Predictive Modelling in Drug Development

Predictive drug development modelling is dependent on the quality and quantity of data on which models are constructed and tested; the industry-wide sharing of data about past development failures, as well as successes, would empower predictive modelling to the benefit of all.

By at Asher D. Schachter at Phorecaster

Asher D. Schachter, MD, MMSc, MS, is a Bioinformatician in the Children’s Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology, a paediatric nephrologist at the Children’s Hospital, Boston, and an Assistant Professor of Pediatrics at Harvard Medical School. He graduated from the University of Toronto Medical School in 1992, and completed his paediatric residency at Toronto’s Hospital for Sick Children. He is board-certified by the American Board of Pediatrics in both paediatrics and paediatric nephrology, and in 2005, was nominated to membership of the Society for Pediatric Research. Dr Schachter’s research focuses on predictive modelling in drug development with the primary goals of increasing the safety and effectiveness of new drugs, and reducing drug development costs. He is an advocate for safe and effective drug development for the paediatric population. His collaborative work with the MIT Center for Biomedical Innovation will focus on the optimisation of pharmaceutical industry R&D productivity.

A significant proportion of the current estimated cost of developing a new drug stems from later phase trials of drugs that ultimately fail to achieve FDA approval. More important than the financial costs is the impact on human subjects in clinical trials who are exposed to drugs that are later determined to be unsafe and/or ineffective. Furthermore, the resources that are invested in conducting clinical trials of eventual failures could be put to better use to develop safer, more effective drugs, if only those drugs could be identified with greater accuracy.

Predictive models can guide compound selection from high-throughput screens at the preclinical phases. Therefore, accurate prediction of new drug success would reduce patient exposure to unsafe drugs, facilitate redirection of research and development resources to trials more likely to succeed, guide adaptive trials, and reduce drug development costs. In addition, investors who rely on market forecasting models operate on the assumption that the drug in question will achieve approval by the US FDA and other regulatory agencies. Methods that predict new drug success would therefore empower market forecasts.

DRUG DEVELOPMENT COSTS

The estimated cost of developing a new drug into a marketable therapeutic agent is reported to be in the range of $802,000,000 (1). The cost of developing a new drug that will ultimately fail is directly proportional to the length of time between IND approval and termination of development. The patent life of a given new drug typically begins at the time of IND approval and lasts for 15 to 20 years, but a financial return does not commence until NDA approval is granted and may be short-lived if competitors release similar agents. Analyses of the distribution of research terminations by clinical phase have shown that over 60% of terminations occur during phases II and III (2) – that is, later in the drug development process. It follows that earlier termination of new drugs 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.

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 requires a delicate balance between financial constraints, proper conduct of clinical trials, good clinical practices, and ensuring that regulatory requirements for approval will be met. However, over-zealous termination of new drugs will impede the development of innovative, breakthrough therapies. The decision process must balance the cost of terminating what would be a successful new drug against allowing an eventual failure to proceed through Phase III. It is important to note the difference between new drug development programmes versus the development of more efficient medications. The latter involves improving on existing, 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 a new drug, and is rarely a major source of lost revenue.

The pharmaceutical industry is, in general, aware of this dilemma and is apparently attempting to adopt models to
Innovations in Pharmaceutical Technology

The adoption of combinatorial chemistry and high-throughput screening for potential new drugs provides exponentially more lead compounds, but also significantly increases the number of early-phase compounds under consideration for further costly development in human clinical trials. Without accurate predictive models, high throughput screens may aggravate the problem of selecting would-be successful drugs from an even larger pool of potential drugs.

The transition from preclinical drug screens to clinical trials has several challenges. Animal models of the target disease may not reliably represent the human form of the disease, which may be heterogeneous in nature. Polypharmacy adds another level of uncertainty in the form of unknown drug interactions with the new drug. Issues with specific drug screening platforms include inter-platform and inter-platform variability, limited database annotation, lack of functional annotation, lack of robust ontologies, complex biological systems with compensatory pathways that circumvent a drug’s effect, gene-gene interactions, and complex gene pathways and mechanisms such as epigenetic gene regulation. All these variables, and others, combine to add multiple levels of complexity to the goal of high throughput identification of safe and effective drugs.

PREDICTIVE MODELS

There are several established approaches to building predictive models, including standard statistical methods and artificial intelligence techniques. While each type of model has characteristics that are well suited to addressing specific questions, any predictive model is only as good as the quality and quantity of the data on which it is constructed. Furthermore, the value of a model can only truly be determined through validation studies on independent test datasets. The reliability of these validation studies also depends on the quality and quantity of the independent test data. Training predictive models requires presenting the model with diverse situations within the domain of interest. In the case of modelling new drugs in development, this means that model construction requires diverse data not only from drugs that have successfully passed FDA scrutiny, but also drugs that have either failed to achieve approval or that were terminated by the sponsor company at some point during the clinical trial phases. Multiply this by the need for data specific to each drug’s therapeutic class and indication, and the extent of the ideal amount of data becomes apparent. Industry-wide sharing of data would certainly advance ongoing efforts toward this goal.

