Quantitative risk modelling for new pharmaceutical compounds

Zhengru Tang, Mark J. Taylor, Paulo Lisboa and Mark Dyas

The process of discovering and developing new drugs is long, costly and risk-laden. Faced with a wealth of newly discovered compounds, industrial scientists need to target resources carefully to discern the key attributes of a drug candidate and to make informed decisions. Here, we describe a quantitative approach to modelling the risk associated with drug development as a tool for scenario analysis concerning the probability of success of a compound as a potential pharmaceutical agent. We bring together the three strands of manufacture, clinical effectiveness and financial returns. This approach involves the application of a Bayesian Network. A simulation model is demonstrated with an implementation in MS Excel using the modelling engine Crystal Ball.

Recent advances in combinatorial chemistry and HTS techniques have vastly increased the number of new molecules that can be generated and tested. The challenge to the pharmaceutical industry is to identify the potentially useful ones that should be advanced to clinical evaluation and commercial development. The ability to identify, rapidly and accurately, the issues associated with the ongoing development of a given candidate are crucial to financial success, yet pharmaceutical R&D is subject to extreme attrition because the vast majority of drug candidates discovered will be discarded at some point during development [1]. There are many reasons for stopping the development of a particular drug and although some, such as safety, are relatively 'black or white', others, such as commercial viability, are grey in nature.

Those compounds emerging from the initial screening process proceed through a sequence of developmental steps, each of which ends with go or no-go decision events. As development progresses, the steps become larger in scope, longer in duration and significantly more costly to complete than the preceding ones. Thus, there is significant pressure to make accurate go or no-go decisions at the end of each step to avoid further commitment of resources to a programme that might ultimately fail before commercial realization. Cocchetto and Nardi [2] commented that successful drug development planning should result in either the drug being approved or justifiably terminated after the least investment possible.

Termination of development is not always a negative result; there are many reasons for a company to cease a development programme, not all of which are necessarily fatal for the drug. For example, a drug that shows poor efficacy against the chosen disease for which it was originally developed could have excellent efficacy against another disease. In this case, the discoverer can exploit their investment by selling the intellectual property (IP) to another company. Smaller companies might not have the finances to take a drug through the full development process or might not have the infrastructure to take a drug from discovery to launch. Essentially, IP is their product and they must negotiate with larger firms to realize the greatest return on their investment by licensing or selling the IP. For such companies, the

Zhengru Tang
Mark J. Taylor*
Paulo Lisboa
School of Computing and Mathematical Sciences, Liverpool John Moores University, Liverpool L3 3AF, UK
Mark Dyas
School of Pharmacy and Chemistry, Liverpool John Moores University, Liverpool L3 3AF, UK
*e-mail: m.j.taylor@livjm.ac.uk

1359-6446/05/$ – see front matter PII:S1359-6446(05)03606-8

challenge is to balance the escalating costs of taking a drug to launch against the exponential rise in its value the further it progresses.

Risk assessment involves the systematic identification and quantitative evaluation of potential costs and benefits to control, and thus minimize, risk exposure. In this article, this approach is extended towards enabling a quantitative scenario analysis to support the decision-maker in exploring the sensitivity of the process to changes in individual elements, as well as balancing probabilistic estimates of risk of failure against potential financial returns.

The research reported in this paper addresses the go or no-go judgements that are required during pharmaceutical development. Monte Carlo simulation, as proposed, offers companies the possibility of quantifying estimates of the likelihood of successful development, time to market, development costs and financial returns, offering the potential to optimize expenditure on drug development [3].

