R&D SENIOR LEADERSHIP BRIEF

A New Tool for Predicting Marketing Approval of Oncology Drugs

Tufts Center for the Study of Drug Development and
Janssen Research & Development

Tufts Center for the
Study of Drug Development
TUFTS UNIVERSITY

R&D SENIOR LEADERSHIP BRIEF

Executive Summary

Despite decades of experience conducting clinical trials and vast improvements in knowledge about disease mechanisms, drug development professionals have been unable to improve the low success rates for new molecular entities (NMEs) and new biologic entities (NBEs). Under the blockbuster drug development paradigm, low success rates account for a large proportion of capitalized cost for each approved compound; but the revenue generated by each commercialized blockbuster drug recoups the high cost of the failures. Under the emerging drug development paradigm, characterized by rising costs for NMEs and NBEs that target stratified patient populations and, therefore, smaller market opportunities, the need to increase new compound success rates and reduce drug development risk intensifies.

To this end, starting in late 2012, the Tufts Center for the Study of Drug Development (Tufts CSDD) and Janssen Research & Development (JRD) collaborated on a pilot study to develop and test a model capable of predicting the likelihood of marketing approval for oncology NMEs and NBEs. The study assessed the predictive power of select new compound characteristics. The result is a simple scoring algorithm (ANDI [Approved New Drug Index]) to assist in the assignment of probability of approval using data available after phase II testing has been conducted.

Oncology drugs that first entered human testing between 1999 and 2007 were analyzed. Logistic regression modeling and machine learning techniques were performed and a four-variable model that is highly predictive of new compound approval was identified.

This Executive Management Brief discusses the pilot study methodology and results in more detail. Tufts CSDD and JRD are now looking to engage a larger number of companies to join in a second study aimed at both refining the current model and expanding the approach to other therapeutic areas.

Problem Statement

The underlying probability of success for the approval of an oncology drug that enters phase III is no better than 50-50, and has not improved in decades despite an explosion in knowledge about the molecular basis of disease. Thus, decisions about the advancement of oncology drugs to phase III are subject to what the behavioral economist Daniel Kahneman describes as

> ***"a puzzling limitation of our mind: our excessive confidence in what we believe we know and our apparent inability to acknowledge the full extent of our ignorance and the uncertainty of the world we live in. We are prone to overestimate how much we understand and to underestimate the role of chance in events."*** *(Thinking Fast and Slow, 2011, pp. 13-14)*

Predictive algorithms can help decision makers overcome these biases. In clinical practice, even the simplest algorithms, such as the APGAR score to evaluate newborns in the delivery room, have been shown to be equal to or better than the judgment of experienced clinicians. The problem for drug developers is that no such algorithm exists for assessing the probability of success for approval at the end of phase II.

From Simple Model to a Probabilistic Algorithm

Working under the hypothesis that a simple model of decision making would perform as well as algorithms used in the clinical setting, oncology drugs developed by Janssen Research & Development (JRD) over the past 10 years were ranked according to a simple APGAR-type score (see Table1). This exercise demonstrated that a simple model could be used to predict the likelihood of approval.

Table 1: Initial Model of the APGAR-Type Score for Phase II-III Decisions in Oncology

	Score		
Factor	0	1	2
Activity	0-29%	30-49%	50% or higher
Pt number in pivotal phase II study	0-34	35-74	>75
Genomic-molecular basis of MOA	Makes no sense	Makes some sense	Clear target drug stoichiometry or validated biomarker
Discontinuation due to adverse event	>15%	6-15%	0-5%
Randomized study data	None	HR>0.67	HR=0.67 or less

In this scoring system, variables were picked based on experience and reason (i.e. intuition) and weighted equally.

On the basis of this finding, a collaboration was formed between JRD and the Tufts Center for the Study of Drug Development (Tufts CSDD). The goal of the collaboration was to build an algorithm that assigns a probability of approval to oncology drugs that have completed phase II of clinical development. The prototype tool should

- help to eliminate bias in decision making;
- weight predictive factors on the basis of underlying probabilities, if necessary;
- hasten the rejection of and lessen the work on candidates with a low predicted probability; and
- allow for comparison to norms when only one project is under consideration, because superior value is usually easier to identify when comparing choices.

The key secondary objective was to compare the ease of use and accuracy of various predictive modeling techniques: APGAR-type scores in which variables are scored and then weighted equally, logistic regression in which weights are assigned, regression tree models in which cuts are determined by the model, and machine learning techniques. We also assessed the feasibility of obtaining the required data from sources that are available to anyone, either for free or by subscription. From the outset, the intent was to share freely the findings of the study, and any subsequent research, with the academic community, the pharmaceutical industry and the public at large.

Pilot Study Methodology

An initial list of 98 oncology drugs was compiled from a Tufts CSDD database of drugs and biologics that first entered clinical development under the sponsorship of one of the top 50 (by 2006 pharmaceutical sales) pharmaceutical companies. The list includes drugs that entered development between 1999 and 2007, and for which entry into phase II was recorded. For each of these drugs, data on over 30 variables—clinical safety and efficacy, preclinical efficacy, operational and market factors—were gathered from subscription-based and publically available sources.

