# Bioinformatics and Pharmacogenomics in Drug Discovery and Development

- ***A Socio-Economic Perspective***

**Chukwuma E. Anyanwu**
**I692: Project in Bioinformatics**
**August 2005**

**IUPUI School of Informatics**
**535 West Michigan Street, Room IT 475**
**Indianapolis IN 46202-3103**

**Thesis Advisor: Dr. Josette Jones**
**Assistant Professor, Informatics**
**Assistant Professor, School of Nursing**
**IUPUI, Indianapolis, Indiana**

**This Thesis is dedicated to the sweet memory of**

**Professor Clement Nnorom Anyanwu (University of Ibadan, Nigeria)**

**My father.**

**Table of Contents**

# ABSTRACT

***Objective:*** Literature review to evaluate the extent to which Bioinformatics has facilitated the drug discovery and development process from an economic perspective

***Problem:*** A plethora of genomic and proteomic information was uncovered by the U.S Human Genome Project (HGP). Despite the projected impact that Bioinformatics and Pharmacogenomics were projected to have in the drug discovery and development process, the challenges facing the pharmaceutical companies – in this regard, still persist.

***Design*****:** An extensive integrated literature review of library resources such as MEDLINE, ERIC, PsychInfo, EconLit, Social Services Abstracts, ABI/INFORM and LISA (all 1990 – Present). These electronic databases were researched because of their focuses on the healthcare sector, medical and scientific innovations, economic modeling and analysis, bioinformatics and computational biology, applied social research and technology applications. Semi-structured interviews of Bioinformatics professionals were also conducted to complement the literature review. Also, Internet-based databases from reliable resources were also researched resulting in serendipitous discoveries.

***Sample:*** Published English language reports of studies and research carried out worldwide from 1990 to 2004, relating to drug discovery and development.

***Selection criteria:*** Primary focus was on research publications and journals that identify and discuss the practice of Bioinformatics, especially in the area of drug discovery and

development. Premium was placed on articles and publications that discussed the economic impacts of Bioinformatics in the drug discovery process.

***Results:*** Though the goals of Bioinformatics have been clearly defined, and the discipline is widely practiced in the pharmaceutical industry, this study has not found any definite attempts to evaluate its economic and regulatory impact specifically in facilitating the drug discovery and development process, and the delivery of personalized drugs.

***Discussion:*** Bioinformatics and Pharmacogenomics are the new facets of the ever-evolving drug discovery and development process. It may still be a while before their full impact and potential is attained.

# CHAPTER 1

## INTRODUCTION

In the April 1953 edition of the scientific journal *Nature*, James Watson and Francis Crick published a landmark paper describing for the first time the structure of the Deoxyribonucleic acid (DNA) molecule. Though the term 'Double Helix' was not used in that paper, the proposed double-helical structure of the DNA came to be known as the Watson and Crick model. Exactly half a century – to the month – after this historical publication, the US Human Genome Project announced that the human genome sequence was "substantially complete" (Malorye, 2003). Fittingly, the inaugural director of the Human Genome Project (1990-1992) was James Watson, one of the two men who a half-century earlier had pioneered the modern concept of the genetic basis of life.

With the advent of the ubiquitous US Human Genome Project (HGP), it became apparent in the scientific community, that computer technology would be required to process the plethora of data uncovered from the human genetic makeup. The immediate need as described by the National Center for Biotechnology and Information (NCBI), was to '**store**, **organize**, **analyze**, and **integrate** vast quantities of diverse data, such as DNA and protein sequences, gene and chromosome maps, and protein structures' that was being uncovered by the project (NCBI, 2003). The information harnessed from this process, the NCBI asserted, was necessary to propel among other things, the efficient discovery of drugs and drug targets, and the elucidation of disease and health conditions (NCBI, April 2003). The unique confluence of molecular and genetic biology and computer technology – otherwise known as Bioinformatics – became the immediate solution as it enabled the "application of information technology to the management of

biological data" (Gibas et al., 2001; NCBI, 2003; Samudhram, 2003). The NCBI defined Bioinformatics as the field of science where "biology, computer science and information technology merge to form a single discipline" (NCBI, 2003). Bioinformatics continues to be a novel concept, generating different but related fields of research and development. Some of those include Pharmacogenomics (the application of genomic approaches and technologies to the identification of drug targets), Proteomics (the qualitative and quantitative studies of gene expression at the level of the functional proteins themselves) and a host of other ***–omics*** such as Transcriptonomics that have characterized the genome-classification era (Hofestadt, 2002; Mendible, 2003; NCBI, 2003; Whittaker,2003). These areas of research are all Bioinformatics applications.

Bioinformatics quickly assumed a larger role in the processing of the genetic information with the hope that the genetic basis of health and disease would be unraveled, resulting in the efficient discovery of tailored and targeted drugs.

Targeted drugs refer to the drugs that have been designed specifically to act on particular genes and their corresponding proteins, where these genes and proteins have been identified to be responsible for certain disease conditions (Boswell, 2002; Whittaker, 2003). Targeted drug development is particularly applicable in the treatment of cancers, where the drugs are designed to directly target molecular abnormalities that lead to the growth of tumors (Prows & Prows, 2004).

Tailored drugs on the other hand refer to drugs designed to address the needs of a specified genetic sub-group of the entire population (Rioux, 2000; Hall, 2003). This practice, also referred to as pharmacogenetics (or pharmacogenomics) is based on the premise that genetic variations within any given population dictates that individuals, or

sub-groups within any given population will respond differently – or not even respond at all – to the same medication (Rioux, 2000; Hall, 2003). Pharmacogenomics is the study of how genes affect a person's response to drugs. This relatively new field combines pharmacology (the science of drugs) and genomics (the study of genes and their functions) to develop effective, safe medications and doses that will be tailored to a person's genetic makeup (Mendible, 2003).

The specific aim of this thesis is to review literature that addresses the socio-economic impact that Bioinformatics and Pharmacogenomics have or can have on the drug discovery and development process. Bioinformatics and other related fields such as Pharmacogenomics are viewed by some writers as distinct areas of research (Lim, 1997; Mahlich et al., 2001), while some others see Pharmacogenomics as the application of Bioinformatics (Myers and Baker, 2001; Whittaker, 2003). Nightingale (2000) classified Bioinformatics and related fields as "bioinformatics and bioinformatics technologies" and posits that they "have allowed pharmaceutical firms to exploit economies of scale in experimentation." In concurrence, Overby (2001) adds that technologies "grouped under the umbrella of bioinformatics … involve the use of computers to store, organize, generate, retrieve, analyze and share genomic, biological and chemical data for drug discovery". The view of this thesis is that Bioinformatics is an all-encompassing field that defines any confluence of molecular biology and computer information science (Counsell, 2004). Bioinformatics as thus, would play a significant role in drug target discovery (the discovery of suitable drug targets in the human DNA) – by mining and analyzing genomic and proteomic data etc – and drug target validation (the validation,

through experimentation that a drug target is the appropriate one) – by linking targets to biological and drug function (Whittaker, 2003).

This thesis is arranged into 6 distinct chapters. The introductory chapter (preceded by an abstract overview) is immediately followed by the Background and Significance. Here, the background is reviewed and the significance of the thesis is discussed.

Next is the problem statement and purpose. The problem, once identified, will enable the formulation of the purpose.

The Methods and Design chapter will reveal the ways and means by which information is gathered and selected. The data sources are also revealed.

The Results of the literature review are presented in the following chapter. The concluding chapter is the Discussion and Future Work.

# CHAPTER 2

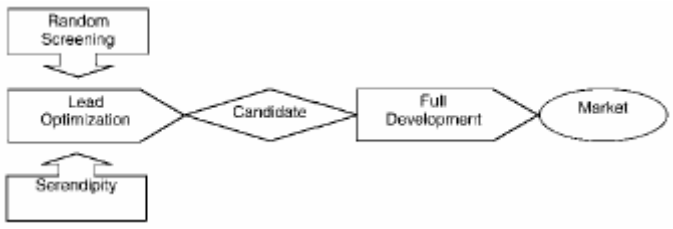
## BACKGROUND AND SIGNIFICANCE

The Merriam-Webster online dictionary defined a drug as “a substance intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease…or as a component of a medication”. The origin of the use of potent compounds for the mitigation of disease conditions predates both the pharmaceutical industry and history. The discovery of drugs historically has been serendipitous (Ratti & Trist, 2001). For most of its history, drug discovery has been a product of trial and error, “guided as much by the intuition and serendipity of chemists, biologists and physicians as by any rational linear process” (Papanikolaw, 1999). Long before the pharmaceutical industry existed, drugs were discovered by accident and their uses passed down by verbal and written records (Ratti & Trist, 2001). According to Boa (2003), “Throughout history people have found by trial and error which berries, roots and barks could be used for medicinal purposes to alleviate symptoms of illness”. For example, the Willow bark, which contains *salicin,* was used as a fever reducer in the same vein as the Cinchona bark from which *quinine* was discovered (Boa, 2003). The drug discovery process would continue to be by trial and error until early - mid 20$^{th}$ century when the pharmacological basis of diseases and drugs were beginning to be defined (Boa, 2003). Even then, serendipity was sought in the pharmaceutical industry as evident by the screening of known compounds or randomly testing any available molecules. Such successful drugs as chlorpromazine, meprobamate, and benzodiazepines were discovered in this way (Ratti & Trist, 2001).

In this process, lead molecules found by chance or from screening the chemical diversity available were then optimized by medicinal chemists to produce candidates, which were passed to development and eventually into the market (Fig 1). However, this approach at that time suffered from a lack of sufficient molecules with high enough structural diversity among other factors.

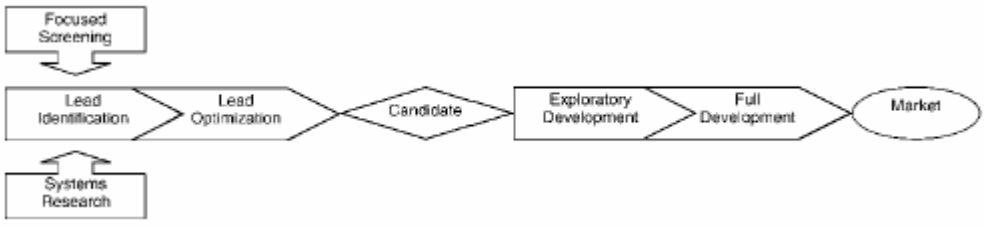
A more rational approach was later developed to improve on this process. In this elongated approach which became mainstream in the 80s, *in vitro* assays using animal tissues (rather than previously used *in vivo* methods) became central in the process for giving valuable information on structure–activity relationships and eventual pharmacophore construction (Fig 2). In this way, if the lead molecule fails there is sufficient information around structure and activity to allow the cause for failure to be extricated from the molecule.

Fig 1: Drug Discovery in the 50s and 60s



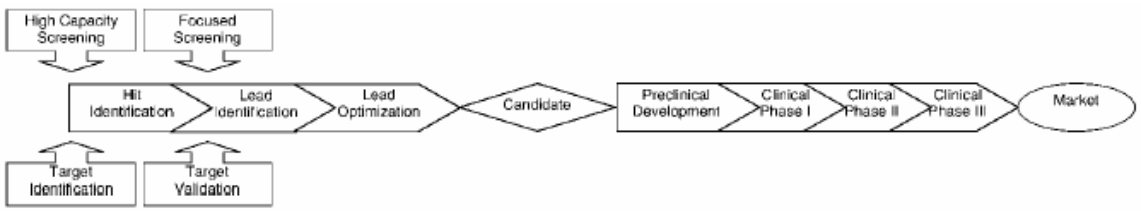

Source: Ratti & Trist (2001)

Fig 2: Drug Discovery in the 80s




Source: Ratti & Trist (2001)

Fig 3: Drug Discovery Today (90s and beyond)




Source: Ratti & Trist (2001)

The projection that Bioinformatics will lead to the era of tailored and targeted drugs has been expressed by many scholars including Van Arnum (1998), Papanikolaw (1999), Attwood et al. (2003), Lindpaintner (2003) and Evans et al. (2004). Evans et al. (2004) described this impact as the development of “subpopulation-specific drug. While the current practice of drug development addresses the needs of the majority of the population, scientific data has shown that people vary in their response to the same drug (Nuffield Council on Bioethics, 2003; Evans et al. 1999, 2004; Prows & Prows, 2004). Few drugs are effective for everyone; all may cause adverse reactions or occasionally death for individuals in sub-groups of the general population (Prows & Prows, 2004). In fact, in 1994, it is estimated that at least 106,000 people died from adverse reactions to "safe," FDA-approved drugs (Lazarou et al., 1998). This figure includes only hospitalized patients and does not include those people who died because of medical error such as the prescribing of the wrong drug, which also accounts for some 100,000 deaths every year (Kohn et al., 1999). Some of the variation between individuals in response to drugs is due to differences in their genetic make-up (Nuffield Council on Bioethics, 2003; Evans et al. 1999, 2004; Prows & Prows, 2004). This variation generally referred to as genetic polymorphisms can be seen as a “stable difference in DNA sequence at the same locus (a

specific position in the genome) among individuals" (Rioux, 2000). The differing DNA sequence at the same position – known as alleles – can lead to a difference in the expression of the genotype (Rioux, 2000). *Genotype* refers to a person's specific allelic composition while the *Phenotype* refers to the observable or measurable manifestation of a person's genotype, either by itself or in coordination with environmental factors (Rioux, 2000; Prows & Prows, 2004). The most common form, or *allele,* of a gene found within a population is known as the *wildtype* allele. Alternate forms result from a change in the gene's chemistry or structure. It is estimated that 99.9% of the human genome sequence is identical in all individuals, despite observable differences. Differences in the human genome, or DNA alterations, are called *mutations* if they are rarely found within a population and *polymorphisms* if they are more common – that is, found in 1% or more of a population (Rioux, 2000; Prows & Prows, 2004). In scientific literature, a gene alteration is mostly regarded as a mutation if it leads to a disease condition or a polymorphism if it leads to no observable effects (Prows & Prows, 2004). Strictly speaking however, both mutations and polymorphisms differ mainly in their relative frequencies of occurrence in the general population, and can have no effect, a beneficial effect, or lead to a disease (Rioux, 2000; Bell, 2004; Prows & Prows, 2004). Similarly, mutations and polymorphisms lead to very diverse differences in the ways that individuals react to drugs when administered (Rioux, 2000; Goldstein et al., 2003; Bell, 2004; Prows & Prows, 2004). Tsai and Hoyme estimate that more than 20 polymorphisms have been identified in drug metabolizing enzymes alone, often with diverse frequencies among individuals from various racial and ethnic backgrounds (Tsai

and Hoyme, 2002). According to Goldstein et al., 2003, the number is far more. In fact they state as follows:

> *Forty-two polymorphisms that have been significantly associated with drug response in at least two studies show that obvious candidate genes, such as drug-metabolizing enzymes and drug targets, often carry important pharmacogenetic polymorphisms, and that such polymorphisms are often owing to common alleles.*

Tsai and Hoyme (2002) went on to identify several mechanisms through which polymorphisms can cause alterations in drug effect. These include:

- Extended pharmacological effect – whereby drugs have more extended effects than anticipated
- Adverse drug reaction
- Lack of pro-drug activators – whereby the drugs become inactivated and therefore ineffective
- Drug toxicity and
- Increased or decreased effective dose

The correlation between genetics and drug responses has been described in scientific literature for many decades. In fact, according to Tsai & Hoyme (2002) Vessel and Page in their 1968 investigation of drug responses in the general population showed "genetic influence on drug metabolism, as drug half-lives were markedly alike in monozygotic *[identical]* twins and varied widely in dizygotic *[non-identical]* twins and the population at large" *[Emphasis mine].* Monozygotic twins share the same identical genomes, while dizygotic twins share genomes just like non-twin siblings.