We describe how Bayesian networks and Monte Carlo simulation can be used together to simulate the probability of success of potential new drugs. Spiegelhalter *et al.* advocated that a Bayesian perspective be adopted for the process of drug development as a means of providing more appropriate and useful tools for aiding decisions regarding whether to continue at each stage of a drug development programme [4]. Goggin *et al.* stated that modelling and simulation have a strong mathematical foundation and have been used extensively in other areas, such as the automobile and aerospace industries, in addition to the pharmaceutical industry [5]. However, Rooney *et al.* commented that the use of simulation is only now beginning to filter through into the mainstream pharmaceutical sector [6]. Bayesian analysis makes a much wider use of probability distributions than traditional statistical methods, in that not only are sampling distributions required for summaries of data but they are also used to represent prior opinion about proportions, event rates and other unknown quantities. Therefore, the shapes of distributions become particularly important because they represent the plausibility of occurrence of values in different ranges [7]. Monte Carlo simulation, the repeated random generation of values for uncertain variables to simulate a model, was used to populate the probability distributions contained within the model described here.

Successful pharmaceutical products undergo three major steps in their life: discovery, development and commercialization [8]. The discovery and development of a new drug could typically take 12 years and cost US$800 million [9]. Drug discovery is relatively cheap when expressed as cost per compound and the challenge is to sift through the data from many hundreds of thousands of simple tests to identify and eliminate the compounds lacking affinity for the chosen target. In drug development, large sums of money and resources are committed to the relatively few compounds fulfilling potency, selectivity, ADME and physiochemical criteria. These are subjected to extensive preclinical and clinical testing to identify those most likely to be clinically effective.

A candidate drug is identified by this process and is then subjected to clinical trial, generally this first phase would be in healthy volunteers (Phase I) and later in progressively larger patient populations (Phases II and III). The decision to take a candidate drug forward for Phase II and III clinical evaluation is crucial because these stages are extremely costly and might be of long duration [10].

The drug discovery process can be divided into six stages, each of which carries its own specific risks:

Drug discovery

Following successful target evaluation and validation, compound libraries would be screened using a validated assay to establish compounds with arbitrarily decided potency. Hutchinson commented that of perhaps 10 000 molecules that might be screened as a potential new drug, only 15 or so might be worthy of further study [11]. These phases are often referred to as hit identification.

Preclinical studies

The hits identified in the previous process are advanced through lead identification and lead optimization phases to identify a candidate with the desired efficacy (*in vitro* and *in vivo*), acceptable ADME properties and without detectable toxicity.

Manufacturing

Risks associated with scaling-up from the pilot-scale batches used to support preclinical studies to the larger requirements for clinical trials and eventually to the high-scale production required for commercial supply can include insufficient quality control and inefficient production control [12], in addition to potential synthetic issues.

Clinical trials stage

Candidates can fail for several reasons, including lack of efficacy, toxicity, inadequate clinical trial design, inappropriate study populations and inappropriate dosage levels [3,13]. Ultimately, if the compound lacks efficacy or is inferior to competitive products, it could be terminated and the project could return to the discovery stage [8].

Regulatory evaluation stage

Drug registration and licensing is a costly and lengthy process involving the regulatory review of a vast amount of data. Risk factors can include inappropriate or insufficient data, as well as recent changes in medical or regulatory opinion.

Marketing stage

After commercial launch, a new product will be subject to ongoing post-market surveillance by regulatory agencies [14]. Risks include the appearance of unpredictable, rare adverse reactions, poor market information, changes to clinical