Logistic regression modeling was used to assess the relationship between the odds of approval and each variable. The variables with the strongest relationship to approval were included in multivariable logistic regression models. Machine learning techniques—neural network, random forest and recursive partitioning tree—were employed to identify the most important variables using an unbiased approach to validate and refine the logistic regression results. Using these techniques, variations of the leave-some-out method were used in lieu of application of the model to a separate validation data set. Finally, classification and regression tree (CART) analysis was used to confirm the key variables and to better define cut points for dichotomous variables and for use in a simple APGAR-type scoring system. Analysis of the area under the receiver operating characteristic (AUROC) curve was used to select the final scoring system that provides the most accurate classification of compounds into successes and failures; and to assess whether the weighted (logistic regression)

analysis provides more useful information to the decision maker than a simple, unweighted scoring system like APGAR.

Pilot Study Findings

The results from all of the techniques converge around variables that describe anti-tumor activity, the certainty of phase II results (i.e. the number of patients included in phase II studies), the rapidity with which phase II was completed, and the size of the patient population under consideration. The bearing of the first two variables on the success of the drug are self-evident. The rapidity with which phase II was completed may be, among other things, an indicator of operational excellence or an indication that the treating community recognizes the drug has the potential to make a significant contribution to patient care. The size of the disease population under study is inversely related to success, which may indicate the lack of a recognized standard of care or a lower hurdle for regulatory approval or both.

Activity, as measured by response rate (partial or complete; or as otherwise defined by the protocol, but excluding stable disease), was the variable that was most strongly associated with approval. When a cutoff of 20% was used, the odds ratio was approximately 20 for high versus low response rate. The influence of response on approval was observed in every model; in every model the effect was large; and the response rate associated with a much increased odds of approval is not that high. Other variables contributed to the ability to classify drugs according to the probability of success, but not to the same extent as the response rate.

Using the results of all of the analyses, we constructed a modified scoring system (ANDI [Approved New Drug Index]) using four variables and scoring each one as 0, 1, or 2. The ANDI scoring system is shown in Table 2. For this model, in which variables are assigned equal weighting, the area under the ROC curve was as high as it was for any of the weighted techniques.

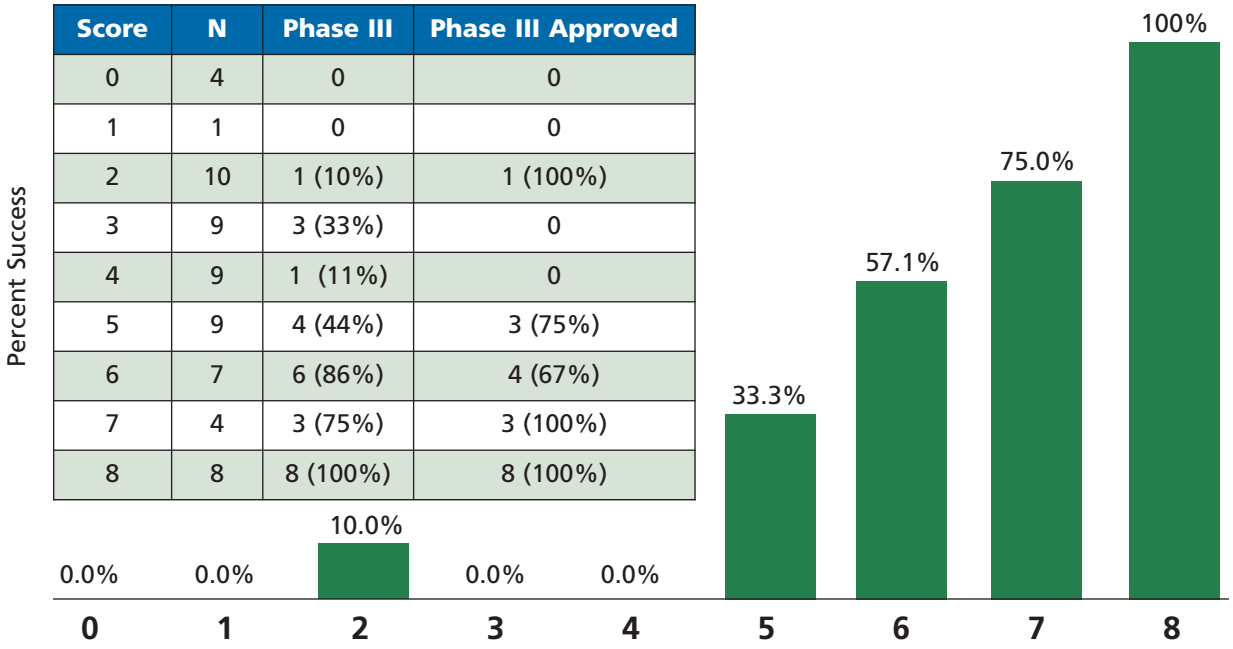
Table 2. ANDI Scoring Cut-Offs for Factors Predictive of Approval After Phase II

	Score		
Factor	0	1	2
Activity	<3% or negative randomized phase 2 study	3-13.8%	>13.8% or positive randomized phase 2 study
Number of patients in pivotal phase II study	<=37	38-49	>50
Number of patients with disease treated (worldwide)	>302,000	50,000-302,000	<50,000
Phase II duration	>44 mos	21-44 mos	<21 mos

Complete data were available for 61 of the drugs in the data base. These were scored according to the factors included in the ANDI scale. The results are shown in Figure 1.

