# Probabilistic Modeling of the Drug Development Domain: A Bayesian Domain-Knowledge Application for Pharmacovigilance

By
Asher Daniel Schachter
MD, University Of Toronto, 1992
MMSc, Harvard-MIT Division of Health Sciences and Technology, 2001

Submitted to the Harvard-MIT Division of Health Sciences and Technology in Partial Fulfillment of the Requirements for Degree of

Master Of Science in Medical Informatics

At the Massachusetts Institute of Technology
Cambridge, Massachusetts

May 2003

[June 2003]



Signature of Author: _____

Asher Daniel Schachter
Harvard-MIT Division of Health Sciences and Technology
May 22, 2003

Certified By: _____

Isaac S. Kohane
Director, Children's Hospital Informatics Program
Thesis Supervisor

Accepted By: _____

Martha L. Gray
J W Kieckhefer Professor of Electrical Engineering
Co-Director, Harvard-MIT Division of Health Sciences and Technology

**Table of Contents**

**ACKNOWLEDGEMENTS**

I wish to extend my gratitude to the following people, for their advice, support and guidance: Isaac Kohane, Marco Ramoni, William Harmon, Robert Rubin, Alan Moses, Peter Szolovits, and Stan Finkelstein. I am also forever thankful for the unwavering support of my wife Leora and my children Noah and Audrey, to whom this work is dedicated. This thesis was supported by NIH grant K23 RR16080.

**ABBREVIATIONS**

BBN: Bayesian belief network

CPT: conditional probability table

IND: Investigational new drug

NDA: New drug application

NCE: New chemical entity

TCSDD: Tufts Center for the Study of Drug Development

TI: Therapeutic index

TI_Disease: minimum therapeutic index for a disease-related undesired effect

TI_Vital: minimum therapeutic index for an undesired effect related to a vital organ

Probabilistic Modeling of the Drug Development Domain:
A Bayesian Domain-Knowledge Application for Pharmacovigilance
By Asher Daniel Schachter
Submitted to the Harvard-MIT Division of Health Science and Technology on May 15, 2003 in Partial Fulfillment of the Requirements for Degree of Master Of Science in Medical Informatics

**ABSTRACT**

A recent analysis by the Tufts Center for the Study of Drug Development estimates that the cost of developing a single new chemical entity (NCE) into a successful therapeutic agent is $802 million. This figure is largely dependent on the expense of investigating NCEs that ultimately fail to be approved for use: between 70 – 90% of NCEs do not achieve New Drug Application (NDA) approval, and many of these failures are identified during the later, more costly phases of drug development. The exponential growth in the number of putative NCEs as a result of combinatorial chemistry and high-throughput screening has only confounded this problem by significantly increasing the number of early-phase NCEs under consideration for further costly development in human clinical trials.

It is widely agreed upon that there are 3 major categories of reasons for drug failure: safety (toxicity), efficacy, and economics. This thesis is concerned with developing a Bayesian domain-knowledge probabilistic model (called Pharminator) to address the first two of these categories, with a goal of predicting clinical success of an NCE. Pharmacoeconomic modeling is a vastly different domain compared to Pharminator's clinical trial domain, and is beyond the scope of this thesis. While several clinical predictive models have been described in the literature over the past 10 years, the ongoing costly failure rate in drug development warrants developing more reliable predictors of NCE clinical success. The number of NDA approvals in 2002 fell to a 5-year low of 18, compared to 30, 35, 27, and 24 in 1998, 1999, 2000, and 2001 respectively, despite rapidly increasing numbers of NCEs as a result of high-throughput screening and combinatorial chemistry. Therefore, previous decision models have had no apparent impact on this problem. The Pharminator model combines knowledge of drug development logistics, existing data on NCE attrition rates, and Bayesian decision theory in a manner that may improve upon the performance of previously described models.

The product of this model is an application to be used by drug development teams at the Phase I/Phase IIa time point for a given NCE that has passed the FDA Investigational New Drug (IND) screening process. The users are prompted to answer several key questions about the NCE. The answers provide information regarding which prior probabilities and conditional probability tables are to be used in the model, as well as the observed data upon which prediction will be made. The output is a numeric and graphical (distribution plot) report of the prior and posterior probability distributions for clinical success, safety and efficacy. The application is demonstrated on one fictional agent and one real agent designed to demonstrate key behaviors of the model. Retrospective and prospective testing and validation will continue beyond the completion of this thesis in order to optimize the performance of this model.

Thesis Supervisor: Isaac S. Kohane, MD, PhD
Title: Director, Children's Hospital Informatics Program, and Associate Professor of Pediatrics, Harvard Medical School

## A Background

### A1. New Chemical Entities: Selection and Development

The USA is known as a world-leader in innovation. The drug development domain is an excellent example of America's innovative potential, with many breakthrough medications having been discovered and developed in the USA. This degree of innovation requires consistently huge research and development expenses, and much of this cost is borne by patients and their insurance plans. A recent analysis by the Tufts Center for the Study of Drug Development estimates that the cost of developing a single new chemical entity (NCE) into a successful therapeutic agent is $802 million (in 2000 dollars), with clinical phase costs of $467 million, and taking into account "time costs" related to the length of time from Investigational New Drug (IND) approval to New Drug Application (NDA) marketing approval [1]. The $802 million figure is crucially dependent on the proportion of NCEs that fail during the clinical trial development phase [2,3]. Confounding this problem is the relatively recent adoption of combinatorial chemistry and high-throughput screening for potential NCEs, significantly increasing the number of early-phase NCEs under consideration for further costly development in human clinical trials. Despite the recent explosion of potential new drugs, the annual rate of NDA approval hit a 5-year low in 2002, with only 18 NDA approvals, compared to 30, 35, 27, and 24 in 1998, 1999, 2000, and 2001 respectively [4]. The only recent improvements in the drug development process are the decreases in mean residence time (the time between IND and NDA approval) by 1.5 years and in median time to research abandonment by 0.8 years, suggesting that drug developers are making faster decisions regarding research failures [5].

The drug development process consists of several phases and milestones: pre-clinical studies, IND approval, clinical trial phases I/II/III, NDA approval, and phase IV (post-marketing surveillance for idiosyncratic adverse events and potential alternate indications). The patent life of a given NCE typically begins at the time of IND approval and lasts for 20 years, but financial

return does not commence until NDA approval is granted and may be short-lived if competitors release similar agents. It is therefore in the pharmaceutical industry's interest to terminate failures early, and to accomplish successful development phases as quickly as possible without compromising the quality of the clinical trials. This is a delicate balance between financial constraints, proper conduct of clinical trials & good clinical practices, and ensuring that regulatory requirements for approval will be met. It is important to note the difference between the innovative development of an NCE and the development of more efficient medications. The latter involves improving on already successful medications by (any or all of) reducing toxicity, increasing potency, reducing the dosing schedule, or by changing to an easier route of administration. Improving a successful agent's efficiency is clearly not as risky as is the development of an NCE, and is rarely a major source of lost revenue.

Analyses of drug development failure consistently reveal that safety, toxicity and economics are the three most important causes of drug failure [5]. Pharmacoeconomic modeling is a vastly different domain compared to the clinical trial domain of the approach described herein, and is beyond the scope of this thesis. However, the impact of safety and toxicity on NCE failure is significant. The cost of an NCE that will ultimately fail is directly proportional to the length of time between IND approval and termination of development. It follows that earlier termination of NCEs destined for failure results in significantly more savings with the added benefits of limiting patient exposure to potentially unsafe and/or ineffective investigational agents, as well as freeing up clinical trial resources for other more promising agents in the development pipeline. Analyses of the distribution of research terminations by clinical phase have shown that over 60% of terminations occur during phases II and III; that is, later in the drug development process [5]. Also, because the later phases are more costly, earlier termination of even a fraction of later phase failures results in a factoring of savings: terminating only 5% of all phase III clinical failures in phase I would reduce out-of-pocket clinical costs by 5.5 –7.1% [6]. However, over-zealous

termination of NCEs will impede the development of innovative, breakthrough therapies. The decision process must balance the cost of terminating what would be a successful NCE against allowing an eventual failure to proceed through phase III. Pharmacovigilance is a difficult and risky task.

**A2. Published Approaches to Decision Analysis in Drug Development**

The aim of this thesis is to devise a Bayesian belief network model (called Pharminator), to calculate the posterior probability that a specific NCE will succeed or fail based on (1) prior data regarding success rates for NCEs of the same therapeutic class and source, and (2) the NCE's therapeutic indices, *in vitro/ in vivo* proof of concept data, and proof of concept data in humans from Phase I and early Phase II studies. The **main distinction between Pharminator and previously described models is that Pharminator focuses on predicting the outcome of a specific NCE**. Other models [2,3,5,7-10] have taken more of a population-based analysis approach, yielding valuable data on overall success rates, but not really addressing the needs of drug developers concerned with the termination decision for a single, specific NCE. Other Bayesian approaches described in the literature differ from Pharminator with respect to the domain to which Bayes theorem is applied. Published Bayesian approaches to pharmacovigilance compare the benefits of Bayesian statistics over frequentist approaches and focus on the utilization of Bayesian statistics for the analysis of clinical trial data which is in turn used to define "stopping boundaries" [11-14] (explored in detail below). Bayesian theory has also been used to facilitate drug development-related tasks such as determining clinical trial sample size [15,16] and designing clinical trials [17]. Yet another proposed use for Bayes theorem is as an alternate approach to utilizing population pharmacokinetic data to predict toxicity in ongoing clinical trials [18]. These are all clearly different tasks from the aim of this thesis, which is neither concerned with the long-standing Bayesian-frequentist debate, nor with the utilization of Bayes theorem to analyze clinical trial results. **Pharminator is specific to individual NCEs rather than individual patients or**