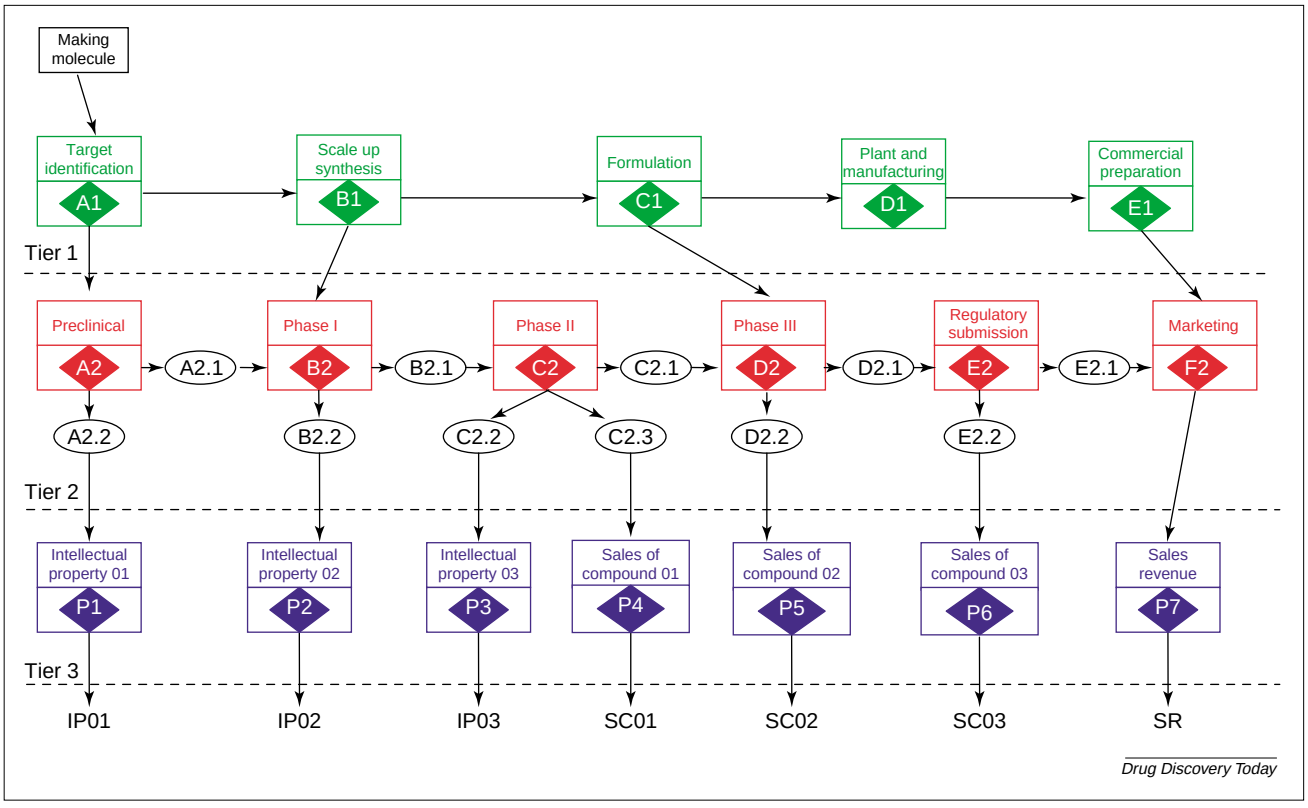

FIGURE 1

Directed graph of the drug discovery and development process showing the possible paths through the process and their respective probabilities. Each stage of the process is represented by a rectangle. Decision points with an associated probability relating to continuation of the process to a further stage, are represented by a diamond. Oval shapes represent management decisions with associated probability distributions.

practise and the appearance on the market of competitor drugs, all of which will impact sales.

Although the programme of clinical studies is probably the most pivotal aspect of drug development, there are other key activities that proceed concurrently with the clinical trials, such as formulation development and scale-up of bulk drug manufacture.

Walley *et al.* [15] and Bingham *et al.* [16] described the use of Bayesian approaches to pharmacoeconomic analysis, which concerns health economics in the field of pharmaceuticals and pharmaceutical policies. Modelling and simulation using Bayesian analysis have also been applied to the clinical trials phase of drug development for pharmocokinetic and pharmocodynamic modelling [3,5,7]. Miller argued that the use of pharmacoeconomic analyses early on in the development process (using clinical trial simulation, option pricing, investment appraisal, threshold analysis and value of information analysis techniques) can provide useful input into the design of clinical development programmes, portfolio management and optimal pricing strategy [17].

Bayesian approaches

What advantages do Bayesian approaches offer?

A Bayesian network is a probabilistic chain model where probability distributions assigned to nodes in a chart structure can propagate according to the standard rules of conditional probability inference. In this way, the probability of observing a given outcome at a node down the processing pipeline can be calculated from knowledge of the states of one or more states of nodes upstream in the process, using the sequence of cause and effect in the process structure. From a mathematical point of view, the proposed model in Figure 1 forms an acyclic graph where the propagation of probabilities from node to node is straightforward to calculate analytically.

