced Bioimaging Techniques

Bioimaging studies are becoming more important in the analysis of how chemical compounds impact biological systems, particularly as FDA increasingly requests that such studies accompany submissions and Phase IV clinical trials. While these studies have the ability to capture and analyze fluorescently coded information to seek out biomarkers, improvements in measurement, calibration, and metrology systems are needed to improve comparability, repeatability, and throughput.

Table 6-2. Stakeholders' Comments on Technology Infrastructure Needs

Technology Focus Area	Within the TFAs studied in this report, company representatives, academics, and government researchers recommended that needed improvements to the technology infrastructure:
Bioimaging	• Include consistent taxonomies for medical and anatomical regions of observations • Standardize image labeling procedures and ontologies • Develop formats for exchanging imaging data among data systems • Improve the image archival, retrieval, and management infrastructure • Improve access to imaging technology, including image capture and interpretation systems
Biomarkers	• Address the need for greater sensitivity in detection of protein expression levels • Develop traceable standards for currently known immunoassayed biomarkers • Standardize existing protocols for generating gene expression results • Develop standardized methods and tools to hasten validation of technology platforms • Standardize statistical methodologies for data analysis in biomarker validation studies
Bioinformatics	• Improve data visualization and analysis techniques, • Develop common (neutral) data formats and analysis tools • Set standards for and transforming and exchanging data • Standardize ontologies for characterizing data
Gene Expression	• Make available reference materials that mimic the biological complexity of tissue and blood samples • Establish sample quality standards, including tools to evaluate the extent to which samples may have degraded • Create sample acquisition, handling, and preparation techniques given the amount of time samples may spend in transit between research sites • Establish systems, data, and analysis mechanisms to benchmark microarray performance • Develop calibration tools and techniques for scanning equipment • Provide standard calibration curves for genes as well as standard control techniques, assays, protocols, and investigative algorithms
Commercial Manufacturing	• Develop standardized data formats for production equipment and instrumentation • Create on-line measurement methodologies to improve process understanding and establish industry standard QA/QC measures • Improve inspection and validation methodologies • Develop reference standards analogous to cellular material for future production cell and gene therapies
Postmarket Surveillance	• Standardize protocols and descriptions for adverse event data to attain greater efficiency in ongoing safety and efficacy monitoring and FDA reporting • Standardize the syntax and interchange between clinical safety databases • Improved statistical methodologies to enable multivariate analysis of safety data • Develop uniform standards for data formats for clinical records

Stakeholders cited the significant benefits of using bioimaging techniques to seek and quantify key biomarkers in large-scale imaging studies. However, identification and quantification of key biomarkers for a disease first requires standard protocols for patient or specimen positioning, instrument calibration, and settings to reduce variability and allow images to be compared across tests (in discovery) or trials (in the clinical testing phases). Given the volume of imaging data that would be analyzed in high-throughput imaging studies, experts also cited the need for automated or computer-assisted interpretation of images, including edge detection and size measurements.

6.2.2 Molecular Biomarkers

New biomarkers such as those that predict toxicity or detect concentrations of antibodies in cells are needed if the potential time and cost benefits quantified as part of this study are to be realized. The industry experts we interviewed stressed the benefits that using validated efficacy biomarkers as surrogate end points in clinical trials could have, particularly the ability to rank and prioritize drug candidates in a firm's portfolio of potential products. Enhancements to researchers' ability to predict biological response to treatment provides researchers with access to information that previously was only available at later stages in the discovery process or clinical trials. Outside of drug R&D, next-generation biomarkers hold promise for patients whose quality of care may improve because doctors have improved foresight into the transition from wellness to disease.

Key improvements in the technology infrastructure to support identification of and widespread use of biomarkers include a standardized and accepted validation process for biomarker discovery. Sophisticated measurement tools and methods are needed as part of providing accurate data if new biomarkers are to assist with ranking drug targets and guide decision-making.

6.2.3 Bioinformatics and *In Silico* Predictive Modeling

Improving the effectiveness and efficiency of bioinformatics as drug discovery shifts away from traditional laboratory science toward computational systems is one of the key challenges facing the industry. Tools and strategies for interpreting, synthesizing, and communicating data have not kept pace with the ability to generate it. Standardization for bioinformatics is essential if the full potential of advances in gene expression, bioimaging, and biomarkers' improvements is to be realized.

Achieving key priorities within informatics would increase the confidence and assurance analysts have in the data being used in analyses. These include developing

- tools for integrating data sets between technology platforms and managing data volumes;

- quality and accuracy measurements and standards for data validation, cleaning, and normalization;

- standardized ontologies for describing and formatting data elements;

- greater availability of currently gathered data through publicly accessible and maintained databases;

- improved data mining applications and algorithms; and

- improved accuracy of *in silico* model predictions, including models of protein structure and binding, cellular localization, PK/PD properties, clinical trial simulation, and disease modeling.

6.2.4 Gene and Protein Expression Analysis

Standardization in gene and protein expression analysis is needed for drug R&D and the practice of medicine to achieve NIH's and FDA's goals of personalized medicine. Expression analysis systems capture and analyze information from organisms' cells that can be used in medical research and health management. It will likely be a decade or more until medical care reaches the point of personalization, but if the underlying technology infrastructure supporting expression analysis remains fragmented and characterized by dissimilar approaches to diagnostics and assays, then availability of personalized medicine will be still further delayed.

The development of key measurements, protocols, standards, and tools underlying the benefits estimates presented in Table 6-1 include

- publicly available synthetic mRNA reference materials for microarray performance assurance,

- standard technical protocols for microarray experiments,

- standard protocols for RNA and DNA extraction,

- data and analysis to benchmark microarray performance,

- microarray scanning equipment calibration tools, and

- techniques for measuring the presence of low-abundance proteins.

6.2.5 Commercial Manufacturing

Improvements in data integration and in-line process measurement techniques are needed in commercial manufacturing. Biopharmaceutical manufacturing is a complex process with a propensity for high levels of variability in reproducing biological substances. Standardization of impurity testing methods and protein aggregation studies will ensure safer and more effective products. Over the next 10 to 20 years, manufacturers will continue to achieve major milestones in improving production yields, demonstrating a better understanding of the bioprocess underlying their products. However, the benefits to consumers resulting from these innovations could be constrained by continued inefficiencies if parallel improvements to the measurement and technical infrastructure that supports the commercial manufacturing process are not made. Particularly these include

- standardized data formats for production equipment and instrumentation,
- creation of on-line measurement methodologies to improve process understanding and establish industry standard QA/QC measures,
- improve inspection and validation methodologies related to immunogenicity and aggregation of proteins during manufacturing and storage, and
- develop reference standards analogous to cellular material for future production cell and gene therapies.

6.2.6 Postmarket Surveillance

Improved standardization of data elements, improved data integration, and more powerful statistical methodologies are needed to improve the efficiency and precision of postmarket surveillance activities. Over the next decade, the scope and quantity of data required for postmarket surveillance is going to increase dramatically largely because of globalization of markets and increasing FDA requirements for long-term safety studies. Robust and efficient technology infrastructure in data collection, retrieval, and analysis methods will be required to ensure that these postmarket studies can be cost effectively implemented. Stakeholders' priorities include

- standardizing protocols and descriptions for adverse event data to attain greater efficiency in ongoing safety and efficacy monitoring and FDA reporting,
- standardizing the syntax and interchange between clinical safety databases,

- improving statistical methodologies to enable multivariate analysis of safety data, and

- developing uniform standards for data formats for clinical records.

6.3 RATIONALE FOR NIST INVOLVEMENT

Interviewees and survey respondents stated that NIST participation in standardization activities for biopharmaceuticals would be welcome. About half of all interviewees were familiar with NIST and cited NIST's historical leadership in standards-setting and coordination. These experts believe that NIST has a natural role, given its status as an independent, neutral body, its greater access to consistent funding, and its mission to provide technical infrastructure to industry.

However, senior scientists and directors at biopharmaceutical companies cited FDA's regulatory authority as a potential constraint and suggested that NIST and FDA collaborate to identify and develop standards that are congruent with the Critical Path initiative. Companies' reporting requirements to FDA add a regulatory driver to nearly all research initiatives. These individuals encouraged NIST leadership and participation in standards-setting but stated that any initiatives should be mindful of and improve companies' ability to report to FDA, as well as improving data communication internally and between organizations.