Despite the available information that closely links drug reaction with genetic identities, the current practice of drug discovery generally does not allow a pharmacogenomic strategy – a strategy that applies the need to have drugs designed to cater for the needs of the various genotypic sub-groups of the population. Drugs are generally designed by applying a 'majority rule' strategy where discovery and development are geared towards the satisfaction of the needs of the major genetic sub-groups of the population. Prows & Prows (2004) identified this application of the 'average dose' to be enormously costly to society. They observe as follows:

> "Currently, most medications are prescribed without assurance of efficacy in a given individual. Often, several drugs must be tried to find one that works. Adverse drug reactions are prevalent and account for hospital expenditures estimated at up to $5.6 billion annually."

The rising costs and inconveniences of prescribing drugs to patients who may well end up as non-responders is generating concern and ultimately interest in research in this area (Burke et al., 2004; Prows & Prows, 2004). The possibility of mitigating – by means of a pharmacogenomic approach to drug discovery – the high healthcare care costs associated with adverse drug responses has generated strong interests not only in medicine, science and the academia, but also within the pharmaceutical and health insurance industries (Burke et al., 2004; Prows & Prows, 2004). The interest expressed by third-party payers – health insurance companies – suggests that this approach may lead to an overall decrease in the cost of healthcare (Prows & Prows, 2004).

Rioux (2000) while championing a pharmacogenomic approach to drug development notes the distinction between the genetic polymorphism in patients resulting in them requiring varying dosages of the same drugs for the desired effects to be achieved, and

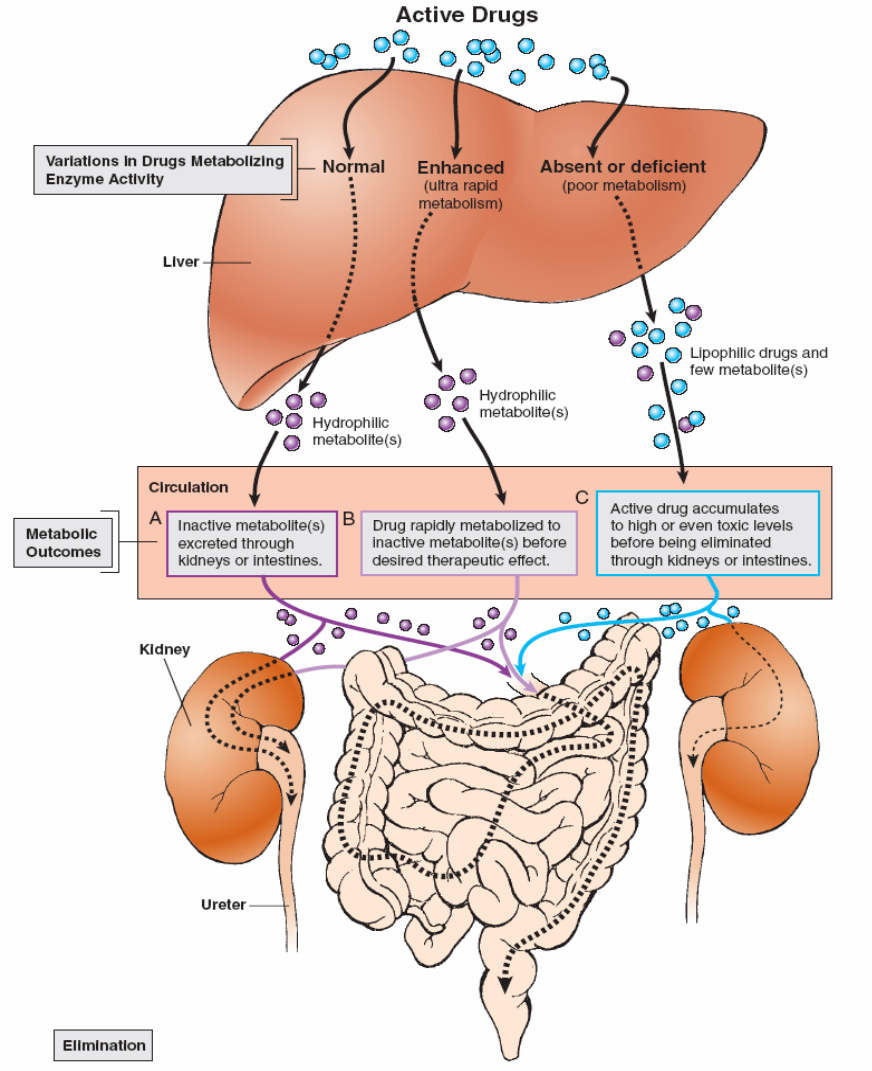
& Prows, 2004).

The genetic variations and the resultant effects on the metabolisms of drug substances are often not trivial, as seen from the observation of Rioux (2000) below:

> "Many drug-metabolizing enzymes exhibit polymorphic expression… but the principal enzymes causing hereditary variations in drug metabolism found in the liver belong to the cytochrome P-450 (CYP) isoenzyme system. Polymorphic expression of at least three of these isoenzymes divides the human population into two groups: "poor" metabolizers, whose genes express dysfunctional or inactive enzymes, and "extensive" metabolizers, whose genes express enzymes with normal activity. Poor metabolizers account for 3-7% of whites ; the distribution of these polymorphisms can be very different in other ethnic groups. Furthermore, the magnitude of variation in drug metabolism between poor metabolizers and extensive metabolizers is normally not trivial; the rate at which many drugs are metabolized may vary 10-fold to 100-fold between these two groups of people"

Prows & Prows (2004) included a third group of metabolizers – the normal, in addition to the enhanced and the deficient. Figure 1 below (Prows & Prows, 2004) shows the drug metabolizing differences between the various drug-metabolizing sub-groups.

**Fig. 1: The Effects of Genetic Variation on Drug-Metabolizing Enzymes (Prows & Prows, 2004)**




Although there seems to be sufficient evidence to show that the traditional drug discovery and development process as commonly practiced today may not cater for everyone's needs (Rioux, 2000; Michalowski, 2001; Evans et al. 2004), there are several factors mitigating the effectiveness of applying Bioinformatics in the quest for targeted and tailored drug development. Some of these factors are economic (Michalowski, 2001; Evans et al. 2004) ethical (Rioux, 2000; Michalowski, 2001; World Health Organization, 2002; Nuffield Council on Bioethics, 2003), legal (Michalowski, 2001; World Health Organization, 2002; McCubbin, 2003) and political (Michalowski, 2001; World Health

Organization, 2002). The ethical, legal and political issues concerning the application of Bioinformatics in pharmacology are too vast to be covered in this thesis. For more information on these, please refer to " Ethical, legal and social implications (ELSI) of human genomics" (http://www.who.int/genomics/elsi/en/) a World Health Organization (WHO) report.

The economic factors however will be reviewed in this paper. The importance of evaluating the economic factors that impact the application of Bioinformatics in drug discovery and development seems evident. When a pharmaceutical company embarks on a search for a new drug or new chemical entity, the overwhelming motivation for this is the quest for profits (Gilbert et al., 2003; Rasmussen, 2003; Tait & Mittra, 2004). A great majority of pharmaceutical companies are publicly traded investor-owned profit-oriented companies (Gilbert et al. 2003; Rasmussen, 2003). The high and rising costs of drug discovery and development combined with lower average margins and shorter exclusivity (patent protection) periods have forced the pharmaceutical companies to rely heavily on blockbuster drugs – drugs that generate sales upwards of $1billion per year (Gilbert et al., 2003; Klein & Tabarrok, 2003; Rasmussen, 2003; Tait & Mittra, 2004). This business model – the quest for blockbuster drugs – is investor driven (Keefer, 2003). With the increasing costs and the need to maintain healthy drug-candidate pipelines, pharmaceutical companies are forced to discover ways of becoming more innovative and cost effective (Papanikolaw, 1999; Ratti & Trist, 2001; Keefer, 2003; Tait & Mittra, 2004). Bioinformatics and related technologies have been widely considered in the pharmaceutical industry in part because they are expected to make the drug discovery process less tedious (Duggan et al., 2000), more time efficient (Ratti & Trist, 2001), more

successful (Zemlo, 2004) and more cost-effective (Overby, 2001). It [Bioinformatics] will allow pharmaceutical firms to exploit “economies of scale in experimentation” (Nightingale, 2000) through high throughput screening (Ratti & Trist, 2001). Investments in Bioinformatics are on the rise. Experts project that the Bioinformatics market would grow from an estimated $290 million in 1998 to about $1.7billion in 2005 (Papanikolaw, 1999).

# CHAPTER 3

## PROBLEM STATEMENT AND PURPOSE

### Problem Statement

Despite all the investments and high expectations from Bioinformatics, the pharmaceutical industry has not shown significant change – if at all, in the drug discovery and development process (Lindpaintner, 2003). This may be because the practice of Bioinformatics is relatively new and has only attained prominence in the years following the particl completion of the HGP (Lindpaintner, 2003). Bioinformatics has not had the impact it was expected to have – an impact that many still expect and project (Lindpaintner, 2003). The pharmaceutical industry continues to witness rising costs (by some estimates, as high as 55% in the last 5 years) and withdrawals of drugs from the market after they had been approved and commercialized – due to multiple documented cases of drug toxicity (Gilbert et al., 2003; Tait & Mittra, 2004).

The drug discovery and development challenges facing the pharmaceutical industry have been noted by several writers (Papanikolaw, 1999; Duggan et al., 2000; Gilbert et al., 2003; Klein & Tabarrok, 2003; Rasmussen, 2003; Tait & Mittra, 2004). The challenges range from – among others, the high cost of drug discovery and development, the lengthy and risky trials and approval process (leading to a shrinkage of the patent protection period), the occasional withdrawal of previously approved drugs from the market and the innovation gap resulting from the dogged quest for blockbuster drugs (Papanikolaw, 1999; Duggan et al., 2000; Gilbert et al., 2003; Klein & Tabarrok, 2003; Rasmussen, 2003; Tait & Mittra, 2004). Bioinformatics was widely projected to invigorate the

identification of drug targets (Papanikolaw, 1999; Overby, 2001; Ratti & Trist, 2001; Keefer, 2003; Tait & Mittra, 2004; Zemlo, 2004). The fact that these problems remain mostly unsolved despite significant Bioinformatics investments is an indication of a larger problem (Lindpaintner, 2003).

The specific problems as identified by this thesis are as follows:

- There are challenges facing the pharmaceutical industry concerning drug discovery and development. The industry relies heavily on a "Blockbuster drug" development strategy, where a drug is designed to meet the needs of several million people. These drugs are more profitable and meet the needs of the profit-seeking shareholders of the pharmaceutical companies.
- A consequence of the blockbuster model is the "innovation gap" in drug discovery and development. Most blockbuster drugs address medical needs for which drugs have already been developed. For the most part, new chemical entities (NCEs) or drugs are not developed in this model.
- The direct impact of the innovation gap is the neglect of drugs for "rare" diseases. The drugs for many diseases that are not prevalent in the general population have been neglected. By definition, a rare disease in the United States is one suffered by 200,000 people or less. In reality, pharmaceutical companies mostly concentrate only on blockbuster drugs, which are taken by 20 million or more people. This means that a significant number of people have no choice but to take drugs that are not targeted for the treatment of their diseases.
- Bioinformatics and related technologies have been widely touted to be the solution to the pharmaceutical industry challenges. To this end, investments in

Bioinformatics have in recent years been on the rise. The number of colleges and universities awarding Bioinformatics degrees in the United States alone has increased from 27 in 2002, to 50 in 2003 and 71 as of July 2004. Despite the increasing interest in Bioinformatics and related technologies, the drug discovery and development problems still persist

- The changes that may be required to ensure the success of Bioinformatics strategy have not been made. The pharmaceutical industry continues to perpetuate the blockbuster model and is still not attentive to the needs of sub-groups of the population, stratified according to their genetic profiles.
- The FDA, the governmental approval authority continues to test all drugs for use by the general population. Potential drugs that may be efficacious for a segment of the population, and be non-effective for other segments are doomed to failure. This testing and approval process is partly responsible for the high cost of drug discovery and development, and the alienation of many sufferers of rare diseases.