**individual studies, and is much more broad in scope in that it aims to predict safety, efficacy and NCE clinical success for a specific NCE in question.**

At the time of this writing, a review of the literature for decision analytic approaches to pharmacovigilance yields several publications of interest. Berry et al adopted a Bayesian decision-theoretic approach to determine "stopping boundaries" for the development of an NCE [12]. Their approach utilizes accumulating information on the NCE's performance to determine at which point the clinical trial's evidence of efficacy is sufficiently negative that the trial should be stopped. The authors argue that if prior data are "positive, then one should be willing to tolerate somewhat more negative results in the current clinical trial than if previous evidence is also negative." Given that this manuscript was published in 1988, without the benefit of hindsight of the past 15 years' 70-90% NCE failure rate, the authors' argument is representative of a dangerous and costly assumption: that the NCE's clinical trial data can be ignored if it is negative in the context of positive prior data. **A counter-argument** could demand that the posterior probability distribution should be relied upon to reflect the updated belief that incorporates prior data and the NCE's most current evidence, and that if this posterior probability distribution reveals poor performance, serious consideration should be given to terminating the NCE's development. Another difference between Berry's approach and Pharminator is that Berry's approach focuses on efficacy in isolation. Pharminator utilizes a Bayesian belief network to relate safety and efficacy as independent variables, conditional on the common parent (root) node, clinical success. The root node's prior probability distribution is constructed based upon extensive data on NCE failure rates stratified by therapeutic class and NCE source (described in detail in Section B). Berry et al also state "the prior distribution is subjective where the 'subject' is the pharmaceutical company." This forces one to predict the posterior probability for a pharmaceutical company rather than for a specific NCE.

Spiegelhalter et al make a cogent argument for the superiority of Bayesian over frequentist models for the analysis of clinical trial data [14]. The authors' are promoting the use of Bayesian statistics to analyze the outcome of specific ongoing clinical trials. Their argument is not relevant to the model of this thesis. Pharminator utilizes the NCE's characteristics within the framework of a Bayes network to predict the outcome for a given NCE; clearly a different use of Bayes theorem compared to approach advocated by Spiegelhalter. Similarly to Spiegelhalter, Johns and Andersen describe the utility of predictive probabilities for interim analyses of phase II and phase III clinical trials [19]. Pharminator focuses on earlier phase decisions so as to avoid costly phase II and phase III trials. Pallay describes the use of Bayes theorem for economically oriented futility analyses of ongoing phase II clinical trials [20]. This is vastly different from the conditional dependencies incorporated into Pharminator, which relates clinical success, efficacy, safety, therapeutic indices and proof of concept data to update prior beliefs pertaining to the NCE's therapeutic class and source. Taken together, the persistently high rate of NCE failures is the strongest evidence that these previously published and widely adopted approaches do not appear to enable drug developers to be sufficiently accurate in their pharmacovigilance decisions.

### A3. Aim

**The aim of this thesis is to devise an approach that will facilitate improving the efficiency of development of NCEs**. The product of this model is an application (Pharminator) to be used by the drug development team at the phase I/early phase IIa time point for a given NCE that has already passed the IND screening process. The user is prompted to answer several key questions about the NCE (detailed below). The answers provide information regarding which prior probabilities and conditional probability tables are to be used in the model, as well as the observed data upon which prediction will be made. The output is a numeric and graphical (binomial distribution plot) report of the prior and posterior probability distributions for clinical success, safety and efficacy.

## B. Model and Algorithm

### B1. Network Structure and Rationale

A **Bayesian Network** (often referred to as a Bayesian Belief Network (**BBN**) or a 'Bayes Net') [21] is defined as a directed acyclic graph encoding assumptions of conditional independence, with stochastic variables represented as nodes within the network, and inter-variable dependencies represented as inter-nodal links. In addition to a graphical model, a BBN also requires certain parameters to be defined in order to be utilized for probabilistic inference. Therefore, it is necessary to specify the conditional probability distribution for each node. For distributions with a **binary outcome (i.e. 2 states)**, the conditional probability distribution can be represented as a 2x2 **conditional probability table (CPT)**. These tables specify the probabilities that the node is in state (0,1) given that its parent is in state (0,1). The CPT for the **top (root) node**, which has no parent node, is that root node's prior probability distribution. Assuming conditional independence, and utilizing the chain rule of probability, the joint probability for a network consisting of a root node, R, that has n child nodes, $C_1$, $C_2$, ... $C_n$ can be calculated:

$$P(R, C_1, C_2, \ldots C_n) = P(R) * P(C_1 \mid R) * P(C_2 \mid R) * \ldots P(C_n \mid R)$$

Likewise, for the 3-layer BBN in Figure 1 below in which the root node has 2 child nodes and each child node has 2 child nodes (i.e. that are "grandchildren" to the root node), the joint probability for the network can be calculated as in *Formula 1* below:

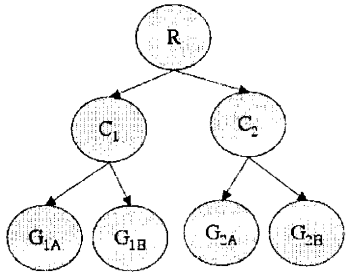



***Figure 1**. An example of a 3-layer BBN. R = root, C= child, G= grandchild*

***Formula 1: Calculation of the Joint Probability Over a Bayesian Belief Network***

$$P(R, C_1, C_2, G_{1A}, G_{1B}, G_{2A}, G_{2B})$$

$$= P(R) * P(C_1 | R) * P(C_2 | R) * P(G_{1A} | C_1) * P(G_{1B} | C_1) * P(G_{2A} | C_2) * P(G_{2B} | C_2)$$

$$= \prod_{i=0}^{n} P(\text{node}_i | \text{parent})$$

**for *n* nodes**

The inner-layer nodes for this network are referred to as **hidden nodes**, and the lowest layer nodes as **leaf nodes**. Calculating the Bayesian posterior probability distribution of the root node given that the leaf nodes are in a specified state is a ratio of the sum of joint distributions. For the example network described above, calculating the probability that the root node is in state 'F" (false), given that all 4 leaf nodes are in state 'T" (true) is achieved as follows:

***Formula 2: Example Calculation of Root Node Posterior Probability***

$$P(R=F | C_1, C_2, G_{1A} = T, G_{1B} = T, G_{2A} = T, G_{2B} = T)$$

$$= \frac{\sum_{v \in \{T, F\}} \sum_{w \in \{T, F\}} P(R=F, C_1 = v, C_2 = w, G_{1A} = T, G_{1B} = T, G_{2A} = T, G_{2B} = T)}{\sum_{x \in \{T, F\}} \sum_{v \in \{T, F\}} \sum_{w \in \{T, F\}} P(R=x, C_1 = v, C_2 = w, G_{1A} = T, G_{1B} = T, G_{2A} = T, G_{2B} = T)}$$

Each of the joint probability distributions in Formula 2 can be calculated utilizing Formula 1.

In this way, Bayes theorem can be applied to a given BBN. At the core of the Pharminator algorithm is a BBN encompassing the most crucial clinical variables in pharmacovigilance (Figure 2). The structure of this network is designed to include only those variables believed to be most critical to predicting NCE failure, based on the literature, the author's training, and consultation with drug development experts.

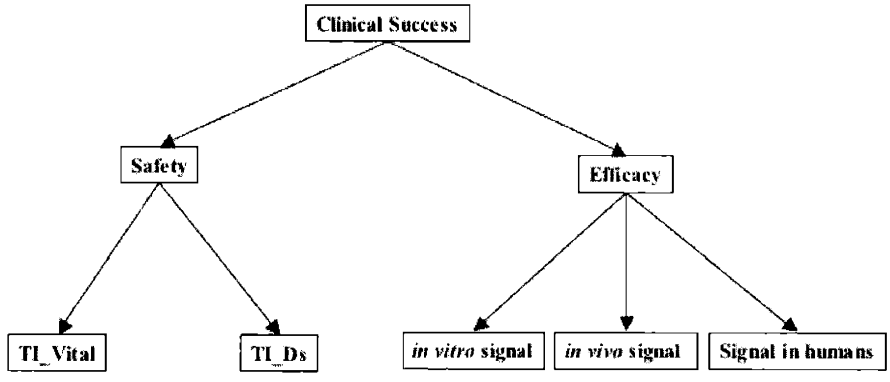



***Figure 2.*** *Clinical variables believed to be most crucial to clinical success for an NCE. This diagram represents a BBN, with implicit assumptions of conditional independence.*

The downward direction of the arrows encodes the NCE's deep or hidden knowledge within "Clinical Success" that is manifested as "Safety" and "Efficacy". The knowledge embedded within each of these two hidden nodes is in turn manifested as therapeutic indices or proof-of-concept signals, respectively. This may seem counter-intuitive, since it may seem logical to believe that safety and efficacy cause clinical success (or failure) rather than represent a

manifestation of an unknown degree of clinical success. However, the goal of Pharminator is to predict the clinical success ***inherent in*** the NCE, rather than to identify causes of clinical success. The NCE has an inherent true degree of "Safety" and "Efficacy". A major goal of clinical trials is to determine what these true values are by studying samples and by assuming that the safety or efficacy in the studied samples are accurate estimates of the NCE's true safety and efficacy. The same explanation can be used to justify the links going from Safety to the therapeutic index nodes, and from Efficacy to the signal nodes. It is this representation that allows Pharminator to predict what the NCE's inherent clinical success is based upon the observed therapeutic indices and proof-of-concept signal data. As this is a BBN, the assumption of conditional independence is required. Therefore, TI_Vital is assumed to be independent of TI_Disease (see Figure 2 and Section B2 for definitions) conditional on the common parent node, Safety. The same is true for the *in vitro*, *in vivo* and human signal nodes, and their common parent, Efficacy.