Although the entire biopharmaceutical industry would benefit from an improved infrastructure, emerging companies that have yet to adopt or develop an internal infrastructure stand to gain the most. Such adoptions by small start-up firms may increase their chances for success and thereby invigorate entrepreneurship.

The broader biopharmaceutical and biotechnology industry would benefit from greater efficiency and effectiveness with a nationally coordinated standardization effort supported by an independent research organization with a proven track record, technical expertise, and access to financial and technical resources. The ultimate beneficiaries are patients who would have access to a broader array of novel therapies whose development was supported by an effective technology infrastructure.

References

Adams, Christopher P., and Van V. Brantner. 2006. "Estimating the Cost of New Drug Development: Is It Really $802 Million?" *Health Affairs* 25(2):421-428.

Aldridge, Susan. 2006. "Bioprocessing: Downstream Processing Needs a Boost." *Genetic Engineering News* 26(1).

Anand, Geeta. 2005. "The Most Expensive Drugs—Rx for an Industry." *Wall Street Journal* December 28:A1.

Battelle Technology Partnership Practice and SSTI. 2006. "Growing the Nation's Bioscience Sector: State Bioscience Initiatives 2006." Available at http://www.bio.org/local/battelle2006/ battelle2006.pdf. As obtained August 2006.

Biotechnology Industry Association (BIO). 2004. "Guide to Biotechnology: Biotechnology Industry Facts." Available at http://www.bio.org/speeches/pubs/er/statistics.asp. Updated June 2004.

Biotechnology Industry Organization (BIO). 2006. "Approved Biotechnology Drugs." Available at: http://www.bio.org/ speeches/pubs/er/approveddrugs.asp. As obtained August 2006.

Burrill and Company. 2006. "Biotech 2006—Life Sciences: A Changing Prescription." Available at http://www.burrillandco.com/bio/ biotech_book.

Chapman, Carolyn Riley. 2005. "In Vivo Molecular Imaging Poised for Growth." *Genetic Engineering News* 25(9):22-23.

Crommelin, Daan J.A., and Robert D. Sindelar (eds.). 2002. *Pharmaceutical Biotechnology.* 2nd Ed. New York: Taylor & Francis Inc.

DiMasi, Joseph A. 2002. "The Value of Improving the Productivity of the Drug Development Process: Faster Times and Better Decisions." *Pharmacoeconomics* 20(3):1-10.

DiMasi, Joseph A., Ronald W. Hansen, Henry G. Grabowski, and Louis Lasagna. 1991. "The Cost of Innovation in the Pharmaceutical Industry." *Journal of Health Economics* 10:107-142.

DiMasi, Joseph A., Ronald W. Hansen, and Henry G. Grabowski. 2003. "The Price of Innovation: New Estimates of Drug Development Costs." *Journal of Health Economics* 22:151-185.

DiMasi, Joseph A., and Henry G. Grabowski. 2007. "The Cost of Biopharmaceutical R&D: Is Biotech Different?" *Managerial and Decision Economics* 28(4-5):469-479.

Frost and Sullivan. 2004. "Strategic Analysis of Downstream Processing in Biopharmaceutical Production." Frost & Sullivan Research Service. Report #F123-52. San Antonio, TX.

Gerhold, D., R. Jensen, and S. Gullans. 2002. "Better Therapeutics through Microarrays." *Nature Genetics* 32(Suppl):547-551

Giersiefen, Helmut, Rolf Hilgenfeld, and Alexander Hillisch. 2003. "Modern Methods of Drug Discovery: An Introduction." In *Modern Methods of Drug Discovery.* Alexander Hillisch and Rolf Hilgenfeld (eds.), pp. 1-18. Verlag, Switzerland: Firkhauser.

Ginsburg, Geoffrey, and Susanne B. Haga. 2006. "Translating genomic biomarkers into clinically useful diagnostics." *Expert Review of Molecular Diagnostics.* 6(2):179-191.

Global Alliance for TB Drug Development. 2001. "The Economics of TB Drug Development." Nancy Pekar (ed.). Available at http://www.tballiance.org/3_2_C_BalancingIncentivesand Access.asp. As obtained on January 10, 2006.

Ho, Rodney J.Y., and Milo Gibaldi. 2003. *Biotechnology and Biopharmaceuticals.* Hoboken, NY: John Wiley & Sons, Inc.

Keating, Sean. 2003. "Data Mining Improves Drugs from Beginning to End." *Discovery & Development* 6(7):53-54.

King, Simon. 2006. "The Evolving Pharmaceutical Value Chain: Forecasting Growth for Small and Large Molecules. Pharmaceutical Technology." Available at http://www.pharmtech.com/pharmtech/article/articleDetail.jsp?id= 385865.

Malik, T. 2003. "Microarrays and the Scanners that Read Them." *Drug Discovery and Development* 6(3):39-44.

Mathieu, Mark. 2005. *PAREXEL's Pharmaceutical R&D Statistical Sourcebook 2005/2006.* Waltham, MA: PAREXEL.

McGee, Patrick. 2005b. "New In Vitro, Modeling Tools May Cut Tox Attrition." *Drug Discovery & Development* 8(8):42-49.

McGee, Patrick. 2005c. "Rethinking Library Synthesis to Do More with Less." *Drug Discovery & Development.* Rockaway, NJ: Reed Business Information. Available at www.ddmag.com. As obtained on March 2, 2006.

McGee, Patrick. 2005d. "Safety Biomarkers Help to Reduce Drug Attrition." *Drug Discovery & Development* 8(9):44-46.

National Institutes of Health (NIH). 2006. "NIH Roadmap by Initiative." Excel spreadsheet sent via e-mail.

O'Connor, A. C., B. Rowe, M. Gallaher., J. Sevinsky., and D. Wood. 2007. "Economic Impact of ATP's Contributions to DNA Diagnostics Technologies." GCR 06-898. Prepared for the Advanced Technology Program, National Institute for Standards & Technology, Gaithersburg, Maryland.

Pharmaceutical Research and Manufacturers of America (PhRMA) March 2005. *Pharmaceutical Industry Profile 2005.* Washington, DC: PhRMA.

Pavlou, Alex K. and Janice M. Reichert. 2004. "Recombinant Protein Therapeutics—Success Rates, Market Trends and Values to 2010." *Nature Biotechnology* 22(12):1513-1519.

Pharmaceutical Research and Manufacturers of America (PhRMA). March 2005. *Pharmaceutical Industry Profile 2005.* Washington, DC: PhRMA.

Pharmaceutical Research and Manufacturers of America (PhRMA). 2006a. "2006 Survey: Medicines in Development: Biotechnology" Available at http://www.phrma.org/files/Biotech%202006.pdf. As obtained August 2006.

Pharmaceutical Research and Manufacturers of America (PhRMA). 2006b. "Pharmaceutical Industry Profile 2005." Washington, DC: PhRMA, March 2006. Available at http://www.phrma.org/files/ 2006%20Industry%20Profile.pdf.

Quackenbush, John. 2004. "Data Standards for 'Omic' Science." *Nature Biotechnology* 22:613–614.

Rader, Ronald A. 2005. "What is a Biopharmaceutical? Part 1: (Bio)Technology-Based Definitions." *Bio-Executive International.* March 2005 (60-65). Available online at: http://www.biopharma .com/BioExec_pt1.pdf

Reichert, Janice M. 2004. "New Biotechnology Therapeutics: Challenges for the Industry." *BIO-Greece.* 10:38-40. Available at http://csdd.tufts.edu/_documents/www/Doc_233_7751_826.pdf. As obtained on March 21, 2005.

Reichert, Janice and Viia Valge-Archer. 2007. "Development Trends in Monoclonal Antibody Cancer Therapeutics." *Nature Reviews Drug Discovery* 6(5):349-356.

Ridley, David B., Judith M. Kramer, Hugh H. Tilson, Henry G. Grabowski, and Kevin A. Schulman. 2006. "Spending on Postapproval Drug Safety." *Health Affairs* 25(2):429-436.