Figure 1. ANDI 4-Factor Scores and Percent Success

Score	N	Phase III	Phase III Approved
0	4	0	0
1	1	0	0
2	10	1 (10%)	1 (100%)
3	9	3 (33%)	0
4	9	1 (11%)	0
5	9	4 (44%)	3 (75%)
6	7	6 (86%)	4 (67%)
7	4	3 (75%)	3 (100%)
8	8	8 (100%)	8 (100%)



Factors: activity, pivotal trial patients, patients treated worldwide, Phase II duration

The overall approval rate (bar graph) at various ANDI scores and the approval rate for those drugs that entered phase III (embedded table) are shown in the figure. Among the companies that are included in the data base, decisions about advancing drugs with scores of 0 through 4 reflect the low probability that these drugs will be approved. Similarly, there seems to be no problem with identifying oncology drugs with a very high probability of approval.

Pilot Study Implications

For oncology drug development the implications for decision making are clear, especially with regard to the assignment of risk-adjusted value. These data support the assignment of much higher than average (i.e. the industry metric) probabilities of success—92% overall—to drugs with high scores (7 and 8); and, much lower probabilities of success to compounds with scores of 0-4, a range in which only 3% of the drugs ultimately are approved and 20% of those that entered phase III ultimately are approved.

Future activities

Based on our initial findings for oncology drugs, we propose to continue this research as a collaborative effort among companies that are willing to contribute data and knowledge in order to refine the model and expand the data base for oncology drugs and build models in the CNS area, on which work is already well underway, and other therapeutic areas. Second, we want to educate decision makers on the usefulness of such models, especially when some might consider them as a limit to intuition and choice or a hindrance to risk taking. Finally, we will use the results of this research, the ANDI score, as a dependent variable in models that relate pre-clinical information to the probability of approval, in an attempt to address the extremely low probability that drugs entering phase II will be approved. We hope that a tool such as the ANDI score will allow us to make better decisions about the advancement of oncology drugs to phase III. If this is accomplished, then patients will benefit from the more efficient development of oncology drugs, the health care system will benefit from a lower cost of failure, and the industry will benefit from a more effective allocation of resources.

If your organization is interested in more information about participating in this working group study, please contact the Tufts Center for the Study of Drug Development.

Tufts Center for the Study of Drug Development

TUFTS UNIVERSITY

75 Kneeland Street, Suite 1100
Boston, MA 02111 USA

Tel 617-636-2170
Fax 617-636-2425
Email csdd@tufts.edu
Web http://csdd.tufts.edu

The research in this brief was conducted jointly by Tufts Center for the Study of Drug Development and Janssen Research & Development. Tufts CSDD Senior Leadership Brief is published by the Tufts Center for the Study of Drug Development; Kenneth I Kaitin, PhD, editor.

About Tufts CSDD Sponsored Research Projects

Our Sponsored Research projects typically focus on operational and R&D management issues, and may entail creation of a multi-company working group, survey creation, leadership/management meetings, collection of company data, and publication of aggregate analyses summarizing key results and insights and benchmark comparisons when applicable. Tufts CSDD also hosts special Senior Leadership Roundtables offering a unique opportunity for senior pharmaceutical and biopharmaceutical leaders to engage in frank and open discussion on a wide variety of strategic issues and to share ideas with industry peers and regulators in a neutral setting.

Visit our website at http://csdd.tufts.edu/sponsored_research to learn how to collaborate with Tufts CSDD on a research initiative and to participate in our sponsored research programs.

Kenneth Getz, MBA
Director of CSDD Sponsored Research Programs
Associate Professor
Tufts Center for the Study of Drug Development

Kenneth I Kaitin, PhD
Director and Professor
Tufts Center for the Study of Drug Development

About Tufts Center for the Study of Drug Development

The Tufts Center for the Study of Drug Development at Tufts University provides strategic information to help drug developers, regulators, and policy makers improve the quality and efficiency of pharmaceutical development, review, and utilization. Tufts CSDD conducts a wide range of in-depth analyses on pharmaceutical issues and hosts symposia, workshops, and public forums.

Tufts Center for the Study of Drug Development

TUFTS UNIVERSITY

75 Kneeland Street, Suite 1100
Boston, MA 02111 USA

Tel 617-636-2170
Fax 617-636-2425
Email csdd@tufts.edu
Web http://csdd.tufts.edu