Figure 2 is designed to be a "best guess" representation of those variables deemed most important in predicting clinical success. It is likely that future versions of Pharminator will encode a BBN with slightly different leaf nodes, based upon accumulated data on NCE failure specific to therapeutic classes. Also absent from the BBN is a representation of idiosyncratic severe adverse events which are unpredictable, by definition. **Future versions of Pharminator will undoubtedly include pharmacogenomic markers of adverse events and drug resistance**.

**B2. Definitions**

Several terms require definition in order to understand the Pharminator algorithm:

- Clinical Success: An NCE is clinically successful if it is still on the market 1 year after NDA approval.
- Efficacy: An NCE is efficacious if it produces a "sufficient" degree of change in a surrogate or true marker, compared to control (placebo or current gold standard

therapy). What constitutes a "sufficient" degree of change depends on (1) the clinical indication and (2) the development phase (during Phase II, the signal need not be statistically significant, while Phase III studies must show statistical significance in at least two separate trials). Certain indications may require only modest effect from an NCE in order to be successful (e.g. acute, relatively benign disorders), while others may require extreme effects (e.g. life-saving therapies).

- Life-saving: An NCE is considered life-saving if
    1. the disease for which it is indicated is fatal and,
    2. there are no alternate life-extending therapies.

  The Pharminator model utilizes the life-saving status of the NCE to determine the influence that the Safety data will have on the calculation of the posterior probability of Clinical Success. The assumption is that a higher degree of toxicity is tolerated for an NCE that is truly life-saving, as defined above, thereby making the probability of clinical success largely dependent on efficacy. For example, an NCE that truly extends life expectancy but causes acute renal failure may still have a high probability of success because it is assumed that the initiation of dialysis is preferable to death. This assumption is open to argument from the point of view of quality of life issues, since truly curative therapies for lethal disorders are rare, however assuming that supportive therapies are preferable to death is reasonable.
- Markers- surrogate vs. true: A marker is an indicator of response to therapy. A surrogate marker is a marker that is not directly or primarily involved in the pathogenesis of the disease, whereas a true marker is primarily integral to the disease mechanism. An example of a surrogate marker is the CD4 count in HIV. A true marker for HIV is viral load.
- NCE source- acquired, self-origin: abroad, self-origin: USA : An acquired NCE is an NCE that a pharmaceutical company has licensed-in from another company, such as

a biotechnology firm, or that has been acquired from some other source. A self-originated NCE is an NCE for which the initial pre-clinical development occurred within the same pharmaceutical company that will assume responsibility for conducting clinical trials. The prior probability of success differs significantly between acquired and self-originated NCEs [5]. Not surprisingly, NCEs that have undergone initial clinical testing abroad (and demonstrate potential effectiveness in humans) are more likely to succeed.

- Prior bias: Pharminator gives the user the option of selecting whether the prior probability should be optimistic or pessimistic. As described in detail in Section B3 "Prior Probabilities and Conditional Probability Tables", the NCE's intended therapeutic class affects the selection of the NCE's prior probability of clinical success. Prior data on NCE success rates are stratified by therapeutic source, and include the total number of NCEs within each therapeutic class, the fraction of the total NCEs that have failed, and the fraction of the total NCEs that are still under development, for a total of 671 NCEs spanning IND filing dates from 1981 to 1992 [5]. However, it should be noted that if an NCE is sufficiently safe and effective, the NCE's posterior probability of success will be high, regardless of the prior probability. The Tufts Center for the Study of Drug Development's (TCSDD) published reports provide the prior probability of failure based on the current failure rate, as well as the probability of failure assuming that all NCEs still under development are successful (i.e. a more optimistic prior probability of failure). When the user sets Pharminator's prior bias to "pessimistic" (the default setting), the former prior probability is used- i.e. the prior probability of failure based on the current failure rate. When the user sets Pharminator's prior bias to "optimistic", the latter prior probability is used- i.e. the prior probability of failure based on the assumption that all NCEs still under development will not fail.

- Safety: An NCE's safety is essentially synonymous with toxicity. Every NCE that is not an inert placebo has some degree of "toxicity" in that even the desired effects of an NCE become toxic if a large enough dose is given. An NCE's safety is therefore defined as a degree of toxicity that is an acceptable balance against the benefit to the patient. Ultimately, an NCE can only be deemed safe once it has undergone Phase IV post-marketing surveillance. Prior to Phase IV, insufficient numbers of patients have received the NCE such that rare but severe idiosyncratic reactions would not likely be detected.
- Therapeutic class: The therapeutic class is the organ system affected by the disease process for which the NCE is indicated. This definition of therapeutic class is utilized rather than the more traditional chemical class because prior probabilities of success are known for a total of 671 NCEs, stratified by therapeutic class [5], and stratifying by chemical class would fractionate the data beyond use with too many categories and too few NCEs in each category. The therapeutic classes included in Pharminator are: Analgesic/Anesthetic, Antimicrobial, Antineoplastic, Cardiovascular, Central Nervous System (CNS), Endocrine, Gastrointestinal (GI), Immunologic, Respiratory, and Miscellaneous. Clearly these are less specific categories than those used by clinicians (e.g. Calcium channel blockers, ACE inhibitors, mono-amine oxidase inhibitors etc.) however. as stated above, there do not appear to be sufficient data to allow for a more specific stratification without fractionating the data beyond utility.
- Therapeutic index- vital organ, disease: The therapeutic index (TI) is the ratio of the NCE dose that produces an undesired effect to the NCE dose that produces the desired effect in a proportion of the study population. The numerator is the $TD_x$ (toxic dose in x% of the population) and the denominator is the $EC_y$ (effective dose in y% of the population). Each NCE has several therapeutic indices, depending on the number of specific adverse events (e.g. the TI for hepatotoxicity is different from the

TI for nephrotoxicity), the number of specific desired effects (e.g. ACE inhibitors reduce blood pressure and reduce proteinuria), and depending on the definition of the proportion of the study population (i.e. the values of x and y). A larger TI represents a generally safer NCE. A smaller TI will be either too unsafe to be used clinically, or will require very close therapeutic drug monitoring in order to ensure safety (e.g. digoxin). The Pharminator model is designed to be inherently pessimistic, given the high rate of NCE failures to date, and the extreme costs associated with these failures. Therefore, the current implementation of Pharminator requires input for two specific TIs: the lowest TI for an undesired effect on any vital organ (brain, heart, lungs, liver, kidney, exocrine pancreas, bone marrow), and the lowest TI for an undesired effect on any organ or system that is already adversely affected by the disease/system for which the NCE is indicated. An example of the latter is retinal toxicity caused by an NCE indicated for the treatment of diabetes mellitus. It is likely that as Pharminator evolves, additional or alternate TI variables will be added to the model.

It is important to note the reasons why therapeutic class and NCE source are included in the model as "prior probability modifiers" and not as stochastic variables (nodes) in Figure 2:

1. As discussed in section B1. "Network Structure and Rationale", Safety and Efficacy do not "cause" clinical success in the Pharminator model. The Safety and Efficacy nodes (and indeed all child nodes in the model) are in fact manifestations of the inherent degree of clinical success of the NCE. Contrary to this, therapeutic class and NCE source have a direct impact on NCE clinical success [5].
2. While the NCE's true safety, therapeutic indices, efficacy and proof-of-concept signals are not known, the NCE's intended therapeutic class and source are known with certainty. It is therefore nonsensical to represent therapeutic class and NCE source as

stochastic variables. However, Figure 3 is provided below as an adjunct to Figure 2, in order to

a. demonstrate the dependencies between therapeutic class, NCE source and clinical success, and
b. demonstrate that the noisy-or assumption [22,23] can be utilized to calculate the prior probability of clinical success from the prior data on NCE failure rates stratified by therapeutic class and NCE source [5].

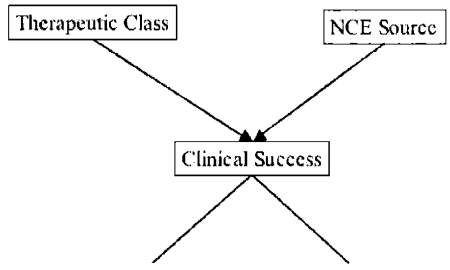



***Figure 3.*** *Illustration of how Therapeutic Class and NCE Source relate to Clinical Success (see Figure 2 as well). The direction of the arrows indicates that Therapeutic Class and NCE Source have a causal effect on Clinical Success. Determining P(Clinical Success | Therapeutic Class, NCE Source) is therefore not a Bayesian posterior probability and the noisy-or assumption is applicable (see text for details).*

## B3. Prior Probabilities and Conditional Probability Tables

### Prior Probability: Source & Selection

TCSDD publications are the most extensive, accessible, and reliable source of the prior probability of NCE success (and failure). DiMasi recently analyzed the causes of failure and reported the success rates for 671 NCEs for which INDs were filed between 1981 and 1992 [5]. In his report, he provides "current and maximum possible success rates" stratified by therapeutic

class for 503 self-originated NCEs. The "current success rate" is the fraction of the number of NCEs (in that class) that have been successful over all NCEs in that class. This is in fact a pessimistic prior because the implicit assumption is that all open NCEs (i.e. NCEs still in development) will fail. The "maximum possible success rate" is the success rate assuming that "all open NCEs will eventually be approved" – an optimistic assumption. DiMasi also provides probabilities of NCE success stratified by NCE source (see definition above, in section B2). Pharminator asks the user to indicate the NCE's therapeutic class, NCE source, as well as the user's desired "prior bias", which may be either pessimistic or optimistic. The prior bias determines which therapeutic class prior probability is utilized: if the user selects "pessimistic" (the default setting), the current success rate is used to calculate the NCE's prior probability of success. Conversely, if the user selects "optimistic", the maximum possible success rate is used.