Tassey, Gregory. 1997. The Economics of R&D Policy. Westport, CT: Quorum Books.

Tollman, Peter, Philipe Guy, Jill Altshuler, Alastair Flanagan, and Michael Steiner. 2001. "A Revolution in R&D: How Genomics and Genetics are Transforming the Biopharmaceutical Industry." Boston: The Boston Consulting Group. Available at http://www.bcg.com/publications/files/eng_genomicsgenetics_rep_11_01.pdf. As obtained on January 12, 2006.

Toole, A.A. 2003. "Understanding Entrepreneurship in the U.S. Biotechnology Industry." In *The Emergence of Entrepreneurship Policy: Governance, Start-Ups, and Growth in the U.S. Knowledge Economy,* D.M. Hart (ed.). Cambridge: Cambridge University Press.

Tufts Center for the Study of Drug Development. March 7, 2005. "Biotechnology Advances Have Improved R&D Success Rates." Available at http://csdd.tufts.edu/NewsEvents/NewsArticle.asp?newsid=52.

U.S. Department of Commerce (DOC). 2003. "A Survey of the Use of Biotechnology in U.S. Industry." Available at http://www.technology.gov/reports/Biotechnology/CD120a_0310.pdf. Obtained August 2006.

U.S. Food and Drug Administration (FDA). 2004a. "Guidance for Industry: PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance." Pharmaceutical cGMPs. Available at http://www.fda.gov/cder/guidance/6419fnl.pdf.

U.S. Food and Drug Administration (FDA). 2004b. "Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products." FDA White Paper. Washington, DC: FDA.

U.S. Food and Drug Administration (FDA). 2006. "Critical Path to Personalized Medicine." Office of Management and Budget Formulation and Presentation. Available at: http://www.fda.gov/oc/oms/ofm/budget/2007/HTML/4CPPOM1.htm.

Walsh, Gary. 2003. *Biopharmaceuticals: Biochemistry and Biotechnology, Second Edition.* West Sussex, England: John Wiley & Sons Ltd.

Wang, Y.-X. J. 2005. "Medical Imaging in Pharmaceutical Clinical Trials: What Radiologists Should Know." *Clinical Radiology* 60:1051-1057.

Woodcock, Janet. 2005. "The Critical Path Initiative: One Year Later." Presentation at the FDA/DIA/BIO Meeting: Use of Medical Imaging as a Drug Development Tool.

Appendix A: Survey

Electronic Survey of the Biopharmaceutical Industry

As part of an economic study for the National Institute of Standards and Technology (NIST), RTI International (RTI) is conducting a survey of biotechnology-related companies involved in research, development, testing, and manufacturing of the next generation in human therapeutics known as biopharmaceuticals.

Currently, NIST devotes $426.3 million to scientific and technical research services in support of the nation's technology infrastructure. The technology infrastructure consists of a set of "technical tools and processes" that enable or increase the efficiency of R&D, manufacturing, and market penetration (e.g., measurement science, standards, and other supporting technologies). NIST has designated biotechnology as a strategic focus area and is planning to expand its research programs to provide the U.S. biotechnology industry with infrastructure technologies and standards that will increase the efficiency of R&D, manufacturing, and market transactions.

You can assist NIST in its strategic planning by providing estimates of in-house and external spending related to the technology infrastructure supporting the biopharmaceutical industry. These estimates will be used to develop a national estimate of private-sector expenditures related to technology infrastructure in the biopharmaceutical industry.

Your participation in this critically important survey is voluntary. However, our study would greatly benefit from your insights and experience, and, in exchange for your participation, you will be given access to the final report. Please note that your responses will be kept strictly **confidential** and will not be available to the public or shared with NIST or other survey participants. All survey responses to questions regarding current expenditures will be aggregated to obtain an estimate of technology infrastructure costs for the entire industry. Only these aggregate numbers will be reported in the final report at the conclusion of this research study.

Thank you for your willingness to participate. The survey should take approximately 20 to 30 minutes to complete, and, if you have any questions, please contact Jeffrey Petrusa at NIST_biotech@rti.org.

Please click here to access the survey or go to https://biotechsurvey.rti.org/.

NIST/RTI Study:
The Technology Infrastructure Needs of the Biopharmaceutical Industry

Thank you for your participation in this brief but critically important survey regarding the technology infrastructure supporting the research, development, production, and marketing of biopharmaceuticals. The results of this survey will be used by RTI International (RTI) as part of a strategic planning study commissioned by the National Institute of Standards and Technology (NIST).

The purpose of this survey is to collect information on the current level of private-sector expenditures on the standards, measurement techniques, and technical tools that make up the technology infrastructure. In addition, the information will be used to assess the existing needs and potential impact of improvements in this infrastructure.

Instructions:
This survey should take approximately 20 to 30 minutes to complete. Your participation is voluntary, and your responses will be kept strictly confidential.[30] Please answer all questions by checking the appropriate box(es) or providing text in the designated space. You do not need to look up any information; simply provide answers based on your best knowledge.

In exchange for your participation, you will have access to a copy of the final report. If you would like to be notified when this report is available, please include your e-mail address and contact information at the end of this survey.

If you have any questions as you complete the survey, please contact Jeffrey Petrusa at NIST_biotech@rti.org.

---

[30] **Nondisclosure policy:**

RTI has well-established practices for handling confidential information as part of numerous projects. Any information we obtain through these surveys will be used solely in aggregate with information from other respondents. In no instance will specific individuals or organizations be identified by name in any reports or as part of information that is released publicly or to NIST.

Part I: Background Information

1. Please provide your professional title: _____ (e.g., director, production supervisor, project manager, head of process engineering)

2. Please indicate whether your parent company's primary or secondary use of biotechnology is for human-health applications (e.g., therapeutics, preventatives, or diagnostics).

   ☐ Yes

   ☐ No

   If they check NO, the survey will terminate with a "Thank you, but you do not qualify for this study" message.

3. What stage of the biopharmaceutical drug development and manufacturing supply chain are you most involved with?

   NOTE: If you work for a company that is diversified, please select the activity about which you personally are most knowledgeable and complete this survey from that perspective. Skip logic for the remainder of the survey will be guided by their response to this question. Based on their selection respondents will be asked to consider questions in 1 of 2 areas: product development, and commercial activities.

   Product Development-Related Activities *(Pre-FDA Approval)*

   ☐ Drug Discovery and Preclinical Development

   ☐ Clinical Trials (includes any activities prior to approval from FDA)

   Commercial Activities *(Post-FDA approval)*

   ☐ Commercial-Scale Manufacturing

   ☐ Post-market Monitoring

4. Approximately how many full-time equivalent (FTE) scientists and engineers are active in <<Activity from Q5>>? << Pop-up definitions:

   FTE: A measurement equal to one staff person working a **full-time** work schedule (35-40 hours per week) for one year.

   Scientists and engineers: Includes scientists, engineers, science and clinical laboratory technicians, and R&D focused computer specialists. >>

   _____ = Total number of scientists and engineers (FTE)

   IF the respondent selects Discovery/Preclinical Development or Clinical Trial Activities in Q5 then the following statement will be given.

   For the remainder of the survey please respond only for the activities related to the <response to Q4 above>scientists and engineers you indicated above.

   THEN GO TO PART II-A

Manufacturing Activities

5. How many biological product lines are currently produced by you group or division?

Product Type	Number of Products in the Development Pipeline	Number of FDA-Approved Products
Therapeutic protein		
Monoclonal antibody		
Vaccine		
Gene therapy		
Cellular therapy		
Tissue therapy		
Other biotechnology based human health applications		
Total	<sum of above #s>	<sum of above #s>

Post-market Monitoring Activities

6. How many biological product lines are currently monitored by your group/division?

Product Type	Number of FDA-Approved Products
Therapeutic proteins	
Monoclonal antibodies	
Vaccines	
Gene therapies	
Cellular therapies	
Tissue therapies	
Other biotechnology based human health applications	
Total	<sum of above #s>

For the remainder of the survey please respond only for the activities related to the <total in table above> product lines currently marketed by your group/division.