## Purpose

This thesis is a review of existing literature on Bioinformatics, and its related sciences, particularly Pharmacogenomics and Pharmacogenetics with an emphasis on publications and research materials that discuss drug discovery and development. The thesis will assess the economic and regulatory factors that determine the current practice of drug discovery and development and how well these approaches fit Bioinformatics-engineered drug discovery and development. From the background information provided, it can be

inferred that Bioinformatics may lead to a more efficient generation of tailored and targeted drugs. This thesis will review the merits or demerits of that perspective.

The literature review is complimented with semi-structured interviews of experts in the field as a means to verify and validate the findings of the literature review.

The review of literature is from a socio-economic perspective. Ultimately, the development of drugs is for the treatment of diseases suffered in the general population. The costs of these drugs are passed on to the final consumers, and by implication, any factors that affect the cost of the development and discovery of these drugs will directly impact – positively or negatively – society.

The specific purposes of the literature review include:

- To understand the current prevailing process of drug discovery. This is an evaluation of the current practices that generally follow a traditional non-Bioinformatics approach to drug discovery including the testing and approval process by the appropriate government agencies.
- To evaluate the current model of cost evaluation of the drug discovery and development process.
- To understand the propositions of a Bioinformatics and Pharmacogenomic approach to tailored and targeted drug discovery and development.
- To review the differences between the traditional approach and the Bioinformatics approach, with an emphasis on overall costs evaluation

# CHAPTER 4

## METHODS AND DESIGN

The findings of this thesis are the result of an integrated literature review. This chapter details the sources for the reviewed data, the search terms and result sets and the selection/inclusion criteria.

### Data Sources

In order to provide an integrated results set, the following different databases were selected as data sources:

- MEDLINE (1990 – 2004)
- ERIC (1990 – 2004)
- EconLit (1990 – 2004)
- ABI/INFORM (1990 – 2004)

The databases are repositories for different genres of data, and were selected for their various focuses on medical science and technology, health information and education, social sciences and economics. The rationale for the selection of each of these databases is explained below.

**MEDLINE** is a bibliographic repository for a wide array of journals, articles and publications on medicine, nursing, dentistry, veterinary medicine, the health care system and pre-clinical sciences. This data source was selected because of its concentration on medical and healthcare related articles. Bioinformatics and related technologies are inclusive of the overall medical sciences genre.

**ERIC**, the Educational Resource Information Center is a national information system supported by the U.S. Department of Education, the National Library of Education, and the Office of Educational Research and Improvement. With Bioinformatics gaining prominence in scientific literature, it is being widely introduced into academic curriculums

**EconLit**, published by the American Economic Association, provides bibliographic coverage of a wide range of economics-related literature. An expanded version of the *Journal of Economic Literature (JEL)* indexes of journals, books, and dissertations, EconLit covers both economic theory and application. Bioinformatics and related technologies can have direct and significant impacts on the economics of healthcare. Health, education, and welfare economics are some of the identified subject coverage areas for the EconLit bibliographic database.

**ABI/INFORM** contains content from thousands of journals that help researchers track business conditions, trends, management techniques, corporate strategies, and industry-specific topics worldwide. This data source was selected because of its publications on the pharmaceutical industry and the business models therein. It also contains up to date information on the news and current affairs of the pharmaceutical industry. With other databases, current affairs concerning Bioinformatics from trade publications and magazine articles are relatively rare.

The **World Wide Web** (www) was also a source for latest, but sometimes fugitive Bioinformatics information. Considering the relative newness of Bioinformatics in mainstream of scientific literature, the WWW, a hyper-linked document repository, independent of geographic locations allows for current research publications and articles

to be readily available. Articles from the World Wide Web however need be considered with care, as the integrity of such can sometimes be difficult to verify.

**Semi Structured Interviews** were also conducted. Bioinformatics has just recently been gaining prominence in the Pharmaceutical industry and is increasingly being employed in the drug discovery and development process. Early adopters of Bioinformatics techniques and practices were interviewed to gather feedback on its application. These interviewees – practicing Bioinformaticians – were very helpful both in complementing the relative scarcity of literature and in validating the results of the literature review.

## Search Terms and Selection Criteria

### Search Terms

A first search was carried out in Medline using the term “Bioinformatics”. The Medline search process mapped the search term to the thesaurus term “Computational Biology”. A review of the scope of this mapped term showed that Computational Biology was introduced as a subject heading in 1997, and encompasses all methods and theories, relating to Molecular Biology, concerned with the development of techniques for the collection and manipulation of biological data, and the use of such data to make biological discoveries or predictions. This search term is used for Bioinformatics and is related to the search term Medical Informatics. Using the “Computational Biology” search term resulted in a recall of 6958 records. The Medline search routine breaks the

overall search results into various subheadings. This allows for the elimination of the undesired subheadings of the parent search term. The following subheadings were selected:

- Classification
- Economics
- Ethics
- History

These together yielded a recall of 93 records, and a precision of 2. This is hereby referred to as result set A.

Since Bioinformatics was not indexed as a subject heading, it could only be searched as a keyword. This search (Bioinformatics as keyword) yielded in Medline, a recall of 2110 records. Using the “Combine” search tool, the two result sets were combined to show the intersection of the two sets. There were 26 common records (intersection) in this result set - B. The 2 records from A were among this set.

Using the term “Medical Informatics” yielded 3098 records. Pertinent subheadings selected reduced this to 130 records (Classification, Economics, Ethics and History). Combining this result set with result set B yielded 1 journal article, “Bioinformatics and medical informatics: collaborations on the road to genomic medicine” by Kulikowski & Maojo.

Introducing Pharmacogenomics as a search term in Medline revealed that the term is indexed under the heading Pharmacogenetics. This term is used for Pharmacogenomics and is related to Drug Resistance and Toxicology. It has been indexed since 1970 and is a branch of genetics which deals with the genetic variability in individual responses to

drugs and drug metabolism. Search term “Pharmacogenetics” gave a recall of 128 records (selected subheadings) and a precision of 3. A Boolean combination of the Bioinformatics and Pharmacogenetics yielded 0 results.

Several other terms were introduced to explore the available literature on Bioinformatics and related disciplines. Such terms include Healthcare Economics, Healthcare Informatics, Pharmaceutical Industry, Drug Discovery, Drug Development, and Genomics among others. These were also searched in combinations. As an example, Drug Discovery (headings: Pharmacology or Biotechnology) yielded 17525 hits. A combination of this and the Pharmacogenetics result set yielded a recall of 16 records of which 3 were reviewed further.

These same searches were carried out in the ABI/INFORM database (hereon after referred simply as ABI). The search for Bioinformatics in ABI resulted in 650 records of which 6 were found to be relevant to this thesis. Of the 6, 1 article was also part of the result set from the Medline search. The ABI search engine presented a few search terms as complements to the Bioinformatics keyword. These include (with the recall/**precision** numbers from searches carried out with these terms):

- Bioinformatics AND Pharmaceutical Industry (36/**3**)
- Bioinformatics AND Research & Development (33/**3**)
- Bioinformatics AND Genomics (28/**6**)

ERIC and EconLit did not generate many articles. A search using the term “Bioinformatics” resulted in 17 total articles. The precision on this was 2. After a few Boolean combinations of various search terms turned up no articles, the use of ERIC and

EconLit was discontinued. Serendipitous discoveries also added to the articles reviewed. When reviewing a selected article, some referenced articles and publications in that article were also sought and reviewed. The reference provides enough information for a focused search in the databases that allows for the article to be easily discovered – if indeed it exists in the searched database.

## Selection Criteria

All searches were configured for English-language only articles and publications were mostly limited to articles from 1990 to 2004 with the exception of a few historical references such as “The cost of developing a new drug” by Steven Wiggins (1987) that was discovered serendipitously. The rationale behind the decision to focus only on articles written or published no earlier than 1990 is based on the assumption that the theory and practice of Bioinformatics and related technologies has only in very recent years gained prominence in the mainstream of scientific literature. Also, the Human Genome Project (HGP) which significantly catalyzed the theory and practice of Bioinformatics and related technologies was inaugurated in 1990. Bioinformatics and related technologies have evolved significantly in the time since the onset of the HGP such that literature prior to 1990 can potentially be obsolete.

Selection was based on a variety of criteria. First was a review of the titles and the abstracts. This was the first gate for articles selected for further review. The title of an article or publication sometimes can be misleading, so a review of the abstracts (abstract differentiation) gives a better indication of the coverage and scope of the article. The

review of abstracts allows for the summarized contents of an entire article or publication to be quickly evaluated. This enables a timely selection process.

Ultimately, selection was based on the contents of the articles. A review of an article will show if the article is can add value to the thesis. The scope of the thesis was always put in consideration when these articles were reviewed. Not all articles that discussed Bioinformatics and related technologies were relevant to the thesis as the thesis primarily reviewed these in relation to the drug discovery and development process. In this regard many articles and publications reviewed were pertinent to the search terms used but were outside of the scope of this thesis.

For the semi-structured interviews, a series of interviews were carried out with two practitioners in the field of Bioinformatics. These experts work in a prominent pharmaceutical company and are directly involved in the application of Bioinformatics techniques in the drug discovery and development process. The application of Bioinformatics techniques in the pharmaceutical industry is geared towards drug discovery through experimentation. It therefore falls within the classification of Research and Development (R&D). R&D in the Pharmaceutical industry is highly confidential, proprietary and is a major component of intellectual property, and is not ready shared. The interviewees therefore requested anonymity and mostly confirmed or validated the information obtained from the literature review. Their perspectives however gave very useful insights into the way Bioinformatics is practiced in the field. With their express permissions, parts of their accounts have been rephrased and paraphrased for the purpose of this thesis.

# CHAPTER 5

## RESULTS

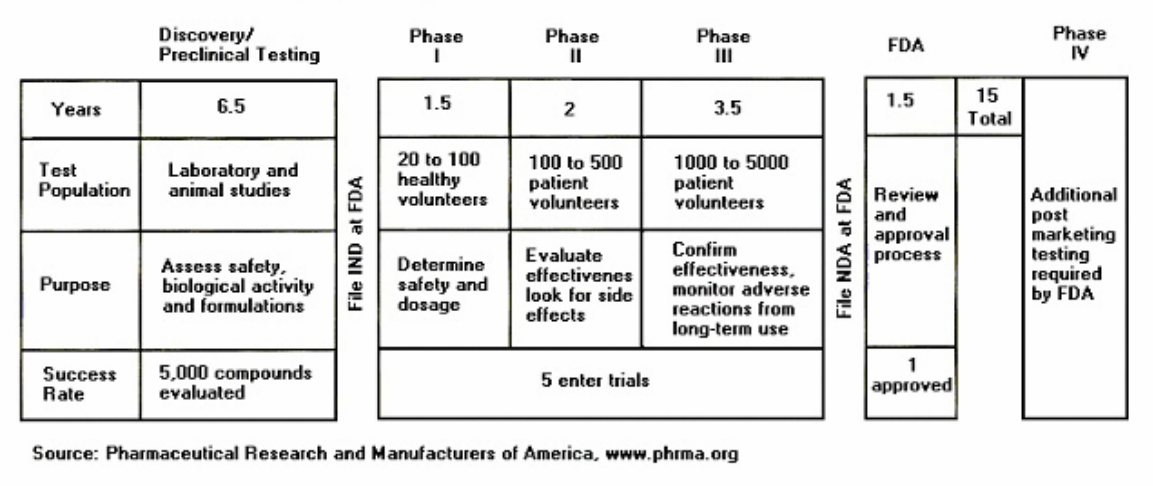
The findings of the literature review are reported in this chapter. The results are arranged in different contiguous sections. The chapter begins with a historical overview of the evolution of drug discovery and development. Next is a discussion of drug discovery and development, as is commonly practiced today. The costs and consequences of the current practice of drug discovery and development is discussed next, followed by a review of the impacts (and potential impacts) of Bioinformatics on this process. The specific impact of Pharmacogenomics is discussed to conclude the chapter.

## Drug Discovery and Development

### *Drug Discovery and Development Today*

Drug discovery today is a very complex, competitive, delicate and risky process. This is partly because the pharmaceutical industry is the epitome of the 'winner-takes-all' philosophy where the first company to patent a drug gets exclusive rights for its use for many years - a passage way to huge financial gains (Duggan et al, 2000). The runner-up gets absolutely nothing (Duggan et al, 2000). The drug discovery, development and approvals process, from conception to market is summarized in Fig. 1 below.

Fig 1: The Drug Discovery, Development and Approvals Process.




Drug companies continuously analyze thousands of compounds, seeking ones of therapeutic value (Klein & Tabarrok, 2003). Drug testing begins at the preclinical stage (Klein & Tabarrok, 2003; Wierenga & Eaton, 2004). In this phase, the manufacturer completes laboratory and animal studies of the compound, to show biological activity against the targeted disease and to verify the safety of the compound (Klein & Tabarrok, 2003; Wierenga & Eaton, 2004). The preclinical testing phase lasts anywhere from three to seven years (Klein & Tabarrok, 2003; Wierenga & Eaton, 2004). Of five thousand compounds tested, approximately five will appear promising enough to induce the company to file an Investigational New Drug Application (IND) (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004). The IND shows results of previous experiments, how, where and by whom the new studies will be conducted; the chemical structure of the compound; how it is thought to work in the body; any toxic effects found in the animal studies; and how the compound is manufactured (Wierenga & Eaton, 2004). If the FDA approves the IND, an Institutional Review Board (IRB) is also required to give its approval before the manufacturer can be permitted to begin the first phase of development (Klein & Tabarrok, 2003; Wierenga & Eaton, 2004). The FDA strictly

monitors the membership of the IRB to ensure that it is composed of “at least five members including at least one scientific member, one nonscientific member, at least one person not affiliated with the research institution, no members with conflicts of interests, both genders if at all possible, and so forth” (Klein & Tabarrok, 2003). The IND stage consists of three phases:

**Phase I**

In phase I, the pharmaceutical companies conduct clinical trials using anywhere from twenty to a hundred healthy volunteers to determine the drug's basic properties and safety in humans (Pratap, 2004). This test stage can last for one to two years (PhRMA, 2003).