Although DiMasi's analysis is quite informative, he did not sub-stratify by therapeutic class ***and*** NCE source combinations. In order to allow Pharminator to choose a prior probability that most accurately reflects the NCE's therapeutic class *and* its source, the algorithm utilizes the noisy-or assumption [22]. Paraphrasing Szolovits [23], the noisy-or assumption states that the probability that some set of variables causes an outcome equals the probability that at least one of the variables does so. The probability of interest is P(Clinical Success | Therapeutic Class, NCE Source). For an explanation as to why this is not a posterior probability distribution, see section B2 and Figure 3. Given the noisy-or assumption, the probability of interest can be calculated:

***Formula 3: Noisy-Or***

1 - P(Clinical Success | Therapeutic Class, Source)

= (1 – P(Clinical Success | Therapeutic Class)) * (1 – P(Clinical Success | Source))

Therefore, P(Clinical Success | Therapeutic Class, Source)

= 1 – [(1 – P(Clinical Success | Therapeutic Class)) * (1 – P(Clinical Success | Source))]

≈ the prior probability, P(Clinical Success) for the NCE in question

and P(Clinical Failure) = 1- P(Clinical Success)

Pharminator utilizes the prior bias selected by the user to determine which prior probability of success to utilize for the selected therapeutic class, then uses this probability along with the probability of success for the selected NCE source to calculate P(Clinical Success | Therapeutic Class, Source), given the noisy-or assumption as in Formula 3 above. This calculated probability is utilized as the prior probability of Clinical Success for the NCE in question (Figure 4).

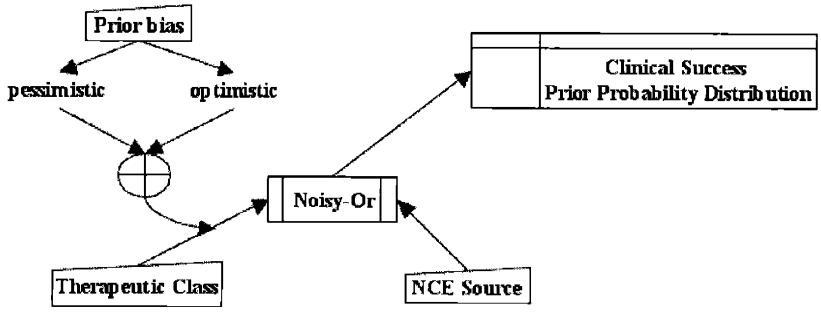



***Figure 4.*** *Flowchart depicting how the prior probability of Clinical Success is determined. See text for details.*

**Conditional Probability Tables**

As described in section B1, a BBN requires CPTs for each node in order to be utilized for probabilistic inference. Although the data provided in the TCSDD reports is valuable, the format of those reports is not directly applicable to the construction of BBN conditional probability tables. **The main limitation of Pharminator is the absence of appropriate conditional probability data and the need to make certain assumptions to allow utilization of the data from the TCSDD published reports. Future versions and revisions of Pharminator will focus on the task of obtaining appropriate data to populate the CPTs**. These assumptions will be tested by sensitivity analyses once characteristics and outcomes for specific successful and failed NCEs become available. In the absence of such data, the methods by which the CPTs are currently constructed are described in this section.

DiMasi analyzed the causes of failure for 348 NCEs that were withdrawn from development [5]. It should be noted that NCEs that proceeded through all clinical trial phases but failed to achieve NDA approval are *not* included in DiMasi's analysis. As well, DiMasi stratified the causes of failure by "primary" cause, thereby not disclosing any degree of overlap – i.e. NCEs that failed primarily due to one reason, but may have also failed for another reason (e.g. an NCE that failed because it was not safe, but was also not very effective). His analysis demonstrated that of a total of 348 NCEs that were terminated, the *primary reason* for termination was efficacy in 121, safety in 72, economics in 109, and "other" in 46. Since Pharminator is concerned only with **safety and efficacy**, the probability that safety is the *primary cause* of failure is 72/(72 +121) = 0.37, and the probability that efficacy is the *primary cause* of failure is 121/(72 + 121) = 0.63. **Assuming that the proportions of causes of failure are consistent across the withdrawn drugs**, the CPT probabilities, P(Safety=F | Clinical Success = F) and P(Efficacy=F | Clinical Success = F), can be calculated by an **"overlap" function**, as follows:

***Formula 4: "Overlap" Function***

P(**Safety=F** | Clinical Success = F)

= total number of primary safety failures +

(total number of primary safety failures * proportion of primary efficacy failures)

= 72 + (72 * 121/193) = **0.606**

P(**Efficacy=F** | Clinical Success = F)

= total number of primary efficacy failures +

(total number of primary efficacy failures * proportion of primary safety failures)

= 121 + (121 * 72/193) = **0.860**

These values (and their respective complement values) occupy the first rows of their respective CPTs.

While the overlap function permits estimation of the first row of each of the Safety and Efficacy CPTs (i.e. P(node = F | parent = F) and P(node = T | parent = F)), currently, there are no adequate, available data for the second rows of the Safety and Efficacy CPTs (P(node = F | parent = T) and P(node = T | parent = T)). For now, these values are currently set as **pessimistic estimates**. For the **Efficacy CPT**, the probability P(Efficacy = F | Clinical Success = T), i.e. the probability that an NCE is not efficacious given that it is clinically successful, is logically estimated to be very low. Until data for sensitivity analyses become available, this value is set at 0.01, and its complement, P(Efficacy = T | Clinical Success = T) is therefore 1 – 0.01 = 0.99. For the **Safety CPT**, the probability P(Safety = F | Clinical Success = T), i.e. the probability that an NCE is not safe given that it is clinically successful is also estimated to be low, but likely not as low as for P(Efficacy=F | Clinical Success = T). Until data for sensitivity analyses become available, this value is set at 0.05 for NCEs that are not life-saving, and for NCEs that are life-saving, this value

is set at 0.5. This difference is to reduce the influence that TIs have on the posterior probability of clinical success for life-saving NCEs.

The construction of the hidden nodes' CPTs is depicted in Figure 5.

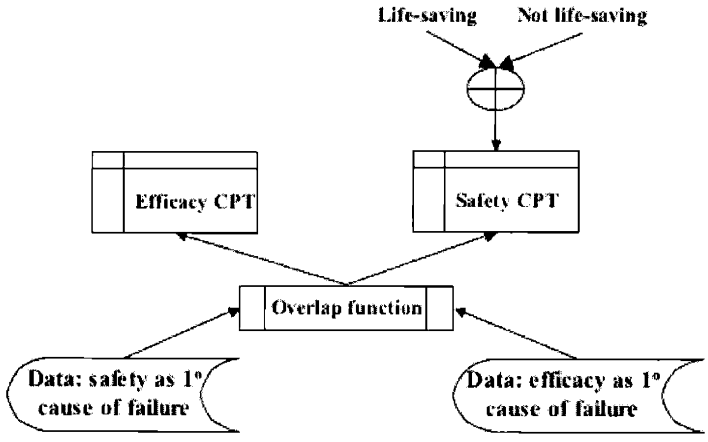



***Figure 5.*** *Algorithm demonstrating how the overlap function (Formula 4) and the life-saving preference are utilized to determine the CPTs for the Safety and Efficacy nodes.*

Just as for the hidden nodes (Safety and Efficacy), there are no easily accessible data on therapeutic indices and proof of concept signal data for NCEs that have failed. Once again, **the task of acquiring appropriate data to inform the CPTs will be the major focus for future versions and optimization of Pharminator**. However, devising models that approximate these relationships is a somewhat less arduous task than for Safety and Efficacy. With respect to the relationship between TI and Safety, it is known that TI is directly proportional to the degree of safety because a larger TI simply means more prescribing "room" between the effective dose and the toxic dose. Most prescription medications have TIs that are in the 8-10 range. Given that TI is a ratio, and that the lowest rational value for a TI is 1, the relationship between TI and safety can be approximated by a logistic sigmoid model (Formula 5, Figure 6):

***Formula 5: Logistic Sigmoid Function***

$$\mathbf{P(TI \mid Safety)} \approx \frac{1}{1 + e^{-(s * (x - i))}}$$

$x$ = **TI**
$s$ = **slope**
$i$ = **intercept**

The slope and intercept of this model were selected to reflect what is believed to be an accurate approximation of the relationship between TI and P(TI | Safety).