Part II-A: Current Expenditures on Technology Infrastructure

The purpose of this section is to obtain estimates from your organization about the resources allocated to developing comparable test methods, standardized software models, standard reference materials used for measurement and calibration, and standard reference data. We will aggregate this information to quantify the proportion of the biopharmaceutical industry's activities that are allocated to improving the "technology infrastructure" supporting product development (R&D).

Technology infrastructure consists of a set of "technical tools and processes" that enable or increase the efficiency of R&D, manufacturing, and market penetration. Collectively, NIST refers to these tools and processes as **the technology infrastructure**. Technology infrastructure provides the technical basis for activities ranging from drug discovery to quality control in the production process. Investments in the technology infrastructure have the potential to significantly lower development, production, and transaction costs and hence are an ongoing activity in most companies. Investment in the technology infrastructure may include both direct and indirect costs in the form of equipment, labor, and material expenditures dedicated to technology infrastructure.

Examples of technology infrastructure include

- certified reference materials, reference methods, and standardized procedures related to measurement or calibration in conducting tests during R&D stages or in any manufacturing process;

- validated methods for interpreting results from different analytical platforms;

- standardized techniques (e.g., for use with spectrometers, imaging systems, and/or measurement devices);

- standardized methods for characterizing scientific and engineering data, and the algorithms to manipulate, search, and analyze this uniform data within a publicly available database; and

- processing techniques (e.g., for use with 2D gel electrophoresis platforms, chromatography separation systems)

Click here for examples of **technology infrastructure costs.**

The following bullets will appear in as a separate browser window if the respondent selects the link in RED above.

Technology infrastructure Cost Examples

- Purchase of measurement-related equipment, software, reference materials, or services.

- Activities required for setup and validation of analytical instruments, reagents, or other research tools. Examples of these activities may include developing test methodologies or process standards in measurement and manufacturing practices.

- In-house customization of technology platforms purchased from third-party vendors.

- Efforts to develop interoperability between different software or equipment systems.

- License fees or any other spending on enhancements to routinely used processes or equipment, intended to increase productivity, reduce redundancy, or improve the confidence in results.

1. Based on the definition and examples of the technology infrastructure provided above, please estimate the current annual level of expenditures on the technology infrastructure supporting your (Fill in from answer to Q5 of Part I) activities. *Please consider all labor, materials, and equipment expenditures for in-house activities, as well as contributions to consortia or partnership initiatives.*

Approximate Annual Expenditure on Technology Infrastructure =

$_____

Technology infrastructure Expenditure Category	Approximate Percentage of Annual Expenditure on Technology infrastructure
Labor (includes all research, implementation support activities, and participation in consortia related to infrastructural technology)	<<value 0 to 100>>
Capital (includes equipment, software, and licensing fees related to infrastructural technology)	<<value 0 to 100>>
Materials (includes all reference materials and reagents and other materials related to infrastructural technology)	<<value 0 to 100>>
Total	100%

Insert "Save & Continue" button after this question. Following Question will appear on a new screen.

2. Approximately what share of total R&D expenditures <<$ value from Q1 above>> in Question 1 represent?

_____ % of total current annual R&D expenditures

SKIP LOGIC: IF Respondent answers $0 or 0% THEN the survey will terminate with a "Thank you for your participation in this important NIST study" message.

3. a. Over the last 5 years, have these expenditures increased or decreased or stayed the same?

☐ **Increased**
☐ **Decreased**
☐ **Stayed the same**

3b. Comments: _____

4. The four technologies listed below have been identified as critical enabling technologies for developing biopharmaceuticals. Please allocate a percentage of your estimated technology infrastructure expenditure of $(Fill in response to Q1 of Part II) to the following enabling technologies. *If you do not use a particular technology, please enter 0%.*

Enabling Technology (not listed in any particular order)	Approximate Percentage of Annual Expenditure on Technology infrastructure
Bioimaging (PET, MRI, mass spec, other imaging modality)	
Informatics (software and databases)	
Molecular biomarkers	
Gene and/or protein expression platforms	
Other	
Total Annual Technology infrastructure Expenditure	100%

If they enter a positive amount for **"Other"** in the box above THEN ask:

4a. What are some examples of the "Other" enabling technologies not listed above?

_____
_____

If they allocate 100 to "Other" category, the survey will terminate with a "Thank you for your participation in this important NIST study" message.

If they allocate a percent greater than zero to bioimaging, they will answer this question.

5. **Bioimaging**

Which of the following imaging modalities does your group support in biotechnology-based drug discovery and development? Please check all that apply.

☐ X-Ray (for research activities other than diagnosis)

☐ Computed tomography (CT)

☐ Magnetic resonance imaging (MRI)

☐ Ultrasound

☐ Single-photon emission computed tomography (SPECT)

☐ Positron emission tomography (PET)

☐ Optical imaging, including fluorescence

☐ Combined methods (such as PET-CT)

Please specify: _____

☐ Other

Please specify: _____

If they allocate a percent greater than zero to informatics, they will answer this question.

6. **Informatics**

Which of the following informatic tools does your organization support in drug discovery and development? Please check all that apply.

- ☐ Bioinformatic databases
- ☐ Cheminformatic databases
- ☐ Systems biology and pathway analysis
- ☐ Data management software
- ☐ Predictive modeling (also referred as *in silico* modeling) software

If they allocate a percent greater than zero to molecular biomarkers, they will answer this question.

7. **Molecular Biomarkers**

Which of the following types of biomarkers does your organization currently support in drug research and development? Please check all that apply.

- ☐ Translation biomarkers
- ☐ Disease biomarkers
- ☐ Efficacy biomarkers
- ☐ Staging biomarkers
- ☐ Surrogate biomarkers
- ☐ Toxicity biomarkers
- ☐ Mechanism biomarkers
- ☐ Target biomarkers

If they allocate a percent greater than zero to Gene/protein expression platforms, they will answer this question.

8. **Gene and/or Protein Expression Platforms**

Which of the following platforms does your organization support in measuring gene and/or protein expression for drug discovery and development? Please check all that apply.

Gene Expression
- ☐ Quantitative, real-time reverse transcriptase polymerase chain reaction (QRT-PCR) assays
- ☐ Microarray experiments
- ☐ Other, please specify: _____

Protein Expression
- ☐ 2D gel electrophoresis
- ☐ Mass spectrometry
- ☐ Protein arrays
- ☐ Other, please specify: _____

Part III-A: Assessment of the Potential Impact of Improvements in Technology Infrastructure

The purpose of this section is to collect information that will help us assess the potential impact that certain improvements to the underlying technology infrastructure could have on the current cost of drug discovery and development.

1. Please indicate which of these four technologies you are MOST knowledgeable about. Select only one of the four boxes.

   ☐   Bioimaging

   ☐   Informatics

   ☐   Molecular biomarkers

   ☐   Gene and/or protein expression analysis

SKIP LOGIC: THE RESPONDENT WILL BE DIRECTED TO THE TECHNOLOGY THAT CORRESPONDS TO THE TECHNOLOGY THEY SELECTED ABOVE in Q1 of Part III.
For Bioimaging, see pages 10-13.
For Informatics, see pages 14-17.
For Molecular Biomarkers, see pages 18-20.
For Gene and/or protein expression analysis, see pages 21-24.

Each Respondent (answering Part III-A) will only answer one of the four Technology Cases listed above.

Assessment of the Potential Impact of Improvements in Bioimaging

In this case, we investigate the potential R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of bioimaging in biopharmaceutical discovery, development, and testing. The advancements listed come from the NIH Technology Road Map, the FDA Critical Path Opportunities List, and other industry sources.

The following six advancements exist in varying degrees today and are expected to improve in the coming years:

B. Qualified imaging biomarkers in your disease or therapeutic category of interest.

C. Availability of labeling and contrast agents that are proven to be stable *in vivo*, nondestructive to the molecule being tested, and nontoxic in humans.

D. Reconstruction algorithms that improve image quality (e.g., reduce fuzziness, define clear borders).

E. Automated or computer-assisted interpretation of images, including edge detection and size measurements.

F. National image test bed and benchmarking methodology to assess the quality of the image analysis algorithms and software.