**Phase II**

In phase II, efficacy trials begin as the drug is administered to several hundred volunteer patients (Pratap, 2004). In this phase, the patients are given the drug to evaluate how effective it is against the observed signs and symptoms of the disease (Pratap, 2004). The possible side effects are also evaluated (Pratap, 2004). At the end of phase II, the manufacturer meets with the FDA to discuss the development process, continued human testing, any concerns the FDA may have, and the protocols for phase III, which is usually the most extensive and most expensive part of drug development (Klein & Tabarrok, 2003).

**Phase III**

Phase III testing involves one to several thousand patient volunteers and is by far the most detailed, time consuming and expensive clinical trial phase (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004). In this phase, the drug is administered to patients to get more information on its effectiveness, safety, optimal dosage and rare side effects (Pratap, 2004; Wierenga & Eaton, 2004). This essentially is a risk-benefit analysis that allows the FDA decide on the overall benefit of a drug, measured against the observed risks and side effects (Klein & Tabarrok, 2003).

Once Phase III is complete, the manufacturer analyzes all of the data and files a New Drug Application (NDA) with FDA if the data successfully demonstrate safety and effectiveness (Klein & Tabarrok, 2003; Wierenga & Eaton, 2004). Mahlich (2001) describes a Phase IV where the new drug, having been successfully registered is continuously monitored "to collect and evaluate information on rare side-effects, to quantify the therapeutic risks and to determine possible new areas of indication".

During the IND phases (Phases I – III) there are a few accommodations that can be allowed the drug manufacturers. The manufacturer can receive an "accelerated development" status, to allow for the treatment of patients with "life-threatening or seriously debilitating conditions, for which no other drug treatment exists", or a "treatment IND" status to allow for treatment of patients with "immediately life-threatening" conditions (Klein & Tabarrok, 2003). The NDA must contain all of the scientific information that the company has gathered and typically run 100,000 pages or more (Wierenga & Eaton, 2004). Review of the NDA typically lasts one to two years, bringing total drug development and approval (that is, the IND and NDA stages) to

approximately nine years (Klein & Tabarrok, 2003). During the NDA stage, the FDA consults advisory committees made of experts to obtain a broader range of advice on drug safety, effectiveness, and labeling (Klein & Tabarrok, 2003). Once approved, the drug may be marketed with FDA regulated labeling (Klein & Tabarrok, 2003). The FDA also gathers safety information as the drug is used and adverse events are reported, and it will occasionally request changes in labeling or will submit press releases as new contraindications arise (Klein & Tabarrok, 2003). If adverse events appear to be systematic and serious, the FDA may withdraw a product from the market (Klein & Tabarrok, 2003).

For every 5,000 or so compounds tested, approximately five will appear promising enough to file an IND (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004). Of those five, approximately one will be approved by the FDA and make it to market (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004).

For a detailed discussion of the drug development process, please refer to the FDA handbook on this subject (http://www.fda.gov/cder/handbook/develop.htm).

## Cost of Drug Discovery and Development

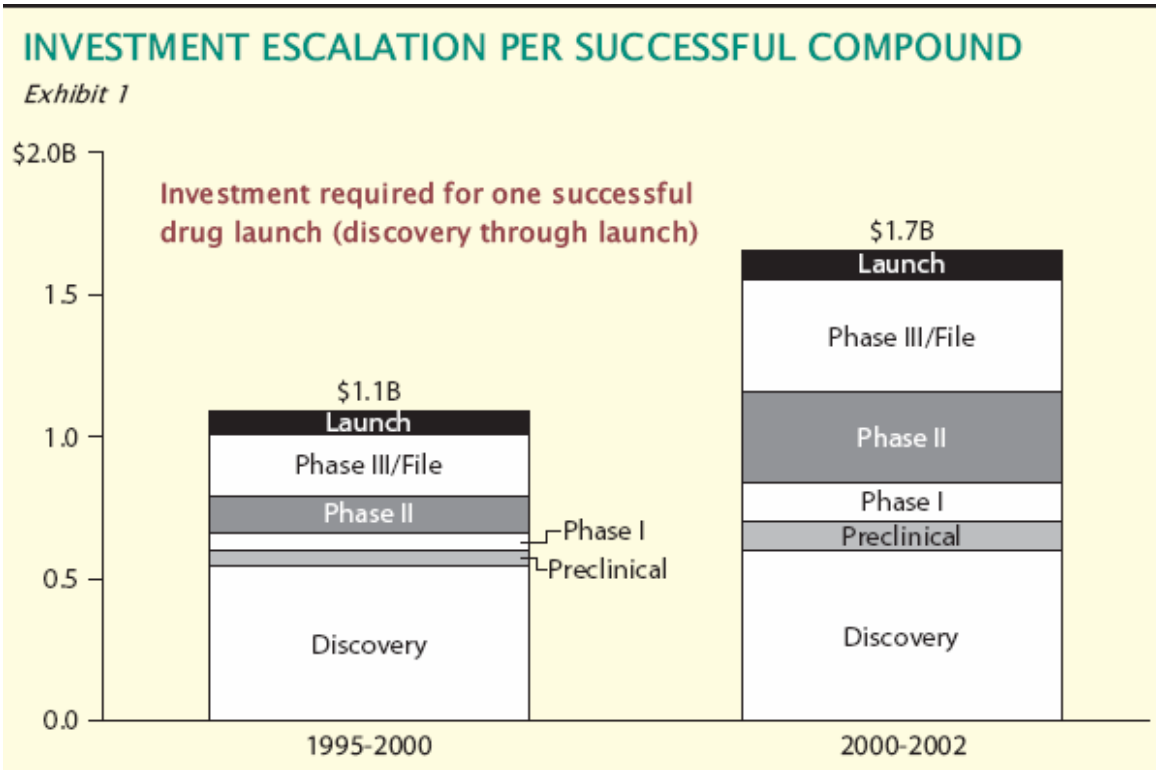
### *Discovery, Development, Testing and Approval Costs*

In addition to the risk and uncertainties, the discovery process is a capital-intensive process, with ever increasing costs (Frank, 2003). There have been differing reports on the actual costs of drug development. Generally speaking, the cost of drug discovery and development is regarded as the total cost from discovery to approval (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004). Wiggins (1987), in his detailed study of

the costs of developing new pharmaceutical drugs showed a substantial and steady increase in the development costs. He reported the cost – at an annual inflation rate of 8% – to be about $54 million in 1976, increasing a decade later to $125 million in 1986 dollars (Wiggins, 1987). The Tufts Center for the Study of Drug Development (CSDD) in 2001 announced that the average cost of developing a new prescription drug was $802 million. The same study had estimated the cost to be $231 million a decade earlier (CSDD, 2001). While this estimate, adjusted in 2003 to $897 million, seems to be the industry standard, the actual cost of drug discovery remains a topic for heated debate (Frank, 2003). *Public Citizen* (2002), a popular nonprofit consumer advocacy organization, referred to the 2001 estimate as the "US$802 million myth". The organization claimed that the drug industry generally "exaggerates the cost of research and development for prescription drugs to justify high prices" (*Public Citizen*, 2002). DiMasi et al (2003) also refuted the $802 million estimate charging that it included "accounting for the time between investment and marketing". After evaluating research and development costs of 68 randomly selected new drugs obtained from a survey of 10 pharmaceutical firms, they proposed the real "out-of-pocket" estimate to be about $403 million, a difference of almost $400 million (DiMasi et al., 2003). The Pharmaceutical Research and Manufacturers Association (PhRMA) however opined that the $802 million estimate was "likely to be conservative" stating that the real cost could be well above the CSDD estimate (PhRMA, 2001). This view was supported by the consulting firm, Bain and Co. in a December 2003 study, that estimated the costs to be approximately $1.7 billion – more than twice the amount announced by the CSDD study just months earlier (Gilbert et al., 2003). The following table shows the costs as proposed by the Bain and

Co.

study.




Source: Gilbert et al.

This new study included the money lost on failed drugs in the clinical trials stages, and money spent on sales and marketing – costs not included in the CSDD report (Gilbert et al., 2003; Mullin, 2003). Many in the drug industry however believe the CSDD estimate to be more accurate (Rogoski, 2004). The FDA concurs with the CSDD view and estimates the costs of drug development to be approximately $800 million with the process from synthesis to approval lasting approximately 15 years (FDA, 2001). While the debate rages on, one fact commonly agreed upon is that the costs of new drug development are enormous, and on the rise. A direct effect of the high and rising cost of drug development is the perception that a drug must be widely successful and be used by a large number of people (Drews & Duyk, 2004).

## Collateral Costs

As was noted in the previous section, the cost of drug discovery and development is generally regarded as the total cost from discovery to approval (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004). There are however other costs that are associated with drug discovery and development. These include – among others – the costs of marketing and commercialization, costs to society as well as costs to pharmaceutical companies as a result of litigations.

### *Commercialization Costs*

The study by Bain Co., which estimated the cost of drug development to be about $1.7 billion included the money lost on failed drugs in the clinical trials stages, and money spent on sales and marketing – costs not included in the generally accepted cost estimate (Gilbert et al., 2003; Mullin, 2003). Gilbert et al., (2003) estimate the average drug launch cost (post approval) to be $250 million.

According to Tait & Mittra (2004), "the products that are easiest and cheapest to develop – and have the largest potential markets – have already been produced and are probably off patent or nearing that point." Pharmaceutical companies in search of immediate profitability often focus on the development of these copycat "me too" drugs that address medical needs for which there are already drugs (Koppal, 2004). The sales and marketing costs are generally high in part because of the "intense promotional "noise" … needed to attract the attention of physicians and patients" who already have access to drugs that offer similar therapeutic benefits (Meyers & Baker, 2001).

Another group of collateral costs include costs associated with drug recalls and litigations.

### *Drug Recalls and Litigations*

The pharmaceutical industry has experienced several withdrawals and recalls of drugs that have hitherto been approved for marketing (Bernard, 2003; Klein & Tabarrok, 2003; Tait & Mittra, 2004). This is often as a result of the emergence of previously unsuspected side effects, or in some cases after accusations of concealment of negative results from clinical trials (Klein & Tabarrok, 2003; Tait & Mittra, 2004). Either way, there usually is an enormous monetary consequence as investors quickly disinvest in the company (where such a company is publicly traded and shareholder owned) and patients flock to the courts with litigations for negligence and pharmaceutical malpractice (Simons et al., 2004). In a 2003 study performed for the U.S. Chamber Institute for Legal Reform, Pendell reported that a large number of medical practitioners have avoided prescribing an appropriate prescription medication because they are aware that such medication may be involved in some product liability litigation (Pendell, 2003). Similarly, patients have stopped taking, or outright refused some medication when they discovered such medication was involved in some product liability litigation (Pendell, 2003). These together ultimately impact – negatively – the revenue potentials of such drugs (Pendell, 2003)

### *Costs to Society*

According to a 1994 study, adverse drug reactions ranked between the fourth and the sixth leading causes of death among hospitalized patients (Lazarou et al., 1998). The authors of this analysis estimated that in 1994 alone, there were 2,216,000 adverse drug reactions of which 106,000 resulted in fatalities (Lazarou et al., 1998). Errors in drug administration, non-compliance, overdose and drug abuse were excluded from this analysis (Lazarou et al., 1998). This is all from adverse reactions to "safe," FDA-approved drugs (Lazarou et al., 1998). A 1999 study by the Institute of Medicine (IOM) stated that “at least 44,000 people and perhaps as many as 98,000 people each year” die in hospitals from preventable medical errors (Kohn et al., 1999). The report indicated that “adverse drug events” was the leading contributor to the overall medical errors, which cost between $17 billion and $29 billion per year in hospitals nationwide (Kohn et al., 1999). The current practice of drug development addresses the needs of the majority of the population. This strategy is counter to scientific data, which has shown that people vary in their response to the same drug due to differences in their genetic make-up (Nuffield Council on Bioethics, 2003; Evans et al. 1999, 2004; Prows & Prows, 2004). Prows & Prows (2004) observed that adverse drug reactions result in annual hospital expenditures estimated at $5.6 billion. These adverse reactions are also costly in terms of the loss of confidence in the healthcare system and a diminishing satisfaction for both patients and health professionals (Kohn et al., 1999).

## Consequences of Current Drug Discovery and Development Strategy

### *The Blockbuster Drug Model*

A blockbuster drug according to Drews & Duyk (2004) is one that is typically marketed for upwards of 20 million people, generating upwards of $1billion a year in sales (Drews & Duyk, 2004). The success of most blockbuster drugs is hinged on their being taken by a large number of people for long periods of time (Tait and Mittra, 2004). All drugs are required to go through the same approval process, whether or not the projected market was two million or twenty thousand people (Klein & Tabarrok, 2003). This means that with the current test and approval process, most drugs will take approximately the same time to get to market (Klein & Tabarrok, 2003), making the "blockbuster drug model" an attractive adaptation. The blockbuster drug model according to Keefer (2003) is that in which drug manufacturers "derive a majority of their revenue from the sales of a few individual patented products". The need by the pharmaceutical companies to recoup the vast sums and still be profitable becomes paramount (Rioux, 2000; Michalowski, 2001). From this perspective, it may be inferred that the blockbuster syndrome is a consequence rather than a cause of the high discovery, development and approval costs (Klein & Tabarrok, 2003). While this may be debatable, the clear result of the high drug discovery and development costs is that pharmaceutical companies rely on blockbusters for their financial sustenance and growth (Rioux, 2000; Michalowski, 2001).

The following table illustrates the results of the blockbuster model for the top 11 pharmaceutical companies in the world in 2002.