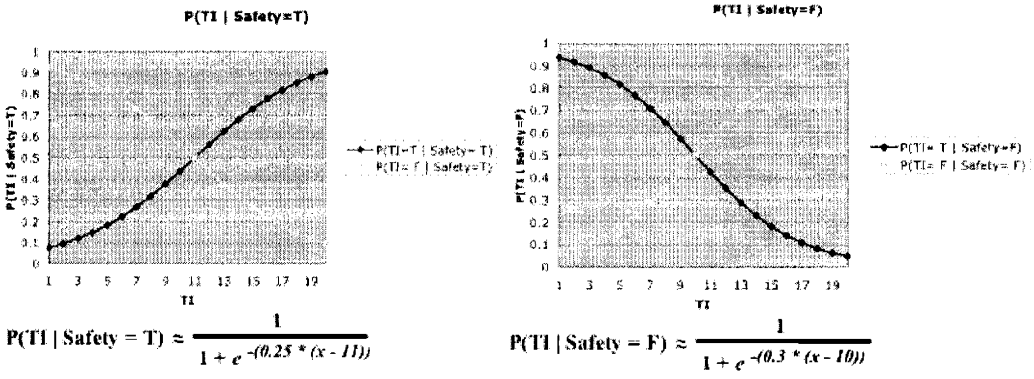



***Figure 6.*** *Graphs showing the logistic sigmoid functions that are used to approximate the P(TI | Safety) CPT values from the TI values.*

A similar assumption is made for the proof-of-concept signal data in that the quantity of the signal is proportional to the degree of efficacy. In contrast to the sigmoid relationship between TI and safety, the relationship between proof-of-concept signal and efficacy is assumed to be a simple linear function ($y = mx + b$). However, the signal data must first be transformed to a standardized measure so that different ranges and scales will not influence the interpretation of the signal. For this purpose a modified signal-to-noise ratio is used (Formula 6), requiring the user to enter the mean and variance (S.D.$^2$) for the control and experimental groups, for each experimental environment (*in vitro*, *in vivo* – highest-order species, human), as well as whether each signal is a true or surrogate marker. This formula provides a variance-corrected measure of the degree of signal as a value between 0 and 1. The resultant signal-to-noise ratio value is

utilized by the linear models to estimate P(Signal | Efficacy) for *in vitro*, *in vivo* and human signals, stratified by true and surrogate markers. The slopes (and intercepts) of the linear models are adjusted to reflect differences between *in vitro*, *in vivo* and human signals, and between true and surrogate markers (Figure 7). Specifically, the slope of the function is proportionate to:

a) the environment order (*in vitro* < *in vivo* < human), and

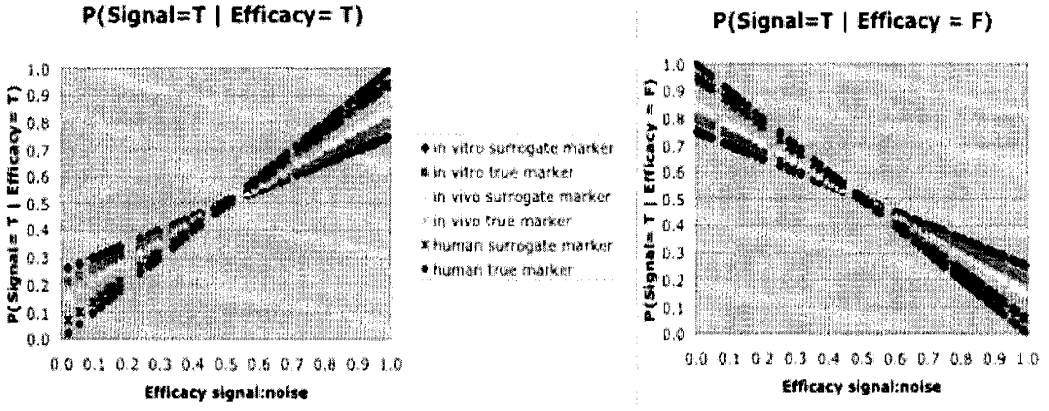
b) the marker (surrogate < true)

***Formula 6: Modified Signal:Noise Ratio***

$$\text{Modified signal:noise} = \frac{\left| \left( \overline{\text{NCE}} - \overline{\text{control}} \right) \right|}{\max\left(\text{NCE 95\% UCL, control 95\% UCL}\right)}$$

**UCL = upper confidence limit**

(Figure 7 shown on next page)




***in vitro***
**surrogate: $y = 0.5x + 0.25$**
**true: $y = 0.6x + 0.2$**

***in vivo***
**surrogate: $y = 0.7x + 0.15$**
**true: $y = 0.8x + 0.1$**

**human**
**surrogate: $y = 0.9x + 0.05$**
**true: $y = x$**

***in vitro***
**surrogate: $y = 1 - 0.5x - 0.25$**
**true: $y = 1 - 0.6x - 0.2$**

***in vivo***
**surrogate: $y = 1 - 0.7x - 0.15$**
**true: $y = 1 - 0.8x - 0.1$**

**human**
**surrogate: $y = 1 - 0.9x - 0.05$**
**true: $y = 1 - x$**

***Figure 7.*** *Linear functions demonstrating how the modified signal:noise ratio value is utilized to approximate P(signal | Efficacy) for a series of randomly-generated NCE & control means & variances. The linear functions are stratified by experimental environment (in vitro, in vivo, human), and type of marker (surrogate vs. true marker). The slopes and intercepts are adjusted to reflect what is believed to be an adequate approximation (remains to be validated).*

The entire algorithm for constructing the leaf node CPTs is summarized in Figure 8.

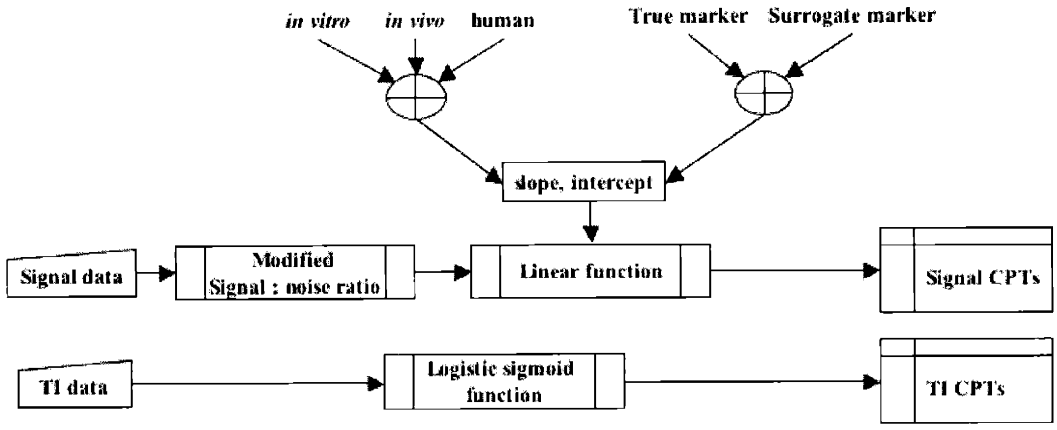



***Figure 8.*** *Overview of algorithm for constructing leaf node CPTs.*

The default state for all leaf nodes is "True", and it is the leaf node CPTs that change in response to the TI and signal data entered by the user. This represents one major departure from BBN methodology: Pharminator's use of the input data to determine the specific CPT to be used for a given set of fixed leaf node states. The acquisition of data on NCEs that have failed will facilitate the modification of the linear and sigmoid functions rather than direct changes to the leaf node CPTs. Justification for this approach is that this can actually work to the advantage of each drug development institution that utilizes Pharminator: Many pharmaceutical companies develop medications in a small number of therapeutic classes and therefore have NCE failure data that is highly specific to that pharmaceutical company's future development projects. Therefore the use of these data to modify Pharminator's CPT functions will result in a company-specific implementation of Pharminator, the predictive accuracy of which will be directly proportionate to the specific company's development history and prior investments in NCE failures (i.e. accuracy proportionate to their losses). Smaller companies with little or no development history will not have the ability to implement a company-specific implementation, but will benefit from the prior knowledge of the entire industry, excluding confidential and privileged information from other companies. Data from a larger pharmaceutical company will remain exclusive to that specific

pharmaceutical company unless that company agrees to allow Pharminator to utilize their data for the benefit of the entire industry, always maintaining confidentiality regarding specific NCEs that have failed.

**B4. Data Input and Output**

Pharminator requires the following input from the user:

- NCE "demographics": NCE name, therapeutic class, source, life-saving status
- User's prior bias preference (pessimistic vs. optimistic; default is pessimistic)
- Signal data- for each of *in vitro*, *in vivo* (highest-order species) and human, the following data are entered ***for the maximal dose given, regardless of toxicity***. Therefore, toxicity (safety) is not taken into account for signal data because Safety and Efficacy are assumed to be independent variables, conditional on the common parent, Clinical Success. The signal nodes' inputs include:
    - NCE mean & variance
    - Control mean & variance
    - Type of marker: true or surrogate
- Minimum TI_Vital (see section B2 for definition)
- Minimum TI_Disease (see section B2 for definition)

The current implementation of Pharminator requires all of these values to be entered in order for the posterior probability to be calculated. Future versions of Pharminator will perform probability inference on partial nets (i.e. nets that are missing one or more leaf node variables).

With this information, Pharminator selects the appropriate prior probability of clinical success, and calculates the posterior probability distribution for Clinical Success, Safety, and Efficacy.

The hidden node "prior" probabilities are required in order to calculate hidden node posterior probabilities. These "prior" probabilities are calculated from the hidden and root nodes' CPTs:

***Formula 7: Calculating a Hidden Node's "Prior" Probability:***

$P(\text{Hidden Node}) = \Sigma\,[P(\text{Hidden Node} \mid \text{Parent Node}) * P(\text{Parent Node})]$

The prior and posterior probability distributions are displayed graphically as binomial distributions. The "n" for the Clinical Success prior probability distribution ("prior N") is the total number of NCEs from which the prior data were attained. The "n" for the Clinical Success posterior probability distribution (post N) is (prior N + 1). This is likewise for the prior N and post N for the Safety and Efficacy probability distributions (see section C. "Implementation and Examples" for a pictorial demonstration).

## B5. Algorithm

Conglomerating Sections B1 through B4, including Figures 4,5 and 8 and Formula 2 results in the algorithm shown in Figure 9.