G. Standard protocols for patient/specimen positioning, instrument calibration, and settings to reduce variability and allow images to be compared across tests (in discovery) or trials (in the clinical testing phases).

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include:

- R&D labor and materials cost savings associated with shorter clinical trials or smaller trial populations.

- Elimination of alternative tests required to identify and validate a target or response.

- Reduction in the number of animal subjects needed (since noninvasive imaging would allow the same animal to be followed for longer periods of time).

- Reduction in the number and skill level of technicians required to verify the analysis of an image.

- Elimination of additional (redundant) tests needed to verify results, due to poor image quality.

- Reduction in R&D labor costs for time spent assessing performance capabilities of image analysis software.

- Reduction in R&D labor costs associated with calibrating equipment due to the availability of standardized procedures and reference phantoms.

Impact on Research and Development

1.   Consider the following "improved world." Suppose there is an immediate advancement in the above areas of 50% over the current state of the art. What would you estimate to be the impact on the process of bringing a candidate biopharmaceutical to the next stage of development? *Please consider impacts over the entire development process.*

Stage	Percentage Reduction in Stage Length (Time until Start of Next Stage)	Percentage Reduction in Stage Cost	Check Here if Not Familiar with this Stage
Discovery & Preclinical	____%	____%	☐
Phase I	____%	____%	☐
Phase II	____%	____%	☐
Phase III	____%	____%	☐

Comments:_____-
_____

Impact on Success Rates

Now consider the success and failure rates for candidate biopharmaceuticals in this same "improved world" (a 50% advancement in the current the state of the art in the above areas).

2a. The table below provides the average success rate for an Investigational New Drug (IND) entering clinical trials, in other words, the probability that an IND will receive FDA approval. Please estimate what impact the advancements above might have on this success rate in the table below.

For example, providing an estimate that is greater than 35% implies that the advancements above allow researchers to make more well-informed decisions in selecting a lead candidate for IND submission. A higher success rate implies improved quality of all INDs; thus, there is a higher probability of receiving FDA approval.

Change in IND Overall Success Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World
Probability that an IND ultimately receives FDA approval	35%		

2b. From the table above, a 35% success rate implies that 35 out of 100 INDs will receive FDA approval. This suggests that the other 65 INDs will fail at some point during clinical trials. The table below provides the average distribution of failures occurring at each clinical trial phase. Please estimate what impact the advancements above might have on this distribution of failures across the clinical trial phases.

For example, if you think the advancements above would allow researchers to make decisions about project failures earlier in clinical trials, this impact would be represented by shifting 5% of clinical failures from Phase III to Phase I. A shift in the occurrence of failures to an earlier phase implies a decrease in the overall cost of drug development.

Change in Clinical Trial Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Of all the INDs that **fail** in clinical trials, **what percentage fail** in each stage?				
Percentage failing in Phase I	15%			☐
Percentage failing in Phase II	46%			☐
Percentage failing in Phase III	39%			☐
Total	**100%**		100%	

2c. The table below provides the average failure rate for biopharmaceuticals that are currently approved and marketed to consumers. Please estimate the impact the advancements above might have on the probability that an approved drug is recalled.

Change in Post-market Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Probability that an approved drug is recalled	0.4%			☐

[a] Based on industry estimates and published literature.
[b] Respond only if your estimates are significantly different from typical percentages provided.

Comments:_____
_____
_____

3. Which of the advancements listed above do you perceive as being the most important to achieving the improvements you noted in the boxes above?

    ☐ Qualified imaging biomarkers in your disease or therapeutic category of interest.

    ☐ Availability of labeling and contrast agents that are proven to be stable *in vivo*, nondestructive to the molecule being tested, and nontoxic in humans.

    ☐ Reconstruction algorithms that improve image quality (e.g., reduce fuzziness, define clear borders).

    ☐ Automated or computer-assisted interpretation of images, including edge detection and size measurements.

    ☐ National image test bed and benchmarking methodology to assess the quality of the image analysis algorithms and software.

    ☐ Standard protocols for patient/specimen positioning, instrument calibration, and settings to reduce variability and allow images to be compared across tests (in discovery) or trials (in the clinical testing phases).

THIS IS THE END OF THE SURVEY for respondents who answered Bioimaging Technology Case.

Assessment of the Potential Impact of Improvements in Informatics

In this case, we investigate the potential R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of informatics in biopharmaceutical discovery, development, and testing. The advancements listed come from the NIH Technology Road Map, the FDA Critical Path Opportunities List, and other industry sources.

The following four advancements exist in varying degrees today and are expected to improve in the coming years:

A. Shared data standards for formatting and content. In addition to formatting standards for storing different types of data (such as image data), this advancement would include standards for the metadata that record information on the conditions of an experiment, similar to the content standards used as Minimum Information About a Microarray Experiment (MIAME).

B. Greater availability of currently gathered data through publicly accessible, curated databases. These databases would include (but not be limited to) data concerning the natural history of rare diseases; adverse events; toxicology properties of drug candidates; and absorption, distribution, metabolism, and elimination (ADME) properties of drug candidate.

C. Improved data mining applications and algorithms. Potential improvements include the incorporation of natural language processing into a text search.

D. Improved accuracy of *in silico* model predictions, including (but not limited to) models of protein structure and binding, cellular localization, pharmacokinetic/pharmacodynamic (PK/PD) properties, clinical trial simulation, and disease modeling.

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include the following:

- Reduction in time and labor costs spent on finding and interpreting data from previous experiments.

- Reduction in labor and material costs related to redundant experiments.

- Reduction in the number of animal subjects needed in preclinical trials (since *in silico* models can be used to predict how drug candidates will interact with the subject).

- R&D labor and material cost savings associated with fewer product redesign-synthesize-test cycles (since more accurate *in silico* models could be used to predict the effect of a proposed structural modification of a therapeutic product).

- Increased clinical trial success rate due to early detection of toxicity, using improved *in silico* models of information obtained from expanded adverse event databases, which eliminates drug candidates before entering clinical trials.

- Reduction in number of trials and patients through the simulation of clinical trials using *in silico* modeling.

Impact on Research and Development

1. Consider the following "improved world." Suppose there is an immediate advancement in the above areas of 50% over the current state of the art. What would you estimate to be the impact on the process of bringing a candidate biopharmaceutical to the next stage of development? *Please consider impacts over the entire development process.*

Stage	Percentage Reduction in Stage Length (Time until Start of Next Stage)	Percentage Reduction in Stage Cost	Check Here if Not Familiar with this Stage
Discovery & Preclinical	____%	____%	☐
Phase I	____%	____%	☐
Phase II	____%	____%	☐
Phase III	____%	____%	☐

Comments:_____
_____

Impact on Success Rates

Now consider the success and failure rates for candidate biopharmaceuticals in this same "improved world" (a 50% advancement in the current the state of the art in the above areas).

2a. The table below provides the average success rate for an Investigational New Drug (IND) entering clinical trials, in other words, the probability that an IND will receive FDA approval. Please estimate what impact the advancements above might have on this success rate in the table below.
For example, providing an estimate that is greater than 35% implies that the advancements above allow researchers to make more well-informed decisions in selecting a lead candidate for IND submission. A higher success rate implies improved quality of all INDs; thus, there is a higher probability of receiving FDA approval.

Change in IND Overall Success Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Probability that an IND ultimately receives FDA approval	35%			☐

2b. From the table above, a 35% success rate implies that 35 out of 100 INDs will receive FDA approval. This suggests that the other 65 INDs will fail at some point during clinical trials. The table below provides the average distribution of failures occurring at each clinical trial phase. Please estimate what impact the advancements above might have on this distribution of failures across the clinical trial phases.
For example, if you think the advancements above would allow researchers to make decisions about project failures earlier in clinical trials, this impact would be represented by shifting 5% of clinical failures from Phase III to Phase I. A shift in the occurrence of failures to an earlier phase implies a decrease in the overall cost of drug development.

Change in Clinical Trial Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Of all the INDs that **fail** in clinical trials, **what percentage fail** in each stage?				
Percentage failing in Phase I	15%			☐
Percentage failing in Phase II	46%			☐
Percentage failing in Phase III	39%			☐
Total	**100%**		100%	

2c. The table below provides the average failure rate for biopharmaceuticals that are currently approved and marketed to consumers. Please estimate the impact the advancements above might have on the probability that an approved drug is recalled.