Table 1: 'Blockbuster sales by major pharmaceutical companies, 2002

| Company | Pharma sales $m | Blockbuster sales (>$US $1b) $m | Blockbuster ratio | Number of blockbuster drugs |
|---|---|---|---|---|
| GlaxoSmithKline | $28,970 | $14,259 | 49.2% | 8 |
| Pfizer | $27,815 | $22,307 | 80.2% | 8 |
| Merck | $21,627 | $14,055 | 65.0% | 5 |
| Aventis | $18,446 | $5,090 | 27.6% | 3 |
| BristolMyer Squibb | $18,119 | $6,056 | 33.4% | 3 |
| AstraZeneca | $17,841 | $7,746 | 43.4% | 3 |
| Johnson & Johnson | $17,151 | $5,900 | 34.4% | 3 |
| Novartis | $15,181 | $3,026 | 19.9% | 2 |
| Pharmacia | $12,037 | $3,050 | 25.3% | 1 |
| Wyeth-Ayerst | $11,733 | $3,143 | 26.8% | 2 |
| Eli Lilly | $11,078 | $4,693 | 42.4% | 2 |
| Total Top 11 | $199,998 | $89,325 | 40.7% | 40 |

Source: Annual reports and Credit Suisse First Boston.

This approach is investor driven and has been applied successfully in both the automobile and movie industries (Keefer, 2003). It is however causing a significant amount of economic pressure on the drug discovery and development process. In fact, Joe Chimera, PhD, the head of the Pharmaceutical Technology Group for BioSignia posits that “the pressure to maintain the existing types of margins to keep shareholders happy is forcing pharmaceutical companies to focus their efforts on developing at a minimum one or two big blockbusters a year” (Keefer, 2003). The blockbuster model has also resulted in a sharp decline in pharmaceutical innovations (Whittaker, 2003). According to Tait and Mittra (2004), the drugs that are “cheapest to develop – and have the largest potential markets have already been produced and are either off-patent or nearing that point”. This position is collaborated by Frank (2003) and Whittaker (2003). Frank (2003) observed that of the 98 NDA applications approved by the FDA in 2000, only 27 (about 28%) were for New Molecular Entities (NME) i.e. new chemical entities that have not hitherto been tested in humans in a quest to seek drug approval. The rest of it was for “products that represent new formulations and new methods of delivering existing drugs” (Frank, 2003).

These ‘new’ drugs have variously been described as “less new” (Frank, 2003), “follower” (Rasmussen, 2003) and “me-too” (Koppal, 2004). The bottom line is they have resulted in an “innovation gap” in the pharmaceutical industry resulting in unmet medical needs remaining unmet (Whittaker, 2003).

Besides the innovation gap, there are other problems that have resulted from the blockbuster drug model. These include the emergence of “orphaned” drugs (drugs for rare diseases that are overlooked or lost in the quest for the blockbusters), and the consequences of patent protection. These are discussed in the following sections.

### *Orphan Drugs*

A direct consequence of the innovation gap is what Klein & Tabarrok (2003) have described as “drug loss”. Drug loss, simply put, is the reduction in the total number of new drugs created due to various reasons including the delay in drug approvals as well as the dogmatic quest for blockbusters (Klein & Tabarrok, 2003). The effect of the drug loss phenomenon is felt mostly in drugs for rare diseases (Klein & Tabarrok, 2003).

Orphan drugs are those designed for diseases or conditions affecting less than two hundred thousand persons in the United States at the time of designation (FDA, 1999). These diseases that fall in this category are usually classified as “rare diseases” (Klein & Tabarrok, 2003). Alternatively, these diseases or conditions can affect more than two hundred thousand persons but the costs of developing the curative drugs cannot be recouped within seven years from sales in the United States (FDA, 1999). In the quest for the blockbusters, many small-market therapies – drugs designed for sub-groups of the general population – are abandoned or orphaned (Klein & Tabarrok, 2003). The medical

needs that these could meet remain mostly unmet. The Orphan Drug Act of 1983 was promulgated to address this problem (FDA, 1999). The idea was to help alleviate the costs incurred by the manufacturers of orphaned drugs by giving them "tax breaks, subsidies, and special exclusivity privileges" as "sponsors of drugs for rare diseases" (Klein & Tabarrok, 2003). Several abuses of the Orphan Drug incentives have been reported (Richardson, 1987; Kenney, 1988). While the Orphan Drug incentives are given based on expected US sales, Klein & Tabarrok (2003) note that "worldwide sales often much exceed U.S. sales" thereby effectively negating the need for such incentives. Overall, the Orphan Drug act has not been very effective in the prevention of drug loss. Its intentions though noble have in reality been significantly more beneficial to the established U.S. drug manufacturers, resulting in a multiplication of their monopoly privileges (Arno et al., 1995; LeBlanc et al., 1996).

***Patent Protection***

If the U.S. government grants a patent to a drug, all other manufacturers are barred for a predefined number of years from producing a product of the same chemical composition, except by permission from the patent holder (Klein & Tabarrok, 2003). The patent protection term for a drug in the US is usually 17 years (FDA, 2001; Klein & Tabarrok, 2003). A patent grants a degree of monopoly power to the patent holder. When developing a new drug, pharmaceutical companies are usually anxious about the possibility that another company may also be working on the drug or similar NCE, or has received news or leaks about the promising incipient drug. They are therefore eager to obtain a patent. Companies therefore apply for and receive drug patents in advance of

final FDA approval to market the drug (Grabowski & Vernon, 1983; Klein & Tabarrok, 2003). Some of the seventeen years of patent protection is therefore dissipated waiting for approval (Grabowski & Vernon, 1983). The "effective patent life" of a new drug is the time from approval to the end of the patent (Grabowski & Vernon, 1983; Gilbert et al., 2003; Klein & Tabarrok, 2003). When a patent expires, other producers are permitted to replicate the product and to sell it as a "generic drug." This competition drives down prices (Gilbert et al., 2003; Klein & Tabarrok, 2003). The 1984 Drug Price Competition and Patent Term Restoration Act, known as the Waxman-Hatch Act further served the generic drug producers by removing some arbitrary constraints on generic drug development (Klein & Tabarrok, 2003). In general, rising prices as well as aggressive patent challenges by competitors limit the total revenue potential of the blockbuster drug (Gilbert et al., 2003)

## Bioinformatics and Pharmacogenomics

### *Bioinformatics in Drug Discovery and Development*

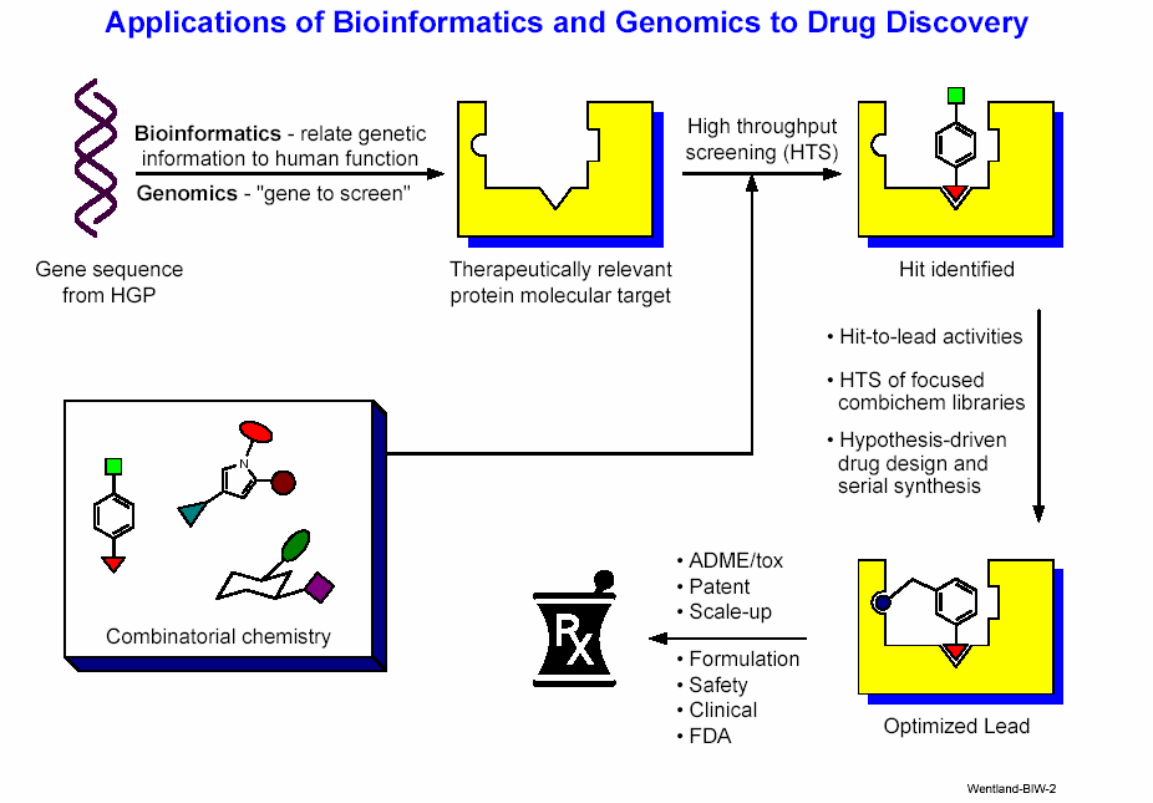
Traditionally, pharmaceutical companies have employed a cautious mostly chemistry and pharmacology-based approach to drug discovery, and are finding it “increasingly difficult to find new compounds that will lead to new drugs” (Lim, 1997; Duggan et al., 2000). In the highly competitive “winner takes all” pharmaceutical industry, the first company to patent a new chemical entity (NCE) for a particular therapy takes all the spoils, leaving other competitors to mostly wait for patent expirations to partake in the largesse (Duggan et al., 2000). Pharmaceutical companies therefore invest heavily in any processes that can accelerate any step of the drug development cycle (Lim, 1997; Van Arnum, 1998; Papanikolaw, 1999; Overby, 2001; Whittaker, 2003). The increasing pressure to generate

more drugs in less time has resulted in remarkable interest in Bioinformatics (Papanikolaw, 1999).

Although Bioinformatics attained prominence because of its leading role in the storage, management and analysis of genomic data, its focus seems to be shifting due to the need of the life sciences to exploit these data (Whittaker, 2003). According to Overby (2001), technologies grouped under the umbrella of **Bioinformatics** involve the use of computers to store, organize, generate, retrieve, analyze and share genomic, biological and chemical data for **Drug Discovery**. Several other writers have made the connection between Bioinformatics and Drug Discovery. Whittaker (2003) posited that Bioinformatics is used in "drug target identification and validation and in the development of biomarkers and toxicogenomic and pharmacogenomic tools to maximize the therapeutic benefit of drugs". Ratti & Trist, (2001) suggest that "today's [drug discovery and development] process … has been enriched by advances in technological developments in screening, synthetic chemistry, and by the increased number of possible targets due to the application of genomics and bioinformatics." The traditional chemistry and pharmacology-based approach to drug discovery has recently given way to a more modernized information-based approach – Bioinformatics (Lim, 1997). The drug discovery and development landscape has changed – for good, with the practice of Bioinformatics becoming prevalent in the drug industry such that the drug industry is one of the major players guiding the development of the Bioinformatics field (Van Arnum, 1998; Duggan et al., 2000; Attwood & Miller, 2003). Duggan et al, (2000) observed that "many (if not all) of the large pharmaceutical companies have established internal bioinformatics groups whose purpose is to beat the competition to solutions of a problem

that may give their company that crucial edge in producing the next major drug." Bioinformatics has certainly come to stay and is now ubiquitous with drug discovery. According to Pollock and Safer (2001) "few if any (drug discovery) projects are computer free". The impacts that Bioinformatics has had and continues to have in the early stages of drug development are encouraging and would only lead to further Bioinformatics investments (Ratti & Trist, 2001). Wentland (2004) provides a graphical illustration of the role of Bioinformatics in drug discovery:

Fig. 1: The Role of Bioinformatics in Drug Discovery




Source: Wentland (2004)

One of the main thrusts of current Bioinformatics methods is the finding of biologically active candidates (Whittaker, 2003). Drugs are usually only developed when the specific drug target for those drugs' actions have been identified and studied (Lim, 1997). Until recently, drug development was restricted to a small fraction of possible targets since the majority of human genes were unknown. The number of potential targets for drug development is increasing dramatically, due mainly to the genome project (Lim, 1997). Mining the human genome sequence using Bioinformatics has helped define and classify the genomic compositions of genes coding the target proteins, in addition to revealing new targets that offer potential for new drugs (Van Arnum, 1998; Southan, 2001; Foord,

2002). This is an area where the human genome information is expected to yield big payoffs (Southan, 2001; Foord, 2002). Drug developers are presented with an unaccustomed luxury of choice as more genes are identified and the drug discovery cycle becomes more data-intensive (Lim, 1997). Of the estimated 35,000 genes in the human genome, Zambrowicz & Sands (2003) contend that the 100 top-selling drugs have targeted only 43 of their encoded proteins. By enabling the identification and analysis of more and more drug targets, Bioinformatics is expected to greatly increase the breath of potential drugs in the pipelines of pharmaceutical companies (Lim, 1997; Overby, 2001; Whittaker, 2003; Zambrowicz & Sands, 2003).

After drug targets – or better still, "potential" drug targets – have been discovered, there is an invaluable need to establish a firm association between a putative target gene or target protein with the disease of interest (Whittaker, 2003). The establishment of such a key relationship provides justification for the drug development. This process, known as target validation, is an area where Bioinformatics is playing a significant role. Target validation not only helps build the case that the drug modulation of such a target will result in beneficial effects on the disease, it also helps mitigate the potential for failure in the testing and approval phases (Ratti & Trist, 2001; Gilbert et al., 2003; Whittaker, 2003).