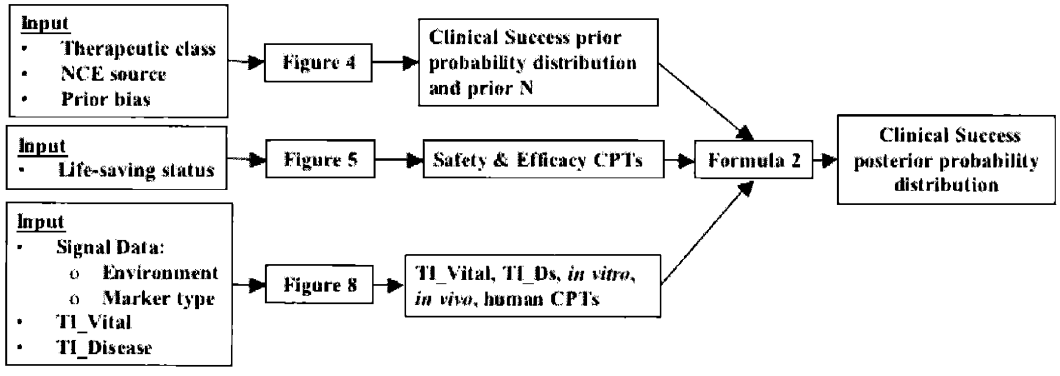



***Figure 9.*** *Overview algorithm, combining the components described in Sections B1 through B4.*

## C. Implementation and Examples

### C1. Implementation

Pharminator is implemented in Java 1.4, using Apple ProjectBuilder v2.1, on Apple OS X.2 Jaguar. An object-oriented, model-view approach [24] was utilized to structure to the program. The accuracy of the BBN was validated against Bayesware Discoverer® (http://bayesware.com).

### C2. Example 1: CurOnc (fictional)

CurOnc is a fictional anti-neoplastic agent devised solely for the purpose of illustrating some key features of Pharminator. CurOnc is self-originated in the USA and meets the definition of "life-saving". The signal inputs, TI inputs, and Clinical Success probability distribution plots are shown in Figure 10. The Safety and Efficacy probability distribution plots are shown in Figure 11. The effect of changing the life-saving option to "Not Life-Saving" is shown in Figure 12. The effect of changing the prior bias to optimistic is shown in Figure 13. Overall, the probability distributions generated by Pharminator suggest that CurOnc has a high probability of efficaciousness (0.7872), but is also very likely to have significant toxicity (P(Safety=T) = 0.0645). Therefore, if CurOnc is indeed "life-saving", it has a probability of Clinical Success of 0.4951 with little overlap between the Clinical Success prior (0.2304) and posterior probability distributions. However, if CurOnc is not truly life-saving, its probability of Clinical Success is 0.2016 (less than the prior probability) when prior bias is pessimistic, and at best (prior bias = optimistic), the probability of Clinical Success is 0.3359, which is still less than the prior probability. Given these results, development of CurOnc should be continued only if it is deemed to be truly life-saving.

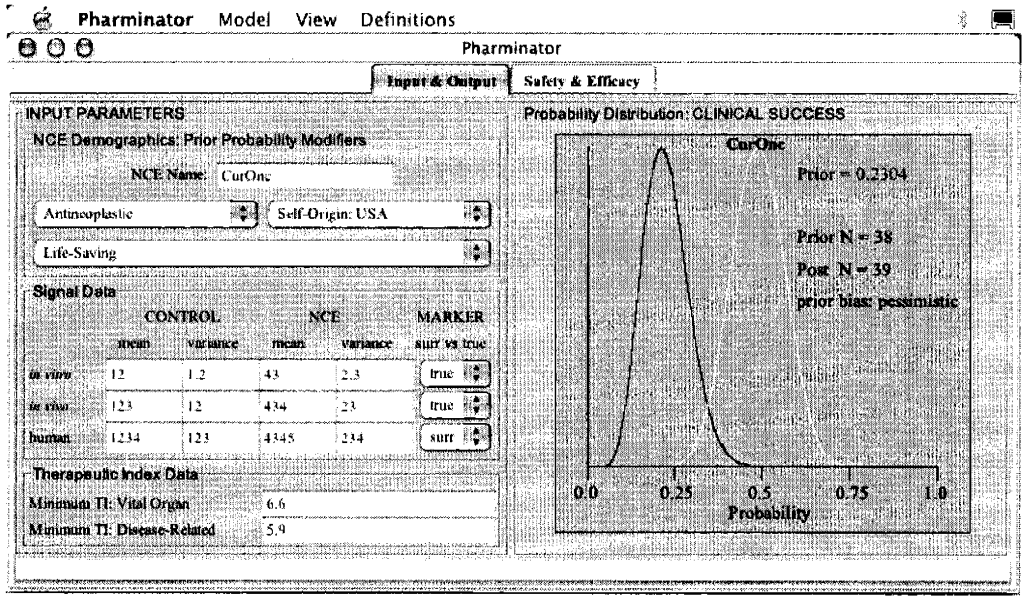



***Figure 10.*** *The prior and posterior probability distributions for Clinical Success are shown for the fictional antineoplastic agent, CurOnc.*

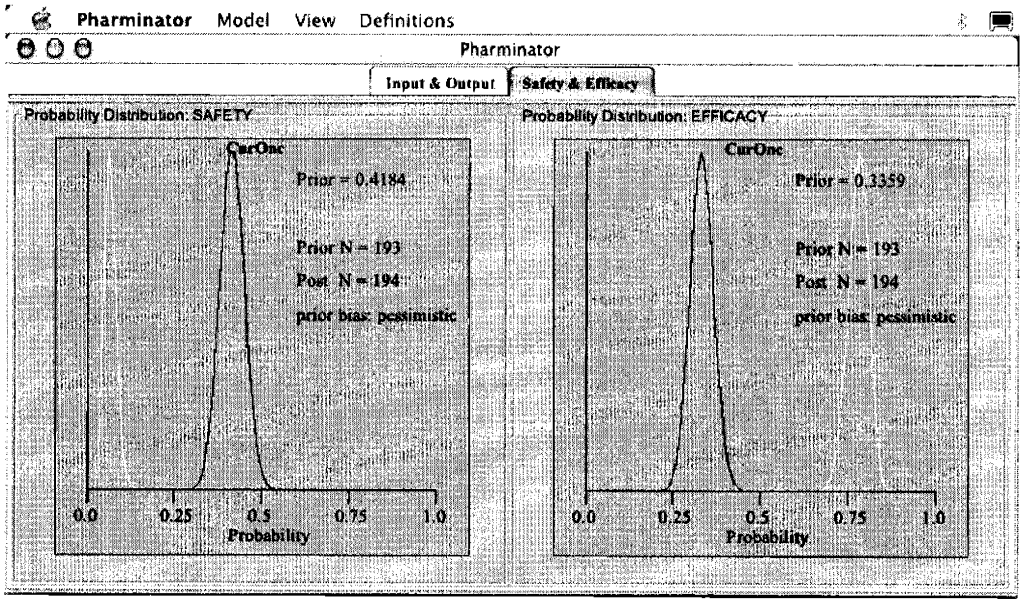



***Figure 11.*** *The prior and posterior probability distributions for Safety and Efficacy are shown for the fictional antineoplastic agent, CurOnc.*

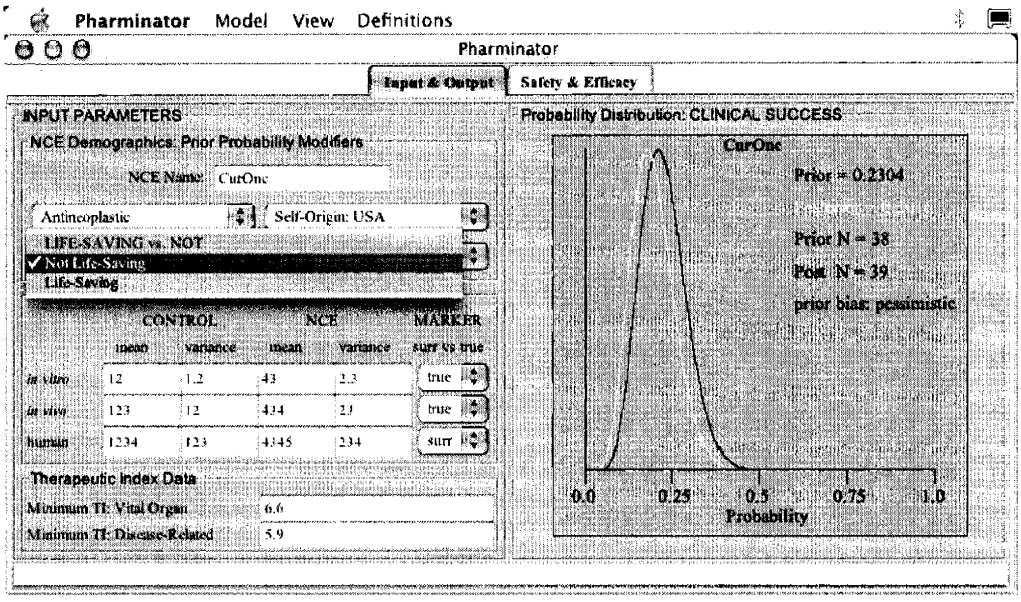



***Figure 12.*** *The effect of selecting the "Not Life-Saving" option on the posterior probability distribution for Clinical Success is shown for the fictional antineoplastic agent, CurOnc. Note that the posterior probability distribution has shifted to the left compared with Figure 10, suggesting that CurOnc is likely not safe- confirmed by the low probability of Safety shown in Figure 11.*

(Figure 13 shown on next page)