Change in Post-market Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Probability that an approved drug is recalled	0.4%			☐

[a] Based on industry estimates and published literature.
[b] Respond only if your estimates are significantly different from typical percentages provided.

Comments:_____
_____

3. Which of the advancements listed above do you perceive as being the most important to achieving the improvements you noted in the boxes above?
   - ☐ Shared data standards for formatting and content.
   - ☐ Greater availability of currently gathered data through publicly accessible, curated databases.
   - ☐ Improved data mining applications and algorithms.
   - ☐ Improved accuracy of *in silico* model predictions.

THIS IS THE END OF THE SURVEY for respondents who answered Informatics Technology Case.

Assessment of the Potential Impact of Improvements in Molecular Biomarkers

In this section, we investigate the potential R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of biomarkers in biopharmaceutical discovery, development, and testing. The advancements listed come from the NIH Technology Road Map, the FDA Critical Path Opportunities List, and other industry sources.

The following four advancements exist in varying degrees today and are expected to improve in the coming years:

A. Standardized and accepted validation process for biomarker discovery.

B. Safety biomarkers that predict human toxicity in preclinical stages of drug development.

C. Validated efficacy biomarkers for diseases or therapeutic categories in your area of interest that could be used as surrogate endpoints in clinical trial activities.

D. Enhanced ability to predict biological response to treatment (e.g., drug efficacy measurement).

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include the following:

- Reduction in number of tests conducted in preclinical studies due to reduced uncertainty of drug efficacy.

- Avoidance of future R&D costs due to early detection of toxicity that eliminates drug candidates before entering clinical trials.

- Shorter clinical trial periods due to faster detection of biologic response to drug.

- Reduced number of patients in clinical studies due to more accurate selection criteria, resulting in improved stratified patient populations.

- Reduction in labor and material costs during clinical trials due to the ability to set dosage levels on a per-patient basis.

Impact on Research and Development

1. Consider the following "improved world." Suppose there is an immediate advancement in the above areas of 50% over the current state of the art. What would you estimate to be the impact on the process of bringing a candidate biopharmaceutical to the next stage of development? *Please consider impacts over the entire development process.*

Stage	Percentage Reduction in Stage Length (Time until Start of Next Stage)	Percentage Reduction in Stage Cost	Check Here if Not Familiar with this Stage
Discovery & Preclinical	____%	____%	☐
Phase I	____%	____%	☐
Phase II	____%	____%	☐
Phase III	____%	____%	☐

Comments:_____

_____

Impact on Success Rates

Now consider the success and failure rates for candidate biopharmaceuticals in this same "improved world" (a 50% advancement in the current the state of the art in the above areas).

2a. The table below provides the average success rate for an Investigational New Drug (IND) entering clinical trials, in other words, the probability that an IND will receive FDA approval. Please estimate what impact the advancements above might have on this success rate in the table below.
For example, providing an estimate that is greater than 35% implies that the advancements above allow researchers to make more well-informed decisions in selecting a lead candidate for IND submission. A higher success rate implies improved quality of all INDs; thus, there is a higher probability of receiving FDA approval.

Change in IND Overall Success Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Probability that an IND ultimately receives FDA approval	35%			☐

2b. From the table above, a 35% success rate implies that 35 out of 100 INDs will receive FDA approval. This suggests that the other 65 INDs will fail at some point during clinical trials. The table below provides the average distribution of failures occurring at each clinical trial phase. Please estimate what impact the advancements above might have on this distribution of failures across the clinical trial phases.
For example, if you think the advancements above would allow researchers to make decisions about project failures earlier in clinical trials, this impact would be represented by shifting 5% of clinical failures from Phase III to Phase I. A shift in the occurrence of failures to an earlier phase implies a decrease in the overall cost of drug development.

Change in Clinical Trial Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Of all the INDs that **fail** in clinical trials, **what percentage fail** in each stage?				
Percentage failing in Phase I	15%			☐
Percentage failing in Phase II	46%			☐
Percentage failing in Phase III	39%			☐
Total	**100%**		100%	

2c. The table below provides the average failure rate for biopharmaceuticals that are currently approved and marketed to consumers. Please estimate the impact the advancements above might have on the probability that an approved drug is recalled.

Change in Post-market Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Probability that an approved drug is recalled	0.4%			☐

[a] Based on industry estimates and published literature.
[b] Respond only if your estimates are significantly different from typical percentages provided.

Comments:_____

_____

3. Which of the advancements listed above do you perceive as being the most important to achieving the improvements you noted in the boxes above?

☐ Standardized and accepted validation process for biomarker discovery.

☐ Safety biomarkers that predict human toxicity in preclinical stages of drug development.

☐ Validated efficacy biomarkers for diseases or therapeutic categories in your area of interest that could be used as surrogate endpoints in clinical trial activities.

☐ Enhanced ability to predict biological response to treatment (e.g., drug efficacy measurement).

THIS IS THE END OF THE SURVEY for respondents who answered Molecular Biomarkers Technology Case.

A-20

Assessment of the Potential Impact of Improvements in Gene and/or Protein Expression Analysis

In this section, we investigate the potential R&D cost savings that could occur given hypothetical improvements in the technology infrastructure supporting the use of gene and/or protein expression analysis in biopharmaceutical discovery, development, and testing. The advancements listed come from the NIH Technology Road Map, the FDA Critical Path Opportunities List, and other industry sources.

The following nine advancements exist in varying degrees today and are expected to improve in the coming years:

For gene expression:
    A. Publicly available synthetic mRNA reference materials for microarray for performance assurance.

    B. Standard technical protocols for microarray experiments.

    C. Standard protocols for RNA and DNA extraction.

    D. Data and analysis to benchmark microarray performance.

    E. Microarray scanning equipment calibration tools.

For protein expression:
    F. Techniques for measuring the presence of low-abundance proteins.

    G. Improved sensitivity and lower coefficients of variation (≤10%) in mass spectrometry analysis to allow for its use in later stages of drug development.

    H. Improved availability of antibodies with high affinity, specificity, and selectivity for use with protein microarrays.

    I. Standards for protein microarray experiments.

Advancements in these areas could lead to a number of impacts on R&D productivity. Examples include the following:

- Reduction in labor, microarray, and consumables costs for redundant experiments and avoided downstream data capture and analysis costs for those experiments.

- Avoided downstream R&D costs for investigating genes and proteins mistakenly identified as potential targets.

- Greater confidence in and comparability among results from microarray experiments allowing for the elimination of redundant tests.

- Reduction in labor costs associated with calibrating equipment due to the availability of standardized procedures and reference phantoms.

- Reduction in or elimination of microarray scanning errors due to poor equipment alignment and calibration and associated downstream impacts of avoidable data capture errors.

- Elimination of additional (redundant) tests needed to verify results due to poor image quality.

Impact on Research and Development

1. Consider the following "improved world." Suppose there is an immediate advancement in the above areas of 50% over the current state of the art. What would you estimate to be the impact on the process of bringing a candidate biopharmaceutical to the next stage of development? *Please consider impacts over the entire development process.*

Stage	Percentage Reduction in Stage Length (Time until Start of Next Stage)	Percentage Reduction in Stage Cost	Check Here if Not Familiar with this Stage
Discovery & Preclinical	____%	____%	☐
Phase I	____%	____%	☐
Phase II	____%	____%	☐
Phase III	____%	____%	☐

Comments:_____
_____

Impact on Success Rates

Now consider the success and failure rates for candidate biopharmaceuticals in this same "improved world" (a 50% advancement in the current the state of the art in the above areas).

2a. The table below provides the average success rate for an Investigational New Drug (IND) entering clinical trials, in other words, the probability that an IND will receive FDA approval. Please estimate what impact the advancements above might have on this success rate in the table below.
For example, providing an estimate that is greater than 35% implies that the advancements above allow researchers to make more well-informed decisions in selecting a lead candidate for IND submission. A higher success rate implies improved quality of all INDs; thus, there is a higher probability of receiving FDA approval.