By using distributions for each nodal state, whether this is the probability of completing that stage of the process successfully, the time required for completion or the task cost, it is possible to encode the uncertainty about the knowledge that is fed into the model. A major strength of this approach is that it will predict a distribution of outcomes for any given node in the model, rather than making point estimates from what is reasonably expected to be a broad, and sometimes multimodal, distribution of expected values.

The Bayesian network for Figure 1 is defined by a pair of knowledge parameters representing structure and local probability distributions at each node. The graphical structure encodes conditional independence assertions about the nature of the process being modelled, by specifying the causal relationships between the process elements, while the probability distributions are defined conditionally at each node in the process acyclic graph. Bayesian networks were used because they offer a comprehensive and robust approach to model estimation, particularly for clinical trials [18].

BOX 1

Model inputs[a]

- D_TI is the duration of the target identification stage, in months. A similar notation is used for the other stages
- C_TI is the cost of the target identification stage, in US$ millions. A similar notation is used for the other stages
- A1 = Probability (TI successfully completed – TI commenced)
  Similarly for A2, B1, B2, C1, C2, D1, D2, E1, E2, F1 and F2
- A2.1 = Probability (deciding to continue to PhI – PC successfully completed)
- A2.2 = Probability (deciding to sell IP – PC successfully completed)
- A2.1 + A2.2 = 1 (because these are two mutually exclusive probabilities)
  Similarly for B2.1 + B2.2, C2.1 + C2.2 + C2.3, D2.1 + D2.2, E2.1 + E2.2
- P1, P2 and P3 are the probabilities of successfully selling IP at each of three stages, given that the choice was made to sell IP at that stage (e.g. immediately after completing the preclinical trials)
- P4, P5 and P6 are the probabilities of successfully achieving sales of the compound at each of three stages given successful completion of Phase II clinical trials
- P7 is the probability of successfully achieving sales revenue for the compound
- Rev_IP01 is the revenue collected from selling IP after choosing to do so
  Similarly for Rev_IP02, Rev_IP03, Rev_SC01, Rev_SC02, Rev_SC03, Rev_SR
- No revenue is collected during the discovery and development process until we choose to sell IP

[a]Model inputs include the duration and cost of the different stages and the probabilities of completing the different stages to realize the potential revenues

BOX 2

Model outputs[a]

- Start_TI and Time_TI are the start time and finish time of the target identification stage
- Cost_TI is the money spent to successfully finish the target identification stage
- C_NPV_TI is the net present value of cost_TI
- Pro_TI is the completion probability of TI stage
- Exp_Rev_IP01 is the expected revenue of IP01
  Similarly for the other stages
- There are seven different ways of potentially making a profit through the candidate compound
- IP01: sale of compound patent at intellectual property 01 stage, after preclinical stage
- IP02: sale of compound patent at intellectual property 02 stage, after completing Phase I stage
- IP03: sale of compound patent at intellectual property 03 stage, after completing Phase II stage
- SC01: sales of compound at sell compound 01 stage, after completing Phase II stage
- SC02: sales of compound at sell compound 02 stage, after completing Phase III stage
- SC03: sales of compound at sell compound 03 stage, after completing regulatory submission stage
- SR: make mature sales of compound at sales revenue stage, after completing marketing stage
- The model outputs apply to all the endpoints

[a]Model outputs include the time, cost, net present value, probability and expected revenue for the different stages

A Bayesian approach enables synthesis of all the available information (i.e. using existing knowledge as well as data about a candidate drug). This information is encoded by the form of the assumed nodal probability distributions (normal, log-normal, multinomial and so on), which, together with the structure, define the joint probability of all states for all nodes in the model. The analytical power of this approach is that from this joint distribution it is possible to query the possible outcomes at every node in the network using any prior knowledge about observed states at any node.

A powerful feature of Bayesian models is that they support backward chaining to identify the possible causes of results by estimating the expected state of nodes upstream in the model, given desired outcomes further downstream. Moreover, a sensitivity analysis can be performed to identify which nodes have the strongest effect on the outcome performance metrics of the model.