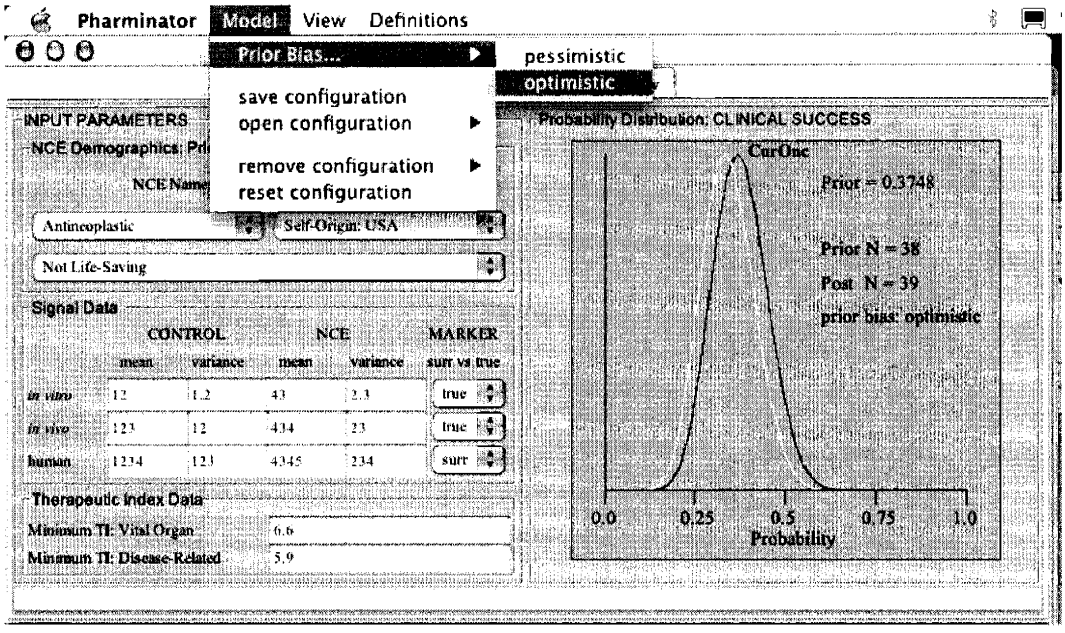



***Figure 13.*** *The effect of setting the prior bias to "optimistic" on the prior and posterior probability distribution for Clinical Success is shown for the fictional antineoplastic agent, CurOnc. Note that both probability distributions have shifted to the right, compared with Figure 12.*

## C3. Example 2: LY203638 (rhAPC)

Recombinant human activated protein C (rhAPC) is a relatively novel agent that is known for its anti-coagulant, pro-fibrinolytic, and anti-inflammatory properties. Eli Lilly™ Research laboratories has developed LY203638 (rhAPC) as a novel therapy for sepsis (Clinical Investigator's Brochure kindly provided by Dr. Robert Rubin). In general, this example is limited in that several unpublished pre-clinical efficacy studies are listed in the Clinical Investigator's Brochure, but no data are accessible. The most relevant *in vitro* study was used; this *in vitro* study was performed prior to the go/no-go decision time point [25]. Bajzar et al reported dose-dependent lysis times, but did not include any measures of variability. Therefore, *in vitro* variance is set to 0 for both the NCE and control (the *in vitro* variance entries are actually set to 0.000001 because

the program's current implementation will not calculate posterior probabilities if any value is 0. This minor problem will be resolved with future implementations). Published *in vivo* studies performed prior to the go/no-go decision time point evaluated pre-clinical efficacy in primates, canines, guinea pigs, and rats. The primate data [26] are used for this example because primates are the highest-order species studied. Early phase II study data in humans were provided in the Clinical Investigator's Brochure. This example is an approximation based upon the accessible information only. Note that the true outcome marker is successful treatment of sepsis. The Phase II endpoints reported are therefore all surrogate markers: organ failure-free days, number of transfusion requirements, ICU-, Hospital-, and Ventilator-free days, and 28-day all-cause mortality. For the purpose of this example, organ failure-free days (shock) was chosen as a good sepsis-specific surrogate marker in that multi-organ failure and sepsis-related morbidity are very tightly correlated. No specific data on therapeutic indices could be found either in the Clinical Investigator's Brochure or in the literature from the go/no-go decision time point. However, the Clinical Investigator's Brochure contains data from Phase I studies at doses ranging from 12 – 48 μg/kg/hour suggesting that the TI is at least 4 (48/12). Toxicology studies in primates demonstrated that the "no-observed-adverse-effect level" was 2 $mg/m^2$/hour with toxic effects observed at a dose of 8 $mg/m^2$/hour. Taken together, these data suggest that the TI is approximately 4. In the absence of more accurate TI data, this value is used in the example. LY203638 is classified as a cardiovascular agent (since there are no prior data for hematologic agents and it is not antimicrobial). The limitations in acquiring appropriate data for LY203638 underscores the requirement to have unfettered access to the NCE's data in order to optimize Pharminator's predictive accuracy.

Figure 14 demonstrates that LY203638 has a very low probability of Clinical Success of 0.0521, much lower than the prior probability of 0.246. Figure 15 shows that the probabilities of safety and efficacy are both very low (0.0211 and 0.0667, respectively). Even when the prior bias is set

to optimistic (Figure 16), the probability of clinical success is only 0.0526, also much lower than the optimistic prior probability of 0.2916. Therefore, based only on data available prior to later Phase II studies: even as a life-saving NCE, and when assuming an optimistic prior probability of success, LY203638 has a very low probability of clinical success based only on data available prior to Phase III studies. Of interest, after LY203638 received NDA approval, subsequent post-approval studies raised several concerns about LY203638's safety and efficacy, calling for Phase IV studies to be performed[27].

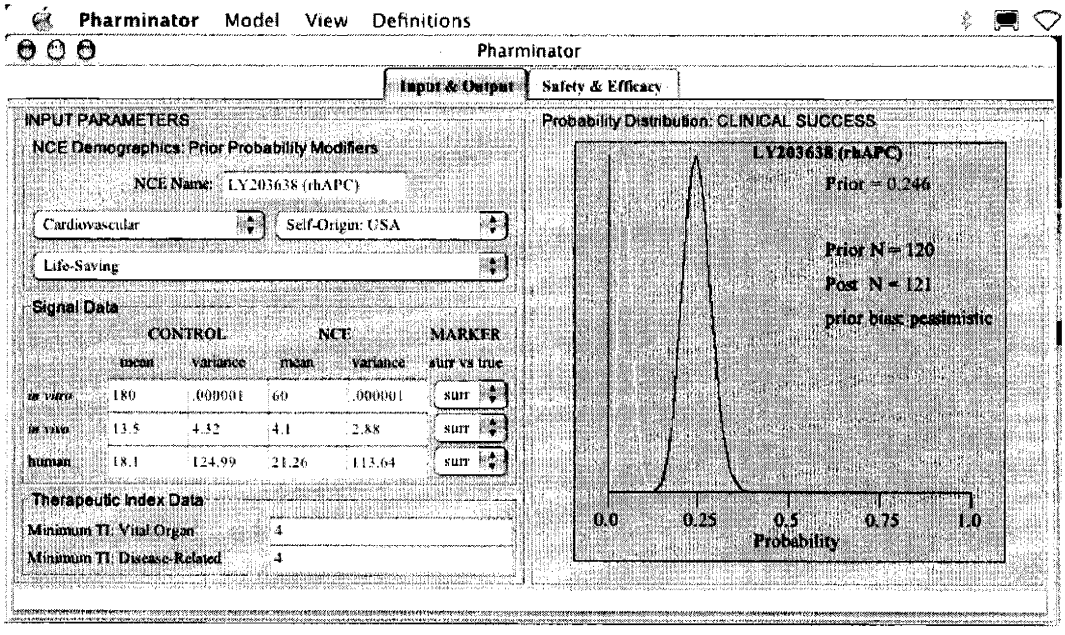



***Figure 14.*** *The prior and posterior probability distributions for Clinical Success are shown for LY203638 (rhAPC).*

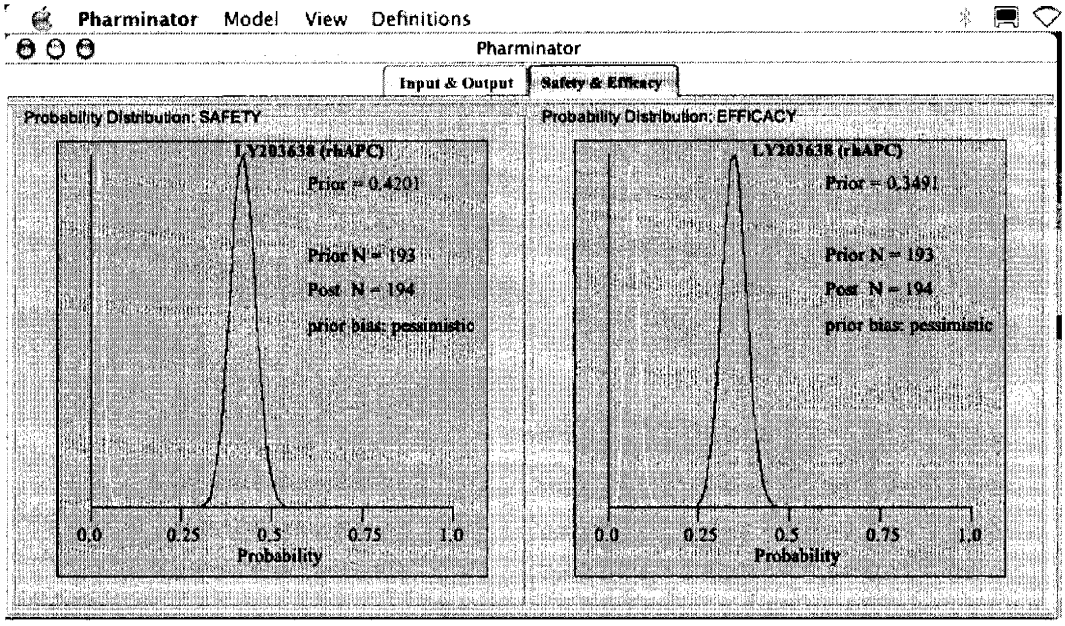



***Figure 15.*** *The prior and posterior probability distributions for Safety and Efficacy are shown for LY203638 (rhAPC).*