Change in IND Overall Success Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Probability that an IND ultimately receives FDA approval	35%			☐

2b. From the table above, a 35% success rate implies that 35 out of 100 INDs will receive FDA approval. This suggests that the other 65 INDs will fail at some point during clinical trials. The table below provides the average distribution of failures occurring at each clinical trial phase. Please estimate what impact the advancements above might have on this distribution of failures across the clinical trial phases.
For example, if you think the advancements above would allow researchers to make decisions about project failures earlier in clinical trials, this impact would be represented by shifting 5% of clinical failures from Phase III to Phase I. A shift in the occurrence of failures to an earlier phase implies a decrease in the overall cost of drug development.

Change in Clinical Trial Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Of all the INDs that **fail** in clinical trials, **what percentage fail** in each stage?				
Percentage failing in Phase I	15%			☐
Percentage failing in Phase II	46%			☐
Percentage failing in Phase III	39%			☐
Total	**100%**		100%	

2c. The table below provides the average failure rate for biopharmaceuticals that are currently approved and marketed to consumers. Please estimate the impact the advancements above might have on the probability that an approved drug is recalled.

Change in Post-market Failure Rates	Typical Rate[a]	Current Estimated Rate for Your Organization[b]	Estimated Rate in Improved World	Check Here if Not Familiar with This Stage
Probability that an approved drug is recalled	0.4%			☐

[a] Based on industry estimates and published literature.
[b] Respond only if your estimates are significantly different from typical percentages provided.

Comments:_____
_____

3. Which of the advancements listed above do you perceive as being the most important to achieving the improvements you noted in the boxes above?

☐ Publicly available synthetic mRNA reference materials for microarray for performance assurance.

☐ Standard technical protocols for microarray experiments.

☐ Standard protocols for RNA and DNA extraction.

☐ Data and analysis to benchmark microarray performance.

☐ Microarray scanning equipment calibration tools.

☐ Techniques for measuring the presence of low-abundance proteins.

☐ Improved sensitivity and lower coefficients of variation (≤10%) in mass spectrometry analysis to allow for its use in later stages of drug development.

☐ Improved availability of antibodies with high affinity, specificity, and selectivity for use with protein microarrays.

☐ Standards for protein microarray experiments.

THIS IS THE END OF THE SURVEY for respondents who answered Gene/Protein Expression Analysis Technology Case.

Part II-B: Current Expenditures on Technology Infrastructure

Manufacturing Activities

The purpose of this section is to obtain estimates from your organization about the resources allocated to developing comparable test methods, standardized software models, standard reference materials used for measurement and calibration, and standard reference data. We will aggregate this information to quantify the proportion of the biopharmaceutical industry's activities that are allocated to improving the "technology infrastructure" supporting commercial-scale manufacturing.

Technology infrastructure consists of a set of "technical tools and processes" that enable or increase the efficiency of R&D, manufacturing, and market penetration. Collectively, NIST refers to these tools and processes as the **technology infrastructure**. The technology infrastructure provides the technical basis for activities ranging from drug discovery to quality control in the production process. Investments in the technology infrastructure have the potential to significantly lower development, production, and transaction costs and hence are an ongoing activity in most companies. Investment in the technology infrastructure may include both direct and indirect costs in the form of equipment, labor, and material expenditures dedicated to technology infrastructure.

Examples of technology infrastructure include the following:

- Certified reference materials (SRM), reference methods, and standardized procedures related to measurement or calibration in conducting tests during R&D stages or in any manufacturing process.

- Validated methods for interpreting results from different analytical platforms.

- Standardized techniques (e.g., for use with spectrometers, imaging systems, and/or measurement devices)

- Standardized methods for characterizing scientific and engineering data and the algorithms to manipulate, search, and analyze these uniform data within a publicly available database.

- Processing techniques (e.g., for use with 2D Gel electrophoresis platforms, chromatography separation systems)

Click here for examples of **technology infrastructure costs.**

The following bullets will appear in a separate browser window if the respondent selects the link in RED above.

Technology infrastructure Cost Examples

- Purchase of measurement-related equipment, software, reference materials, or services.

- Activities required for setup and validation of analytical instruments, reagents, or other research tools. Examples of these activities may include developing test methodologies or process standards in measurement and manufacturing practices.

- In-house customization of technology platforms purchased from third-party vendors.

- Efforts to develop interoperability between different software or equipment systems.

- License fees or any other spending on enhancements to routinely used processes or equipment, intended to increase productivity, reduce redundancy, or improve the confidence in results.

1. Based on the definition and examples of technology infrastructure provided above, please estimate the expenditures for the technology infrastructure supporting your commercial–scale manufacturing activities. *Please consider all labor, materials, and equipment expenditures for in-house activities, as well as contributions to consortia or partnership initiatives.*

Approximate Annual Expenditure on Technology Infrastructure =

$_____

Technology infrastructure Expenditure Category	Approximate Percentage of Annual Expenditure on Technology infrastructure
Labor (includes all research, implementation support activities, and participation in consortia related to infrastructure technologies)	<<value 0 to 100>>
Capital (includes all equipment, licensing fees, and …)	<<value 0 to 100>>
Materials (includes all reagents, reference materials)	<<value 0 to 100>>
Total	100%

2. Approximately what share of total "**sales**" does the dollar value reported in Question 1 represent?

   ☐ _____ % of total current annual sales

   SKIP LOGIC: IF Respondent answers $0 or 0% THEN the survey will terminate with a "Thank you for your participation in this important NIST study" message.

3. Over the last 5 years, have these expenditures increased, decreased, or stayed the same?

   ☐ **Increased**

   ☐ **Decreased**

   ☐ **Constant**

   ☐ Comments: _____

Part III-B: Assessment of the Potential Impact of Improvements in Technology Infrastructure

The purpose of this section is to collect information that will help us assess the potential impact that certain improvements to the underlying technology infrastructure could have on the current cost of biopharmaceutical manufacturing.

1. Please estimate the current share of total manufacturing costs accounted for by the following phases of production for a typical biopharmaceutical like those your organization manufactures.

Production Phase/Activity	Approximate Percentage of Total Annual Manufacturing Costs
Preproduction (includes *process engineering* and pilot *scale-up* activities to move from clinical trial to commercial volumes)	
Upstream processing	
Downstream processing	
Process monitoring across all phases of production	
Postproduction market transactions (e.g., testing, handling, and distribution)	
Total Manufacturing Costs	**100%**

2. To help us establish a baseline figure for these manufacturing costs, please estimate total manufacturing costs as a share of annual sales. The estimate may be for a particular product line(s) or an industry estimate for firms similar to yours.

    Manufacturing costs account for approximately _____% of annual sales.

3. Based on your knowledge of emerging technologies and other developments in the biomanufacturing industry to improve the productivity, quality, and efficiency of the biopharmaceutical production process, which phase/activity has the greatest potential for cost reductions over the next 5 years? Please only check one.

    □ Preproduction

    □ Upstream processing

    □ Downstream processing

    □ Process monitoring (across all phases of production)

    □ Postproduction market transactions (e.g., testing, handling, and distribution)

    □ Other, please specify: _____

4. The NIH [Technology Road Map](), the FDA [Critical Path]() Opportunities List, and other industry sources have identified a number of areas that could benefit from additional research. Consider advancements in the following areas related to the technology infrastructure supporting different stages of manufacturing. These exist in varying degrees today and are expected to improve in the coming years.

Preproduction:

    A. Predictive modeling and underlying data for determining batch yield given a specified level of inputs and production parameters (which would reduce the number of test batches required during process scale-up).

Upstream Processing:

    B. Robust expression systems that produce raw proteins in higher yields with fewer impurities (for example, the use of transgenic plants and animals).

    C. Technologies, such as (but not limited to) disposable bioreactors and mixing systems, that can accommodate rapid changes in manufacturing processes and can reduce the risk of biological product contamination.

Downstream Processing:

    D. Purification technology that can improve flow rates and increase capacity over current methods. Examples of new technology might include membrane chromatography or improved, high-pressure affinity chromatography among other methods.