In this simulation framework it is possible to consider scenarios obtained by varying the conditional probabilities specified at each node, and to assess their impact on the bottom line figures at the exit points in the development process. Given a range of candidate compounds for development, the proposed framework provides a quantitative model of the effects of the different scenarios on the relative ranking of their cost-benefit figures. The model was implemented in an Excel and Crystal Ball spreadsheet.

Although there has been application of Bayesian analysis techniques at both the macro scale of pharmaceutical activities (in terms of pharmocoeconomics) and at the micro scale (in terms of pharmocokinetic and pharmacodynamic modelling of particular compounds for clinical trials studies), little modelling and simulation appears to have been applied to the mid-range process of drug discovery and development

Monte Carlo approaches

In the current implementation, the conditional probabilities are specified by expert users and are varied as part of the scenario analysis. It is straightforward to set these distributions on the basis of prior knowledge derived from historical data, where they are available, just as the distribution parameters for the model can, in principle, be estimated directly from the data.

Integrating over the multiple probability distributions defined at each node in the process pipeline, which are the degrees of freedom of the model, cannot be done analytically. However, this integral is efficiently conducted using Monte Carlo methods, where the nodal probability distributions are sampled repeatedly until the outcome distributions in the nodes at the end of the pipeline stabilize [19,20]. Formally, the Monte Carlo process for forward propagation of conditional probabilities across the process pipeline consists of drawing random samples from each of the conditional distributions specified locally at each node in the graph, whose product along the directions of the arrows then generates the required probabilistic values for the likelihood of achieving the model outcomes successfully at each of the possible exit points.

The main risk elements in the development process of a new pharmacological compound are represented by the directed acyclic graph shown in Figure 1. Mathematically,

TABLE 1

Entries in the Excel and Crystal Ball spreadsheet model

Compound A	Duration (months)	Time to start (months)	Time to complete (months)	Cost (m)	Total cost (m)	Revenue (m)	Monthly discount rate (%)	NPV cost (m)	NPV revenue (m)	NPV profit (m)	Transitional probability (%)	Compound probability (%)
Target identification	12	0	12	5	5	NA	0.01	5.00	0	−5.00	90	90.00
Scale up synthesis	12	12	24	5	10	NA	0.01	9.98	0	−9.98	90	81.00
Formulation	12	24	36	10	20	NA	0.01	19.95	0	−19.95	90	72.90
Plant and manufacturing	14	36	50	10	30	NA	0.01	29.89	0	−29.89	90	65.61
Commercial preparation	12	50	62	2	32	NA	0.01	31.84	0	−31.84	90	59.04
PC	20	12	32	35	45	NA	0.01	44.94	0	−44.94	90	81.00
PC outcome 1											40	
PC outcome 2											60	
Phase I	10	32	42	4	59	NA	0.01	58.81	0	−58.81	90	26.24
Phase I outcome 1											50	
Phase I outcome 2											50	
Phase II	12	42	54	8	67	NA	0.01	66.71	0	−66.71	50	6.56
Phase II outcome 1											30	
Phase II outcome 2											20	
Phase II outcome 3											50	
Phase III	30	54	84	12	79	NA	0.01	78.57	0	−78.57	70	1.24
Phase III outcome 1											60	
Phase III outcome 2											40	

Abbreviations: NA, not available; m, US$ millions; PC, preclinical.

the requirement is for an estimation of the probabilities of successfully reaching the leaf nodes in this graph, from which expectations of income generation can be obtained, alongside measures of cost and duration for the process as a whole.

Development of the process model

The process model described in this paper is divided into three main tiers:

- Tier 1: Chemical manufacturing
- Tier 2: Preclinical and clinical trial stages I–IV
- Tier 3: Financial returns

Each stage of the process (Figure 1) is represented by a rectangle [e.g. Target Identification (TI)]. Decision points with an associated probability relating to continuation of the process to a further stage (e.g. A1 = 'does *in vitro* activity occur?') are represented by a diamond. Oval shapes represent management decisions with associated probability distributions (e.g. A2.2 = 'decision to sell the intellectual property rights upon successful completion of preclinical trials').