The current high cost of drug discovery and development is a major cause for concern among pharmaceutical companies (Papanikolaw, 1999). Along with increasing productivity, pharmaceutical companies always aim to reduce the high failure rate in the

drug discovery process thereby increasing the number of drugs coming to market (Papanikolaw, 1999). The cost of clinical trials limits the number of drugs a pharmaceutical company can develop, and hence selecting the compounds with the best chances for approval is critical (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004). The costs of drug discovery and development – generally include total costs from discovery to approval (Klein & Tabarrok, 2003; PhRMA, 2003; Wierenga & Eaton, 2004) though some studies have included the costs of failed drugs and the costs for commercialization (Gilbert et al., 2003). There is also a cost associated with the elongated process, beginning from discovery all the way to final approval (Lim, 1997; Klein & Tabarrok, 2003; PhRMA, 2003). Advances in Bioinformatics have allowed for marked efficiencies, beginning with target identification and validation, to assay development, and high-throughput screening (HTS) – all with the goal of identifying new chemical entities (Belkowski, 2003). With more efficient target discovery and validation, Bioinformatics can help ensure that more drug candidates are successful during the approval process as well as shortening the discovery and development cycle, making it more cost-effective (Lim, 1997).

There are some other non-discovery/development costs – collateral costs – that plague the pharmaceutical industry. These costs include commercialization costs (Gilbert et al., 2003; Mullin, 2003), litigation and drug-recall costs (Klein & Tabarrok, 2003; Tait & Mittra, 2004), and general costs to society (Lazarou et al., 1998; Kohn et al., 1999). Commercialization costs, estimated by Gilbert et al., (2003) to be about $250 million per approved drugs, are high mainly because most “new” drugs approved are essentially

functional replicas of drugs that already exist (Koppal, 2004; Tait & Mittra, 2004). Because these mostly copycat drugs are being commercialized to abate ailments for which there are already drugs, there is a need for what Meyers & Baker (2001) describe as a high and "intense promotional noise" in their (pharmaceutical companies) efforts to attract the attention of both patients and physicians who already have access to similar medication. Bioinformatics, by enabling the more efficient discovery and identification of drug components and targets, will bridge the innovation gap, thereby allowing pharmaceutical companies the opportunity to efficiently discover and develop novel drugs and chemical entities (Whittaker, 2003). Commercialization costs are then expected to fall significantly, as drugs are commercialized not in competition with already existing equivalents, but in announcement and advertisement of new drugs that offer new therapeutic benefits for hitherto unmet medical needs (Meyers & Baker, 2001)

***Pharmacogenomics in Drug Discovery and Development***

The new knowledge of genetic biomarkers for diseases is spurring the development of pharmacogenomic-based drugs discovery and development strategies that allow pharmaceutical companies design more individualized drug regimens and dosages (Wechsler, 2004). This is accomplished by identifying genetic conditions that allow individuals to be more likely to respond to certain drugs, not respond at all or be susceptible to adverse reactions (Bernard, 2003; Whittaker, 2003; Wechsler, 2004). Several studies have cited adverse drug reactions to be one of the leading causes of death among hospitalized patients, contributing the majority of the $17 to $29 billion annual costs of medical errors (Lazarou et al., 1998; Kohn et al., 1999; Prows & Prows, 2004).

Kohn et al., (1999) also note the cost associated with the loss of confidence in the healthcare system and the diminishing satisfaction of both patients and health professionals.

Pharmacogenomics is expected to reverse this trend by enabling pharmaceutical companies to design drugs that meet the needs of pre-defined genetic sub-groups of the general population (Rioux, 2000). Pharmacogenomics cannot and does not improve the efficacy of a given drug; it simply helps in selecting patients who are likely to respond well (Rioux, 2000). The main interest is in identifying patients for whom drug efficacy can be predicted, and to spare others from avoidable adverse effects (Rioux, 2000). The promise of pharmacogenomics is that physicians may soon be able to prescribe drugs on the basis of their patients' genetic profiles (Prows & Prows, 2004; Zemlo, 2004). This would take away the guesswork in drug prescriptions, increase both physician and patient confidence and radically modify the prevailing approaches to drug discovery and development, diagnostics, therapies and disease prevention strategies (Prows & Prows, 2004; Zemlo, 2004). There is also a benefit to society as the use of expensive drugs is avoided in patients whose ailments clearly would not have been abated or cured by these drugs (Rioux, 2000).

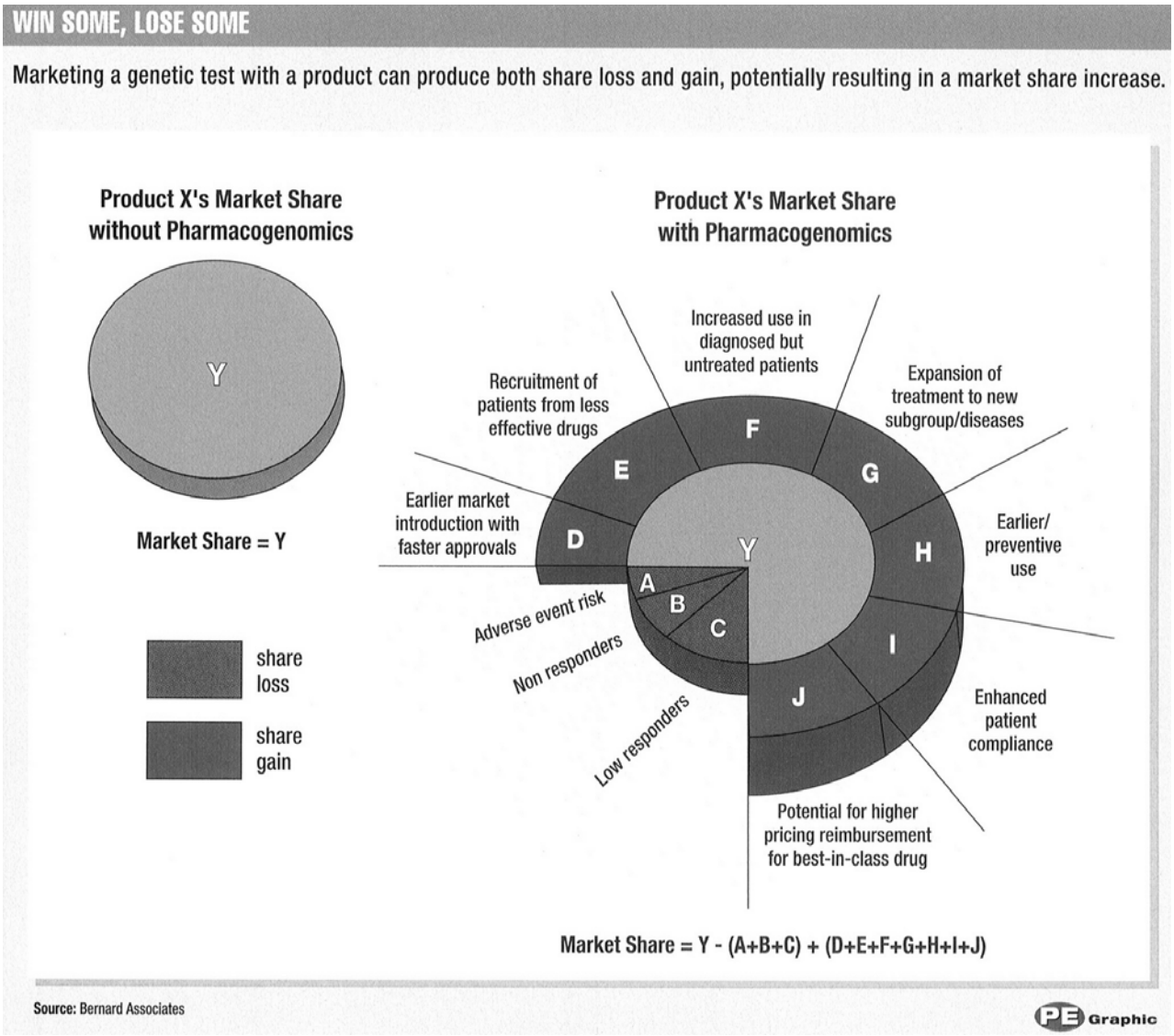
As noted earlier, adverse drug reactions in the general population often results in hospitalizations and in some cases fatalities (Lazarou et al., 1998; Kohn et al., 1999; Prows & Prows, 2004). When these drug reactions are investigated and documented, and the culprit drug identified, the result commonly is a withdrawal (or recall) of the drug

(Bernard, 2003). This is immediately followed by a torrent of lawsuits for negligence and pharmaceutical malpractice (Simons et al., 2004). A pharmacogenomic drug development strategy presents an opportunity to reverse this trend (Bernard, 2003; Zemlo, 2004). The promise of pharmacogenomics may lead to the "accelerated development of precision pharmaceuticals" (Zemlo, 2004). Precision pharmaceuticals refer to drugs (and dosages of these drugs) that are tailored to an individual's genetic composition (Bernard, 2003; Zemlo, 2004). These drugs according to Zemlo (2003) can be can be evaluated in simplified and shortened clinical trials and because of their customization will show little or no adverse effects. With drug customizations, some form of genetic testing or verification may be required prior to prescription (Bernard, 2003). This ensures that the chances for wrong prescriptions are greatly minimized (Bernard, 2003).

In the course of drug discovery and development, pharmaceutical companies stay mostly focused on the major blockbuster drugs that are prescribed to upwards of 20 million people (Klein & Tabarrok, 2003; Drews & Duyk, 2004; Tait and Mittra, 2004). The result of this is a drug loss – a loss of drugs that may have been developed to cure diseases that affect only a small number of people (Klein & Tabarrok, 2003). These drugs – or potential drugs are abandoned or orphaned ((Klein & Tabarrok, 2003). A pharmacogenomic strategy to drug development may revive these orphaned drugs if it can be demonstrated that there are potential beneficiaries for these drugs (Bernard, 2003). Similarly, Rioux (2000) has suggested that "From an economics standpoint, if the pharmaceutical company could benefit from something like orphan drug status for its product, this would help encourage stratification of populations on the basis of

pharmacogenomics, since the reduction in the size of a population to be treated could be offset by preference for the drug". This according to Rioux (2000) is the only way that pharmaceutical companies can be encouraged to forgo the blockbuster dogma.

The FDA in recent years has recognized and even encouraged Pharmacogenomic approaches to drug discovery and delivery (Rioux, 2000; Wechsler, 2004). As pharmacogenomic technologies continue to emerge and mature, the FDA as well as other international regulatory bodies is developing pharmacogenomic guidelines and regulations (Bernard, 2003). In fact, the FDA is encouraging pharmaceutical companies to conduct pharmacogenomic research and submit results to a proposed "Interdisciplinary Pharmacogenomic Review Group" – a group that isn't directly involved in the drug approval process (Bernard, 2003). Notwithstanding the increased attention and the documented potential and promise of pharmacogenomics-based drug development strategies, there has been continuous resistance to this approach on the part of pharmaceutical companies (Rioux, 2000; Bernard, 2003; Wechsler, 2004). This resistance is borne out of the perception that a pharmacogenomics strategy will lead to a significant loss of revenues resulting from the fragmentation of the drug market (Rioux, 2000; Bernard, 2003). Bernard (2003) in his article "The 5 Myths of Pharmacogenomics" lists as the 4$^{th}$ myth, the perception that pharmacogenomics will diminish the much coveted blockbuster drug model. He goes on to debunk this "myth" by demonstrating the potential market-increase impacts of pharmacogenomics. In the diagram below, Bernard (2003) shows a holistic representation of the possibilities with pharmacogenomics.




Bernard (2003) clearly indicates that a pharmacogenomic strategy may or may not reduce a drug's market size, and actually has the potential to increase it, depending on a variety of factors. For a full review of the factors illustrated in the chart above, please refer to "The 5 Myths of Pharmacogenomics" (Pharmaceutical Executive. Eugene: Oct 2003.Vol.23, Iss. 10; pg. 70)

# CHAPTER 6

## *DISCUSSION AND FUTURE WORK*

This chapter is a review of the stated purpose, and a discussion of the results reported in previous chapters.

In summary, the stated aims of this thesis are as follows:

- To review the historical trend and evolution of the drug discovery and development process
- To review the current model of cost evaluation of the drug discovery and development process.
- To understand the propositions of a Bioinformatics and Pharmacogenomic approach to tailored and targeted drug discovery
- To review the differences between the traditional approach and the Bioinformatics approach to drug discovery and development

The findings of the literature review have been presented in the previous chapter.

A review of the history of drug discovery and development reveals a process that has greatly evolved over time. Papanikolaw (1999), Ratti & Trist (2001) and Boa (2003) describe a process driven – up till the mid 20$^{th}$ century – mainly by trial and error, intuition and serendipity of chemists, biologists and physicians. During the second half of the 20$^{th}$ century, the pharmacological bases of drugs and diseases were beginning to be defined, leading to a more rational linear drug development process. *In vitro* methods (experiments carried out using animal tissues, but in an artificially simulated environment) gradually replaced the more invasive *in vivo* methods allowing for a higher

economy of scale in experimentation. Drug discovery today has generally followed this trend, though a significant amount of emphasis is placed not only on the efficacy, but also on the safety of the discovered and developed drug. The drug discovery process today is very heavily regulated, mainly to ensure the safety and protection of the general population. A major obstacle to the goal of drug safety is the inherent genetic variations in the general populations. Rioux (2000) has described this variation – generally referred to as genetic polymorphisms – as a "stable difference in DNA sequence at the same locus (a specific position in the genome) among individuals". This difference is essentially what makes very few drugs to be effective for everyone. Hence, in the U.S., despite the elaborate testing and approval process as instituted by the FDA, Lazarou et al. (1998), Klein & Tabarrok (2003), Evans et al. (1999, 2004) and Prows & Prows, (2004) report adverse effects from supposedly "safe" FDA approved drugs – adverse effects that sometimes lead to fatalities. Mahlich et al, (2001) and Klein & Tabarrok, (2003) observe that the development, testing and approval process has increased from about 8 years in 1960 to approximately 15 years in 1996 due mainly to the more stringent safety requirements imposed by the FDA during toxicity testing.