Unfortunately, the pharmaceutical industry has a tendency to bury ‘negative’ data, which has intrinsic value in terms of learning from prior experience. There are several possible reasons for this. It is human nature to focus on winners and ignore losers. For example, stock investors are much more likely to think about and discuss their winning investments, and avoid acknowledging their losing stocks. There is also likely to be some concern that release of less than impressive data will result in reduced confidence by investors in the pharmaceutical industry. There may also be fear that competitors will gain an edge, or that the company will be put at risk of litigation.

However, pharmaceutical and biotechnology companies have more to gain than lose from industry-wide sharing of data (3). Learning from prior failures can markedly improve the accuracy of predictions, particularly for drugs in the same therapeutic class. A single company would essentially relinquish data in exchange for the benefit of experience from multiple other companies, in an ethos of ‘one for all and all for one’. ‘De-identification’ of drugs and/or sponsor companies could address concerns about lost competitive advantage, since the identities of neither the compound nor the company is necessary for predictive modelling in this context. The Massachusetts Institute of Technology’s Center for Biomedical Innovation (http://web.mit.edu/cbi/) is working to create a ‘safe haven’ to facilitate sharing of information between industry, academics and government regulators. In addition to the benefit of improved productivity, the economic benefits are not insignificant, as discussed in the next section.

ECONOMIC IMPACT OF A PREDICTIVE MODEL

In collaboration with the Program on the Pharmaceutical Industry at the Massachusetts Institute of Technology, I have reported the potential economic benefits of a Bayesian model that predicts the probability of clinical success for a new drug (4). Marco Ramoni, an expert in Bayesian modelling, collaborated with me in order to
develop a model that predicts the success of a new drug given: (a) prior information on successful and failed compounds in the same class; and (b) pre-Phase III safety and efficacy data for the new drug in question. For the most part, we utilised publicly available information from the Tuft Center for the Study of Drug Development and other diverse sources to train and test the model. We found that the model could predict success or failure beyond Phase III with impressive accuracy.

Next, we aimed to determine the economic impact that would result from applying this model to a virtual pharmaceutical company. We utilised Monte Carlo simulation methods and publicly available detailed data on expenditures and revenues across the pharmaceutical industry to compare our virtual pharmaceutical company’s performance with the recent, actual performance across the pharmaceutical industry. We found that our model reduced the cost of developing a new drug by an average of $283,000,000 per compound in comparison with the pharmaceutical industry at large. In addition, our model increased revenues by an average of $160,000,000 per new drug in comparison with the industry norm. While the reduction in costs was the direct result of earlier termination of eventual failures, the increase in revenues was due to our model’s improved false negative rate – that is, the continued development of more would-be successes that would otherwise have been terminated. Although it is impossible to know the true extent to which the pharmaceutical industry is terminating would-be successful new drugs, we configured the Monte Carlo simulation to account for both pessimistic and optimistic distributions.

The bottom line is that there is, in all likelihood, a lot of ‘low hanging fruit’ in the form of reducing even a fraction of erroneous decisions with significant clinical and financial benefits. Predicting on the basis of pre-clinical data alone presents an even greater challenge with substantially more potential reward.

PREDICTIVE MODELS AND PAEDIATRIC DRUG DEVELOPMENT

Although the US FDA was initially created in large part as a result of tragic events involving children, the paediatric population receives inadequate attention in the drug development domain. However over the past ten years, legislators have sought to improve drug development for children with the introduction of incentives that grant six-month patent extensions to companies that perform studies providing sufficient data for paediatric labeling. Li et al (5) recently demonstrated that this incentive is profitable for companies, and advocates renewal of the legislation, which is set to expire this year. However there is still a reluctance to include children in clinical trials prior to NDA approval. In the absence of approval for drug use in children, off-label prescribing predominates clinical practices, with a significant increase in adverse events as a result (6,7). Established pharmacokinetic/pharmacodynamic models in conjunction with models that are accurate predictors of safety and toxicity, and enhanced Phase IV surveillance technologies could improve the industry’s comfort level for including children in pre-NDA clinical trials.

CONCLUSION

Predictive modelling is dependent on the quality and quantity of data on which models are constructed and tested. The information, data and expertise all currently exist to markedly improve the efficiency of the development of safe and effective drugs. The benefits to all involved parties far outweigh any risks, either real or imagined. Ongoing efforts that facilitate industry-wide sharing of diverse data from past and present pre-clinical and clinical trial development programmes will empower predictive modelling to the benefit of all.

The author can be contacted at asher@phorecaster.com

References