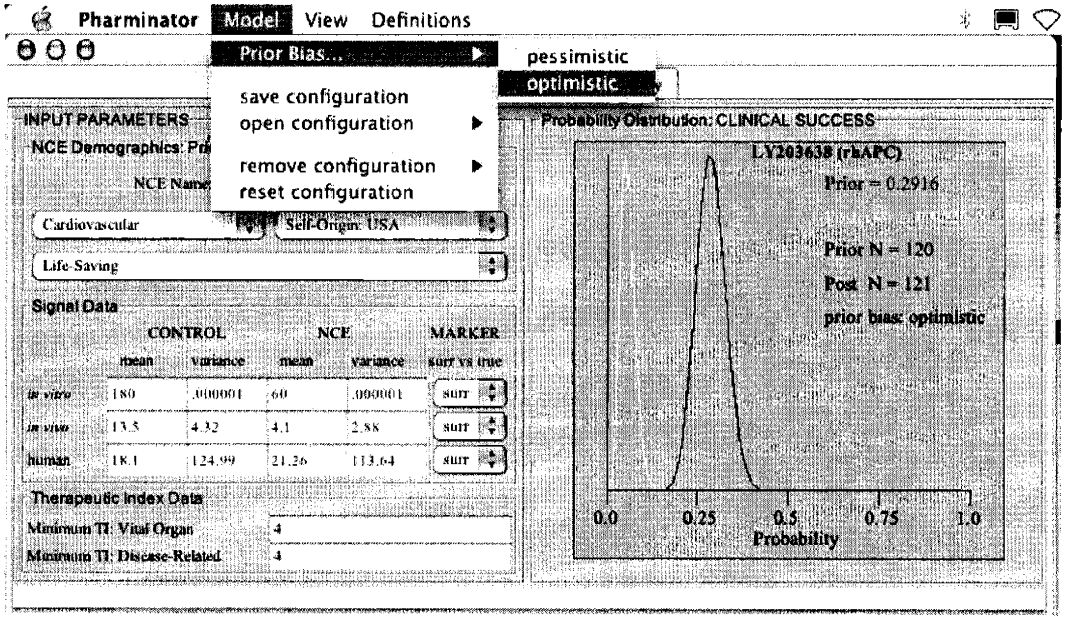



***Figure 16.*** *. The effect of setting the prior bias to "optimistic" on the prior and posterior probability distribution for Clinical Success is shown for LY203638 (rhAPC).*

**D. Future Directions: Plans for Testing, Validation and Optimization**

While Pharminator is an apparently novel model for pharmacovigilance, it is unknown how accurate the model is. As stated several times above, Pharminator's major limitation is the lack of access to appropriate data to populate the CPTs. The specific types of data needed include safety and efficacy data from:

- ***NCEs: IND-approved but withdrawn prior to NDA filing***
- NCEs: IND-approved but NDA denied (and not subsequently approved)
- Successful NCEs

Fortunately, Pharminator will continue development as the primary focus of a Sloan Industry Fellowship at the Program on the Pharmaceutical Industry at MIT, (7/2003 – 6/2004). The predominant aims of the Sloan Fellowship are (1) to acquire data on NCEs that have failed, (2) to validate Pharminator, and (3) to optimize Pharminator. Along the lines of Aim (3), Pharminator can be optimized for a specific pharmaceutical company, leveraging the strengths and extent of the company's development history. As stated in Section B3, Pharminator could be trained on company-specific data such that its predictive accuracy will be directly proportional to the company's development history and prior investments in NCE failures - i.e. the accuracy may be proportional to the extent of capital lost in NCE failures.

**E. Summary and Conclusions**

In summary, NCE failure rates are increasing despite concomitant increases in research & development expenditures. Previously utilized approaches to pharmacovigilance have not been successful. Pharminator's approach is novel in that individual NCE characteristics are modeled on the background of prior data specific to those characteristics. Pharminator does not include pharmacoeconomic parameters in its model, and therefore Pharminator should be used in

conjunction with pharmacoeconomic analyses. Pharminator's main limitation is the lack of appropriate data for populating CPTs; this limitation may be leveraged to optimize company-specific implementations of Pharminator. Acquisition of appropriate data, validation and optimization are the aims that will be addressed during a Sloan Industry Fellowship at the Program on the Pharmaceutical Industry at MIT.

## F. References


1. DiMasi, J.A., Hansen, R.W. & Grabowski, H.G. The price of innovation: new estimates of drug development costs. *J Health Econ* **22**, 151-85 (2003).
2. DiMasi, J.A., Hansen, R.W., Grabowski, H.G. & Lasagna, L. Cost of innovation in the pharmaceutical industry. *J Health Econ* **10**, 107-42 (1991).
3. DiMasi, J.A., Hansen, R.W., Grabowski, H.G. & Lasagna, L. Research and development costs for new drugs by therapeutic category. A study of the US pharmaceutical industry. *Pharmacoeconomics* 7, 152-69 (1995).
4. Frantz, S. & Smith, A. New drug approvals for 2002. *Nat Rev Drug Discov* **2**, 95-6 (2003).
5. DiMasi, J.A. Risks in new drug development: approval success rates for investigational drugs. *Clin Pharmacol Ther* **69**, 297-307 (2001).
6. DiMasi, J.A. The value of improving the productivity of the drug development process: faster times and better decisions. *Pharmacoeconomics* **20 Suppl 3**, 1-10 (2002).
7. Sheck, L. et al. Success rates in the United States drug development system. *Clin Pharmacol Ther* **36**, 574-83 (1984).
8. Bienz-Tadmor, B., Dicerbo, P.A., Tadmor, G. & Lasagna, L. Biopharmaceuticals and conventional drugs: clinical success rates. *Biotechnology (N Y)* **10**, 521-5 (1992).
9. Struck, M.M. Biopharmaceutical R&D success rates and development times. A new analysis provides benchmarks for the future. *Biotechnology (N Y)* **12**, 674-7 (1994).

10. DiMasi, J.A. Success rates for new drugs entering clinical testing in the United States. *Clin Pharmacol Ther* **58**, 1-14 (1995).

11. Berry, D.A. Interim analyses in clinical trials: classical vs. Bayesian approaches. *Stat Med* **4**, 521-6 (1985).

12. Berry, D.A. & Ho, C.H. One-sided sequential stopping boundaries for clinical trials: a decision-theoretic approach. *Biometrics* **44**, 219-27 (1988).

13. Berry, D.A. A case for Bayesianism in clinical trials. *Stat Med* **12**, 1377-93; discussion 1395-404 (1993).

14. Spiegelhalter, D.J., Freedman, L.S. & Parmar, M.K.B. Applying Bayesian ideas in drug development and clinical trials. *Statistics in Medicine* **12**, 1501-11 (1993).

15. Cressie, N. & Biele, J. A sample-size-optimal Bayesian procedure for sequential pharmaceutical trials. *Biometrics* **50**, 700-11 (1994).

16. Gittins, J.C. & Pezeshk, H. A decision theoretic approach to sample size determination in clinical trials. *J Biopharm Stat* **12**, 535-51 (2002).

17. Backhouse, M.E. An investment appraisal approach to clinical trial design. *Health Econ* 7, 605-19 (1998).

18. Schumacher, G.E. & Barr, J.T. Using population-based serum drug concentration cutoff values to predict toxicity: test performance and limitations compared with Bayesian interpretation. *Clin Pharm* **9**, 788-96 (1990).

19. Johns, D. & Andersen, J.S. Use of predictive probabilities in phase II and phase III clinical trials. *Journal of Biopharmaceutical Statistics* **9**, 67-79 (1999).

20. Pallay, A. A decision analytic approach to a futility analysis of a phase II pharmaceutical study. *Journal of Biopharmaceutical Statistics* **11**, 209-25 (2001).

21. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference,* (Morgan Kaufmann, San Mateo, CA, 1988).

22. Pauker, S.G. & Kassirer, J.P. Decision analysis. *N Engl J Med* **316**, 250-8 (1987).

23. Szolovits, P. Uncertainty and decisions in medical informatics. *Methods Inf Med* **34**, 111-21 (1995).

24. Winston, P.H. & Narasimhan, S. *On To Java,* (Addison-Wesley, Boston, 2001).

25. Bajzar, L., Nesheim, M.E. & Tracy, P.B. The profibrinolytic effect of activated protein C in clots formed from plasma is TAFI-dependent. *Blood* **88**, 2093-100 (1996).

26. Gruber, A. et al. Inhibition of thrombus formation by activated recombinant protein C in a primate model of arterial thrombosis. *Circulation* **82**, 578-85 (1990).

27. Eichacker, P.Q. & Natanson, C. Recombinant human activated protein C in sepsis: inconsistent trial results, an unclear mechanism of action, and safety concerns resulted in labeling restrictions and the need for phase IV trials. *Crit Care Med* **31**, S94-6 (2003).