There appears to be a conflict or disconnect between the full or intensive application of Bioinformatics in drug discovery and development and the expectations from the regulatory bodies. Though the FDA has – in principle, encouraged and in fact promoted pharmacogenomic testing during the discovery and development process, the regulatory machinery has not been put in place to fully support such initiatives. Primarily, pharmaceutical companies may benefit from a shortened testing and approval process to allow for their drugs to enter the market quicker. The shortened length of the approval

process can be justified by the fact that the target market for the concerned drug has been clearly and specifically defined along the lines of genetic composition. In the documented evolution of the drug discovery and development process, the position of this thesis is that the regulations need to be reformed to allow the full integration of the advanced Bioinformatics technologies in the drug discovery process. When this happens, the concept of personalized drugs may transform from myth to reality.

The variations in the genetic composition among individuals in the general population was further revealed by the findings of the Human Genome Project – a project funded by the U.S. Government with the primary charge to research and document the genetic composition of the human being. The results from the Human Genome Project (HGP) further accentuated the knowledge of the differences in nucleotide sequences of individual DNA, and the effects of such differences. This newfound knowledge further elucidated the genetic bases of health and disease conditions.

With the partial completion of the HGP in April 2003, Bioinformatics received a high amount of recognition in the scientific community. While it initially attained prominence because of its central role in the genome data storage, Whittaker (2003) notes that Bioinformatics has shifted focus as the life sciences attempt to exploit the data generated from the HGP. The pharmaceutical industry is perhaps the earliest and most ardent adopter of Bioinformatics and related technologies. Two chief challenges facing the pharmaceutical industry were identified in this thesis. They are:

- The high cost of drug discovery and development, arising mainly from the extensive and time consuming trial and approval process, and

- The innovation gap - arising mainly in the quest to produce widely applicable blockbuster drugs – drugs that address needs for which there are already therapies.

The high and rising cost of drug discovery and development has been highlighted in this thesis. Pharmaceutical companies, in competition continue to invest more and more resources in their respective bids to develop the next commercial blockbuster. A significant portion of the high cost is as a direct result of the time and resource intensive trials and approval process.

The high cost of drugs is passed on to society, the end users of the drugs, and can have very far-reaching ripple effects on the society. Though the costing of the drug discovery and development process is a controversial subject, the effects of the high costs of the end product are often clear-cut. For example, a high cost for newly approved drugs would result in an increase in the cost of insurance to cover these drugs, in cases where these drugs must be prescribed and consumed. An increase in any insurance costs can potentially affect the attitudes of employers towards employees. Some analyses have shown that the high cost of insurance can translate to the high cost for employers who often are required to provide such benefits to their employees. In order to avoid incurring high costs for benefits such as employee insurance, employers have been known to resort to out-sourcing jobs to countries where they do not have to provide such benefits. Bioinformatics is expected to contribute to cutting down the costs of drug discovery and development, therefore resulting in cheaper medication and by implication, more affordable health insurance. Burke et al, (2004) and Prows & Prows (2004) note the high interest expressed by health insurance companies to a pharmacogenomic approach to

drug discovery, and suggest this to be and indication that they (the health insurance companies) believe this approach may lead to an overall decrease in the cost of healthcare. This perspective needs to be emphasized. The position of this thesis is that the cost-savings effect of Bioinformatics and related technologies need to be further and fully explored. The cost of drugs and healthcare is a very serious topic of public and political debate. This is ample justification for future work in this area.

Another effect of the lengthy drug discovery and development process is the relatively quick expiration of patent protections, and the introduction of generic drugs.
In the pharmaceutical industry, patents have to be taken out at a fairly early stage of drug development. This means that every additional day spent in drug development is a day lost in revenues under patent protection. In the U.S. patent protection for new chemical entities are usually given for a period of 17 years. Depending on when a pharmaceutical company applies for and receives patent protection, the 15-year (average) drug discovery and development period can erode much of the protection such that the effective patent protection time – the period from when the drug is approved and marketed to the end period of the patent – is significantly shortened. The patent protection is meant to give the original manufacturer a market exclusivity to enable such a company substantially recoup the high development costs and benefit from its “intellectual property”. The reviewed literature did not sufficiently address the impact that Bioinformatics and Pharmacogenomics can have on the patent protection that pharmaceutical companies receive. This subject matter is one that needs to be evaluated further. With the implementation of a Bioinformatics and Pharmacogenomics strategy, it is expected that

the drug development time frame and costs will be significantly reduced. If pharmaceutical companies are spending less time and money to develop drugs, perhaps the patent protection time periods ought to be reduced to allow quicker introduction of generic equivalents. While proponents may cite the advantages of competition, as manufacturers of generics are opportune to introduce their cheaper brand-name equivalents sooner, giving consumers the opportunity to select cheaper drugs without losing on efficacy. The stringent testing requirements for new drugs are mostly waived for the generic equivalents making them even cheaper to develop and market. They usually have the exact same chemical composition as the brand-name versions and are favored by many health insurance providers. Opponents will decry the shortening of an already "short" patent protection period. Besides, they will likely propagate the notion that the manufacturers of generics are benefiting unduly from the expense incurred by the original drug manufacturer.

Bioinformatics and Pharmacogenomics are also expected to close the innovation gap leading to a wider array of truly new drugs. A good indication of the innovation gap is the availability of 3 high-profile drugs for erectile dysfunction (ED). When Pfizer in 1998 introduced Viagra as a cure for ED, it was geared to capture and dominate a huge market of an estimated 30 million sufferers in the U.S. alone. After 5 years of market monopoly, two rivals were introduced in quick succession. GlaxoSmithKline PLC and Bayer AG introduced Levitra in August of 2003, with Eli Lilly and Icos Corp. following suite with Cialis in November of 2003. Viagra, Cialis and Levitra each are effective on 70 percent to 80 percent of men, but at different speeds. All three drugs work similarly – they block

an enzyme called PDE-5, which helps relax the penis and allow the blood flow needed to produce an erection. The rival companies have had to spend millions of dollars to market the differences among their drugs. Cialis rang up U.S. sales of $203 million in its first full year and had a $165 million advertisement expenditure. Levitra sold just $128 million worth of pills, well below what the manufacturers spent on TV, print and other media. Though there are differences in the ways that these drugs operate, they essentially cure the same condition leading to the same end results. The gargantuan advertisement budgets of these "me too" drugs is not yielding the desired results. With the introduction of Bioinformatics, literarily thousands of hitherto unknown drug targets have been identified in the human genome. Drugs can be developed to address very specific diseases based on the genetic composition of its sufferers. This thesis foresees the result to be the diversification of pharmaceutical companies' drug profiles and a movement away from the blockbuster approach.

The emergence of Bioinformatics and related technologies was widely believed to be the precursor for the age of targeted and tailored drug development. The testing and approval process would be shortened and the failure rate significantly reduced as drugs are designed for specific sub-sets of the population who have been identified and tested. Several writers alluded to the potentials of Bioinformatics and projected the development of subpopulation-specific drugs – drugs that would be tailored to meet the needs of sub-groups, stratified according to their genetic compositions. More specifically, Pharmacogenomics, as projected, would combine pharmacology with Bioinformatics to yield what Zemlo (2004) described as "precision pharmaceuticals". The development for

drugs for sub-groups of the general population has generated concern among pharmaceutical companies as this signifies a shift from the current practice. The coveted blockbuster drugs strategy is perceived to be at risk with the emergence of sub-population specific drugs. Bernard (2003) however described a system where the introduction of Bioinformatics would potentially increase the market share of a pharmaceutical product. For as much attention that Bioinformatics has received in recent years, the impact has been, in the most part, slow in manifesting. Lindpainter (2003) observed that "…the interface between genetics/genomics and pharmacology, pharmacogenetics and pharmacogenomics…. are commonly touted as heralding a 'revolution' in medicine, yet as soon as one begins to probe more carefully, little substance is yet to be found to support these enthusiastic claims". Most of the enthusiastic projections remain just that – enthusiastic projections.

There are a handful of exceptions. One outstanding example is the Genetech drug, Herceptin (trastuzumab), the product approved by the FDA in September 1998 for use in women with metastatic breast cancer who have tumors that overexpress the HER-2 protein. This represents only about 25% of women with breast cancer. In the course of the testing and approval process, it was recognized that the drug did not meet the needs of the general population. It was therefore doomed for failure. Buoyed by public advocacy, Genetech saved the "failing" drug by coupling it with a pharmacogenomic test – HercepTest, to identify potential responders. The approval of this drug for marketing is an indication that pharmacogenomic approaches to drug discovery and delivery are being recognized. It is a striking example of how identifying patient-population subsets can bring a new measure of safety and efficacy.

As observed earlier, this measure needs to be expanded across the board to allow for the full integration of Bioinformatics and related technologies into the drug discovery and development process.

Conversely, there are several cases where the genetic differences in the general population have negatively impacted a drug's commercial success. Bayer's cholesterol agent Baycol, Wyeth's appetite suppressant Redux, GlaxoSmithKline's oral diabetes agent Rezulin and more recently Merck's arthritis drug Vioxx are some of the very recent drug withdrawals. Mainly because these drugs have been designed, developed and marketed for the general population, adverse reactions from a very small section of the population would sometimes lead to the withdrawal of a drug. As an example, the FDA in 1998 forced Hoechst Marion Roussel (now Aventis) to withdraw its $600 million-a-year anti-allergy drug Seldane (terfenadine) from the market because less than 0.5 percent of the population has a variant of the CYP3A gene that makes them unable to metabolize the drug in the presence of erythromycin, an antibiotic. These people usually suffered severe cardiotoxicity. A pharmacogenomic test, as was the case of Herceptin, may have kept the drug in the market.

Besides limiting the incidents of drug withdrawals, a pharmacogenomic approach will also mitigate the incidence of orphaned drugs. Most drugs become orphaned during the testing and approval process when it becomes apparent – thanks to the non-responders in the general population – that they cannot be effective or in many cases injurious to certain groups of the population. If the responder groups are identified in the same way,

these drugs could be evaluated and approved following simplified clinical trials, as the identified responders would show very little or no side effects and/or unwanted complications.

Bioinformatics and Pharmacogenomics have more potential than the current results have shown. McHale (2003) attested to this when he noted that Bioinformatics has not yet begun to have the impact on discovery that it can have. He cited the reasons to include a lack of standards, ontology and technological integration. The lack of standards and ontology was particularly apparent during the review of literature for this thesis. There are almost as many definitions of Bioinformatics (and related technologies) as there are Bioinformatician (also referred to as Bioinformaticists). There are often divergent views of the discipline and an absence of a hierarchical structuring of knowledge in the genre. Though this could be associated with the fact that the discipline is still relatively new in scientific literature, there is an urgent need to ensure that practitioners are "speaking the same language". This will allow for better collaboration and integration with other disciplines. In fact, McHale estimated that a better defined and organized Bioinformatics combined with Chemical Informatics (Cheminformatics) will yield an integrated drug discovery informatics solution that could lead to an estimated $282 million potential savings per drug.

Finally, Bioinformatics and Pharmacogenomics are the latest facets of the continuous evolution of the drug discovery and development process. Their adoption in the drug discovery and development process is synonymous with the future survival of the major

pharmaceutical companies. According to Zemlo (2004), Pharmacogenomics should be considered as the gateway by which 21st century medicine is ushered into the present. This gateway however is only wide enough to allow the harbingers of change.

## Acknowledgements

**References:**


1. Altman, R.B., (2000): The interactions between clinical informatics and bioinformatics: a case study, *Stanford University*
2. Annemans et al. (2002): A Health Economic Model to Access the long-Term Effects and Cost-Effectiveness of Orlistat in Obese Type 2 Diabetic Patients. *DIABETES CARE, Volume 25, Number 2. February 2002*
3. Anonymous (see pg 39.)
4. Arno, P. S., Bonuck, K., & Davis, M. (1995). Rare Diseases, Drug Development, and AIDS: The Impact of the Orphan Drug Act. *Milbank Quarterly* 73, no. 2: 231–52.
5. Attwood, T.K., and Parry-Smith, D.J., (1999): Introduction to Bioinformatics. *Prentice Hall*
6. Attwood, T.K., and Miller, C.J. (2003): Bioinformatics Goes Back to the Future. *NATURE REVIEWS | MOLECULAR CELL BIOLOGY*. Vol. 4; Feb 2003; pgs. 157-162
7. Belkowski, Stanley. (2003): Functional informatics: Drug Discovery and Development. *Highland Ranch, Oct. 2003*
8. Bell, John (2004). Predicting disease using genomics. Nature | vol 429 | 27 may, 2004 | www.nature.com/nature
9. Bernard, Stan (2003). The 5 Myths of Pharmacogenomics. Pharmaceutical Executive. Eugene: Oct 2003.Vol.23, Iss. 10; pg. 70

*10.* Boa, A. N. (2003). Introduction to Drug Discovery. *Department of Chemistry, University of Hull*

11. Boguski, M. S. (1998). Trends Guide to Bioinformatics. *Elseiver*

12. Boswell, Clay. (2002): The future of pharmaceuticals is in targeted treatment, says IBM. *Chemical Market Reporter;* Dec 16-30, 2002; 262, 22; ABI/INFORM Global pg. 12

13. Brown, D. R. (2000). A Review of Bio-Economic Models. Cornell University

14. Burke, Wylie & Haga, Susanne B. (2004). Using Pharmacogenomics to Improve Dug Safety and Efficacy. *Journal of American Medical Association (JAMA).* Reprinted JAMA June 16 2004 – Vol. 291 No. 23

15. Chitty, Mary (2003): Overviews: Genomics & proteomics for Drug discovery & development and Molecular Medicine. *Cambridge Healthtech Institute (CHI).* December 2003.