Process Monitoring (crossing over all phases of production):

    E. Implementation of process analytics technology (PAT) to better understand, monitor, and control production processes in real time. For example, this might involve the following:
        a. Improved detection of contamination in biological products (e.g., viruses, bacteria, and other organisms) through the use of microarrays or proteomics technologies.
        b. Uniform standards for spectroscopic instruments. For example, standards for appropriate instrument qualification and calibration standards for techniques such as Raman and Terahertz spectroscopy.

    F. Improved methods for product characterization, including enhanced potency assays and appropriate statistical and sampling techniques. For example, this might involve the following:
        a. More reliable and quantitative nonanimal-based tests of vaccine potency.
        b. Potency measurements that provide reliable information about the quality of cells or tissues to be used in therapies.

Postproduction:

    G. Improved certainty about product characteristics, including potency, sterility, purity, and handling requirements.

Enhancements in these areas could lead to a number of impacts on manufacturing productivity, such as the following:

- Reduced time to market as a result of better scale-up procedures, shorter downstream processing purification periods, and faster and more reliable batch testing procedures.

- Improved product yields for a given amount of raw material.

- Reduced labor and material costs with fewer trial and error iterations and fewer contaminated batches.

- Labor saved in sterilization activities between batches.

- Lower transactions costs to customers as a result of greater reliability of data and fewer instances of inactive product released.

- Smaller amount of product required for testing.

Impact on Manufacturing

Consider the following "improved world." Suppose there is an immediate advancement in the above areas of 50% over the current state of the art. What would you estimate the impact on the manufacturing process to be for a typical biopharmaceutical?

Manufacturing Costs	Percentage Reduction in Cost by Phase	Check Here if Not Familiar with This Activity
Preproduction (scale-up from clinical trial volumes) costs		☐
Upstream processing costs		☐
Downstream processing costs		☐
Process monitoring and quality assurance testing costs		☐
Postmanufacturing market transactions costs		☐

Comments:_____
_____

Insert "Save & Continue" button after this question. Following Question will appear on a new screen.

THIS IS THE END PART III-B, Respondents will now be asked questions related to their post-market monitoring activities.

Part II-C: Current Expenditures on Technology Infrastructure

Post-market Monitoring Activities

The purpose of this section is to obtain estimates from your organization about the resources allocated to developing comparable test methods, standardized software models, standard reference materials used for measurement and calibration, and standard reference data. We will aggregate this information to quantify the proportion of the biopharmaceutical industry's activities that are allocated to improving the "technology infrastructure" supporting post-market monitoring.

1. Based on the same definition and examples of the technology infrastructure provided in Part II-B, please estimate the expenditures for the technology infrastructure supporting your post-market activities. *Please consider all labor, materials, and equipment expenditures, both for internal use and contributions to consortia or partnerships.*

Approximate Annual Expenditure on Technology Infrastructure =

$_____ _____

Technology infrastructure Expenditure Category	Approximate Percentage of Annual Expenditure on Technology infrastructure
Labor (includes all research, implementation support activities and participation in consortia related to infrastructure technologies)	<<value 0 to 100>>
Capital (includes all equipment, licensing fees, and …)	<<value 0 to 100>>
Materials (includes all reagents, reference materials)	<<value 0 to 100>>
Total	**100%**

2. What share of total sales does the dollar value in Question 1 represent?

    ☐ _____ % of total current annual sales

    SKIP LOGIC: IF Respondent answers $0 or 0% THEN the survey will terminate with a "Thank you for your participation in this important NIST study" message.

3. Over the last 5 years, have these expenditures increased, decreased, or staved constant?

    ☐ **Increased**

    ☐ **Decreased**

    ☐ **Constant**

3b. Comments: _____

Part III-C: Assessment of the Potential Impact of Improvements in Technology Infrastructure

The purpose of this section is to collect information that will help us assess the potential impact that certain improvements to the underlying technology infrastructure could have on the current cost of biopharmaceutical post-market monitoring activities.

1. Please estimate the current share of total post-market monitoring costs accounted for by the following activities for a typical biopharmaceutical like those your organization monitors.

Post-market Monitoring Activities	Approximate Percentage of Total Annual Post-market Monitoring Costs
Adverse event (AE) monitoring	
AE and safety data management	
Regulatory drug safety reporting	
Product label updating	
Other related activities	
Total	**100%**

2. Please estimate total post-market monitoring costs as a share of annual sales. The estimate may be for a particular product line(s) or it may be an industry estimate for firms similar to yours.

    Post-market monitoring costs account for approximately _____% of annual sales.

3. Based on your knowledge of emerging technologies and other developments in the biopharmaceutical industry to improve the productivity, quality, and efficiency of the biopharmaceutical development process, which post-market monitoring activity has the greatest potential for cost reductions over the next 5 years? Please check only one.

    ☐   AE monitoring

    ☐   AE and safety data management

    ☐   Regulatory drug safety report

    ☐   Product label updating

    ☐   Other, please specify: _____

4. The NIH Technology Road Map, the FDA Critical Path Opportunities List, and other industry sources have identified a number of areas that could benefit from additional research. Consider advancements in the following areas related to the technology infrastructure supporting post-market monitoring activities. These exist in varying degrees today and are expected to improve in the coming years.

   A. Improved statistical methods for signal detection in AE monitoring.

   B. Development of standards for AE reporting systems.

   C. A standardized process and protocol for ensuring data consistency across international AE databases.

   D. Improved interoperability between safety and clinical database management systems.

   E. Automated monitoring software tools to provide notification of required product label changes based on changes in AE frequencies.

   F. Improved integration linking data entry, clinical databases, randomization databases, and regulatory systems.

Enhancements in these areas could lead to a number of impacts on the productivity of post-market monitoring activities, such as the following:

- Reduced labor costs of physicians required to conduct AE monitoring and data analysis.

- Reduced labor costs in reconciling differences across international AE databases.

- Reduced transaction costs in updating product labels based on AE frequency changes.

- Reduced labor costs to generate electronic submissions to regulator agencies.

- Reduced risk due to the mitigation of potential errors in analyzing and interpreting AE data.

Impact on Post-market Monitoring Activities
Consider the following "improved world." Suppose there is an immediate advancement in the above areas of 50% over the current state of the art. What would you estimate to be the impact on the post-market monitoring costs for a typical biopharmaceutical?

Post-market Monitoring Activities	Percentage Reduction in Cost by Activity	Check Here if Not Familiar with This Activity
AE monitoring		☐
AE and safety data management		☐
Regulatory drug safety report		☐
Product label updating		☐
Other related activities		☐

Comments:_____
_____

THIS IS THE END OF THE SURVEY for respondents who answered PART III-C.

A-32

5. Do you see a role for the National Institute of Standards and Technology (NIST) in helping to achieve the technological advancements discussed in this survey? If Yes, what role should NIST play?

   ☐ Yes

   ☐ No

   ☐ **If Yes, what role should NIST play?_____**

6. What is the total employment of your parent company? Please indicate the approximate range.

   ☐ 1–10

   ☐ 11–50

   ☐ 51–500

   ☐ 501–2,500

   ☐ 2,501–15,000

   ☐ >15,000

7. Approximately how many full-time equivalent (FTE) scientists and engineers were represented in your responses to this survey?

   This may be a subset of your company's total employment.

   _____ # of FTEs

On behalf of NIST, RTI International would like to thank you for your participation in this important survey of the biopharmaceutical industry.

If you would like to be contacted when a final report is available and/or if we can contact you with any questions we may have, please indicate such:

        ☐   I am willing to have RTI contact me about my responses to this survey.
        ☐   Please send me a copy of the final report when it is available.

If you clicked either box above, please provide the following contact information and any additional comments or questions you would like to discuss with us (<u>NOTE</u>: your e-mail address and contact information will not be shared with anyone outside RTI or used for any other purpose outside this project):

Name: _____
Company Name: _____
E-mail Address: _____
Phone Number: _____(optional)

NOTE: If you would like to refer this survey to a colleague who is active in other areas of biopharmaceutical development, please forward them the e-mail invitation you received or enter their e-mail address in the space below. An e-mail will be sent to them with the survey material. Please be sure to include your name and e-mail address above, so we can indicate that you asked that this e-mail be sent.

    Intended recipient's e-mail address: _____