16. Counsell, Damian. (2004) Bioinformatics Frequently Asked Questions.

17. DiMasi, J.A., Hansen, R.W., Grabowski, H.G. (2003): The price of innovation: new estimates of drug development costs. Journal of Health Economics. vol. 22, no. 2, March 2003, pp. 151-85

18. Drews J., & Duyk G. (2004). Fixing the Drugs Pipeline. Economist; 3/13/2004, Vol. 370 Issue 8366, p37, 2p, 1 graph, 1c

19. Drummond, M. F., O'Brien, B., Stoddart, G. L., Torrance, G. W. (1997): Methods for the Evaluation of Health Care Programmes. *Oxford University Press*

20. Dryden, Jim. (1998). Researchers Discover New Genetic Risk for Alzheimer's Disease

21. Duggan G., Fejes A., Ingram C., Salazar M., Sirski M. (2000): Drug Discovery in the Bioinformatics Age

22. Durbin R., Eddy S., Krogh A., & Mitchison G. (1998). Biological Sequence Analysis. *Cambridge University Press*

*23.* Evans Williams E. & McLeod H.L. (2003). Pharmacogenomics – Drug Dispositions, Drug Targets and Side Effects. New England Journal of Medicine, pgs 348, 538-549 (2003)

*24.* Evans William E. & Relling Mary V. (2004). Moving Towards Individualized Medicine with pharmacogenomics. *Nature **429**, 464 - 468 (27 May 2004); doi:10.1038/nature02626*

25. First Genetic Trust. (2000): Individual Privacy and Security Requirements for Genetic Research and Personalized Care. *First Genetic Trust*

26. Food and Drug Administration (FDA). (1999). Orphan Drug Act (as amended). http://www.fda.gov/orphan/oda.htm

27. Food and Drug Administration (FDA). C.D.E.R., (2001): Report to the Nation: Improving Health Through Human Drugs. FDA, Rockville, MD

*28.* Foord, S.M. (2002). Receptor Classification: Post genome. *Curr. Opin. Pharmacol. 2, 561-566*

29. Fox, Joanne. (2003). The Human Genome Project: The impact of genome sequencing technology on human health. *BioTeach Reviews and Readings*

30. Frank, Richard G. (2003). New Estimates of Drug Development Costs. Journal of Health Economics 22 (2003) 325-330

31. Gibas, Cynthia and Jambeck, Per. (2001). Developing Bioinformatics Computing Skills. *O'Reilly Associates*

32. Goldstein, D., Tate, S.K., Sisodiya, S.M. (2003). Pharmacogenetics goes Genomic. *Nature Reviews Genetics* **4**, 937 -947 (2003); doi:10.1038/nrg1229. **December 2003 Vol 4 No 12**

33. Gilbert, J. Henske, P. & Singh, A. (2003). Rebuilding Big Pharma's Business Model. IN VIVO Business and Medicine Report. Vol 21 #10

34. Grabowski, Henry G. & Vernon, John M., (1983). *The Regulation of Pharmaceuticals: Balancing the Benefits and Risks*. Washington D.C.: American Enterprise Institute for Public Policy Research.

35. Hall, Stephen S. (2003): Personalized Medicine's Bitter Pill. *Technology Review;* Feb 2003; 106, 1; ABI/INFORM Global pg. 62

36. Henriksson, Freddie. (2002). Applications of Economic Models in Healthcare. *Pharmacogenomics 2002: 20 Suppl. 1: 43-53*

37. Higgins, Des & Taylor, Willie. (2000). Bioinformatics Sequence, Structure and Databanks. *Oxford University Press*

38. Hofestadt, R. (2002). Bioinformatics. 2002 Yearbook of Medical Informatics. *Schattauer*

*39.* Jones, A. M. & O'Donnell, O. (2002) Econometric Analysis of Health Data. *New York: John Wiley*

40. Keefer, Robert C. (2003). What Ails the Pharmaceutical Market? Competition, Costs and Collaboration Ills. Vignette Publications. http://www.vignette.com/Downloads/WP_PH1_WhatAilsPharma.pdf

41. Kennedy, Dennis. (2003). Bioinformatics Law Resources

42. Klein, D.B. & Tabarrok, A., (2003). The Drug Discovery, Development and Approval process. www.fdareview.org

43. Kohn, Linda T., Janet M. Corrigan, and Molla S. Donaldson, eds. (1999). *To Err Is Human: Building a Safer Health System*. Institute of Medicine (IOM) Washington, D.C.: National Academy Press.

*44.* Koppal, Tanuja (2002). Genomic tools step up flow of drug pipeline. Drug Discovery & Development. *Highland Ranch: Dec 2002. Vol. 5, Iss. 11; pg. 28, 5 pgs*

45. Koppal, Tanuja (2004). Playing it Safe. Drug Discovery & Development. *Highland Ranch: Nov 2004. Vol. 7, Iss. 11; pg. 9, 1 pgs*

46. Kumar, Saran. (2002): Post-genomic drug development: Is pharma ready? *Pharmaceutical Executive;* Mar 2002; 22, 3; ABI/INFORM Global pg. 98

47. Lanfear, Jerry. (2001). Industry Opportunities: Pharmaceuticals and Designer Materials

48. Lazarou, J., Pomeranz B.H., & Corey P.N. (1998). Incidence of Adverse Drug Reactions in Hospitalized Patients. *Journal of the American Medical Association* (JAMA) 279, no. 15: 1200–1205.

49. LeBlanc, S., & Sabados, R. (1996). Oxandrolone Background. *Bay Area Reporter*, July 11.

50. Lim, H.A., (1997). Bioinformatics and Cheminformatics in the Drug Discovery Cycle. In: Lecture Notes in computer Science 1278, Bioinformatics, Ralf Hofestädt, Thomas Lengauer, Markus Löffler, and Dietmer Schomburg (eds.), (Springer-Verlag, Berlin, 1997), pp. 30–43

51. Lindpaintner, Klaus. (2003). The Impact of Pharmacogenetics and Pharmacogenomics. *Journal of Commercial Biotechnology*; Sept 2003; 10, 1; *ABI/INFORM Global pg 60.*

52. Luscombe, N. M, Greenbaum, D., Gernstein, M. (2001). What is Bioinformatics? A Proposed Definition and Overview of the field. *Schattauer GmbH*

53. Mahlich, Jörg & Roediger-Schluga, Thomas. (2001). The Changing Nature of Pharmaceutical R&D – Opportunities for Asia?

54. Malorye, Branca. (2003). Beyond the Blueprint: 50th Anniversary of the Double Helix. *Bio-IT World April 2003*

55. Mathews, Ann W. (2003). FDA Will Issue Rules on New Era of "Personalized Medicine". **Wall Street Journal**. (Eastern edition). New York, N.Y.: Nov 3, 2003. pg. B.1

56. McCubbin, Caroline. (2003). Legal issues in Bioinformatics. **Journal of Commercial Biotechnology**. London: Mar 2003.Vol.9, Iss. 3; pg. 249

*57.* McHale, Phil. (2003). Can Informatics Really Save Drug Discovery? *IUPUI Guest Lecture. Summer 2003*

58. Mendible, Juan. (2003): The Promises of the Human Genome Project. Life and Molecular Medicine (http://www.suite101.com/article.cfm/molecular_medicine/103424) September 2003.

59. Meyers, Scott & Baker, Ann. (2001). Drug Discovery – an operating model for a new era. *http://biotech.nature.com* August 2001, Volume 19 nature biotechnology

60. Michalowski, Stefan. (2001): The Genomics Revolution. Organization for *Economic Cooperation and Development (OECD) Observer*. Paris: Summer 2001. , Iss. 226/227; pg. 65, 3 pgs

*61.* Miller, Perry L. (2002). Opportunities at the Intersection of Bioinformatics and Health Informatics: a case study. *2002 Yearbook of Medical Informatics: Schattauer*

***62.*** Mirasol, Feliza (2004). Orphan Drugs Open Door to Opportunity in Small Markets. **Chemical Market Reporter**. New York: Mar 8, 2004.Vol.265, Iss. 10; pg. 8, 1 pgs

*63.* Mullin, Rick. (2003). Drug Development Costs About $1.7 Billion. *Chemical and Engineering News April 2003*

64. **National Center for Biotechnology Information** (NCBI) April, 2003: Building a Genomic Information Infrastructure at NCBI,

*65.* Nuffield Council on Bioethics (2003): Pharmacogenetics and Ethical Issues. Henry Ling Limited. The Dorset Press. 2003

*66.* Nussbaum, R.L., McInnes, R.R. and Willard, H.F. (2001): Genetics in Medicine. *6th Edition pgs 18-27. Thompson and Thompson*

*67.* Nightingale, Paul. (2000). Economies of scale in experimentation: Knowledge and technology in pharmaceutical R&D. *Oxford: Jun 2000. Vol. 9, Iss. 2; pg. 315*

68. Obenshain, Andrew. (2004). Targeted Drug Development and the Commercialization Process. Kellogg Biotech Conference.

69. Overby, Stephanie. (2001). Drug Companies on Speed. **CIO**. Framingham: Oct 15, 2001. Vol. 15, Iss. 2; pg. 114

70. Papanikolaw, Jim. (1999). Bioinformatics Emerges as Key Technology for Developing New Drugs. *Chemical Market Reporter; May 24, 1999; 255, 21; ABI/INFORM Global pg 22*

71. Pendell, Judyth (2003). Pharmaceutical Liability Study – Report on Findings. AEI-Brookings Joint Center. Harris Interactive Market Research (www.harrisinteractive.com) July, 2003.

72. Pollock, Sarah and Safer, M. Hershel. (2001). Bioinformatics in the Drug Discovery Process.

73. Pratap, Ravindram. (2004). Clinical Trials and Good Practices. **Businessline**. Chennai: Nov 2, 2004. pg. 1

74. Prows Cynthia A. & Prows Daniel R. (2004). Medication Selection By Genotype – how genetics is changing drug prescription and efficacy. AJN. May 2004. Vol. 104. No. 5.

75. *Public Citizen* (2002). America's Other Drug Problem: A Briefing Book on the Rx Drug Debate. *Public Citizen, Washington DC.*

76. Rasmussen, Bruce (2003). Aspects of the Pharmaceutical Business Mode: Implications for Australia. Center for Strategic Economic Studies, Victoria University of Technology.

77. Ratti, Emiliangelo & Trist, David. (2001). Continuing Evolution of the Drug Discovery Process. *Pure Appl. Chem., Vol. 73 No. 1, pp. 67-75*

78. Rioux, Patrice P. (2000). Clinical Trials in Pharmacogenetics and Pharmacogenomics: Methods and Applications. *American Journal of Health System Pharmacy. Vol. 57(9) May 1, 2000. Pgs 887-901.*

79. Rogoski, Richard. (2003). Technology: Clinical Trials Drive Up Drug Development Costs. *IN DEPTH: BIOTECH JOURNAL*

80. Samudhram, A. (2003). Advancing Bioinformatics Research. Computimes Malaysia. *New York: Nov 21, 2002. pg. 1*

*81.* Simmons, John., Stipp, David., and Parloff, Roger. (2004). Will Merck Survive Vioxx? **Fortune**. New York: Nov 1, 2004.Vol.150, Iss. 9; pg. 90, 9 pgs

*82.* Smith, Thomas M, Emmeluth, Donald M. (2002). Introducing Bioinformatics into the Biology Curriculum: Exploring the National Center for Biotechnology Information. American Biology Teacher; v64 n2 p93-99 Feb 2002

*83.* Southan, C. (2001). A genomic perspective on human proteases. *FEBS Lett. 798 214-218*

*84.* Southwick, Karen. (2001). The good life. Forbes. Sept 10 2001. NY Page 70

*85.* Stankus, Tony. (2003). The Promise of Proteomics. *The Alert Collector – Diana Zabel, Editor.*

86. Suzuki, D., Knudtson, P. (1998). Genethics: The Clash Between the New Genetics and Human Values. *Harvard University Press, Boston 2003.*

87. Swiss-America Business Council (SABC). (2002). Competence Center for Biotechnology, Biomedicine and Bioinformatics, December 2002.

*88.* Tait, Joyce and Mittra, James (2004). Complexity and Innovation in the Pharmaceutical Industry. Innogen policy Brief. Chemistry and Industry, In Press

89. Tsai Y.J., & Hoyme H.E. (2002). Pharmacogenomics: The Future of Drug Therapy. *Clin Genet 2002: 52: 257-264. Copyright Blackwell Munksgaard, 2002*.

*90.* US National Library of Medicine (2004). Genetics Home Reference; Your guide to understanding genetic conditions. October 8, 2004. http://ghr.nlm.nih.gov

*91.* Van Arnum, Patricia. (1998): Bioinformatics. *Chemical Market Reporter*; *Mar 9, 1998; 253, 10; ABI/INFORM Global pg FR19*

92. Verbeek, M. (2000) A Guide to Modern Econometrics. *New York: John Wiley*

93. Wechsler, Jill (2001): Genomic Discoveries Challenge Drug Manufacturers. *Pharmaceutical Technology North America*. Duluth: April 2001, Vol. 25, Iss. 4; pg. 12, 5 pgs

94. Wechsler, Jill (2004): Drug Development Linked More Closely to Diagnostics. *Pharmaceutical Technology; Oct 2004; 28, 10; ABI/INFORM Global pg 28.*

95. Whittaker, Paul. (2003) What is the Relevance of Bioinformatics to Pharmacology? *Trends in Pharmacological Sciences. Vol. 24 No.8 August 2003*

96. Wierenga, Dale E & Eaton, Robert C. (2004). DRUG DEVELOPMENT AND APPROVAL PROCESS. Alliance Pharmaceutical Company. http://www.allp.com/drug_dev.htm

97. Wiggins, Steven N. (1987): The cost of developing a new drug. *Washington, DC: Pharmaceutical Manufacturers Association*

***98.*** Wooldridge, J. M. (2003) Introductory Econometrics (2nd edition). *Madison, OH: South-Western Publishing.*

***99.*** World Health Organization (WHO) (2002). GENOMICS AND WORLD HEALTH. WHO, Geneva, 2002.

***100.*** Zambrowicz , B.P. & Sands, A.T. (2003). Knockouts model the 100 best-selling drugs – will they model the next 100? *Nat. Rev. Drug. Dicov*. 2, 38-51

101. Zemlo, Tara. (2004). Pharmacogenomic Promise, Pitfalls. **Drug Discovery & Development**. Highland Ranch: May 2004.Vol.7, Iss. 5; pg. 13, 1 pgs