Model inputs

The inputs required for the model include, for each stage, the expected values of:

- Duration (in months)
- Cost – at current value (in US$ millions)
- Revenue – at current value (in US$ millions)
- Rate – monthly discount rate
- Transitional probability – probability of successful completion of a stage, given completion of the previous stage
- Optional transitional probability – probability of deciding to proceed a stage, given the previous stage being successfully completed.

Box 1 describes the model inputs in more detail.

Model outputs

The outputs generated by the model include, for each stage:

- Start date
- Finish date
- Total cost
- Net present value (NPV) cost
- NPV revenue
- NPV profit
- Completion probability

Box 2 describes the model outputs in more detail.

Algorithm for calculating the cost–benefit ratio

Total cost is the money that a company spent during the whole process, which is the sum of the cost of all the stages. NPV is calculated as follows:

$$NPV = \text{Value in year } t / (1 + r)^{t}$$

Where r is the discount rate, and t is the number of years into the future that the cash flow occurs. The expected NPV for a project is obtained by discounting each cash flow (both negative and positive) and summing the discounted values.

Parameters that need to be specified by the user of the model include: total cost, discount rate and total revenue.

For example, for the stage Intellectual Property01:

TABLE 2

The required inputs for a typical sample compound for the model[a]

Stages	Duration (monthly)	Cost (US$ millions)	Transitional probability (%)	Transitional probability (names)	Revenue (US$ millions)
Target identification	12	5	90	A1	NA
Scale-up synthesis	12	5	90	B1	NA
Formulation	12	10	90	C1	NA
Plant and manufacturing	14	10	90	D1	NA
Commercial preparation	12	2	90	E1	NA
PC	20	35	90	A2	NA
PC outcome 1			40	A2.1	NA
PC outcome 2			60	A2.2	NA
Phase I	10	4	90	B2	NA
Phase I outcome 1			50	B2.1	NA
Phase I outcome 2			50	B2.2	NA
Phase II	12	8	50	C2	NA
Phase II outcome 1			30	C2.1	NA
Phase II outcome 2			20	C2.2	NA
Phase II outcome 3			50	C2.3	NA
Phase III	30	12	70	D2	NA
Phase III outcome 1			60	D2.1	NA
Phase III outcome 2			40	D2.2	NA
RS	12	4	90	E2	NA
RS outcome 1			50	E2.1	NA
RS outcome 2			50	E2.2	NA
Marketing	10	4	90	F2	NA
Intellectual property 01	11	1	90	P1	2
Intellectual property 02	11	1	80	P2	10
Intellectual property 03	11	1	85	P3	100
Sell compound 01	13	1	80	P4	20
Sell compound 02	13	1	75	P5	200
Sell compound 03	13	1	75	P6	800
Sales revenue	14	1	70	P7	1500

[a]Compound A: total time to launch, 10 years; total cost, US$100 million; total revenue, US$1500 million; monthly discount rate, (0.01%)

Abbreviation: NA, not available; PC, preclinical; RS, regulatory submission.

Total cost of IP01 = IP01 cost * Transitional Probability A2.2 + TI cost + Scale-up Synthesis (SS) cost + Formulation (FM) cost + Preclinical Trial (PC) cost

Expected NPV cost of IP01 = IP01 total cost / (1+discount rate) ^ time required thus far to begin IP01

Expected NPV profit of IP01 = Expected NPV revenue – Expected NPV cost

Algorithm for calculating completion time

The total launch time is the time required to launch the drug to market, which is calculated as the sum of successful completion times of each stage.

For example, for the stage PC:

Finish (PC) = duration (PC) + time of completing earlier stages

= D_PC + Max [finish (TI), finish (SS)]

= D_PC + MAX (Finish_TI, D_SS + Finish_ TI)

Algorithm for calculating the probability of success

This follows directly from user-specified transitional probability and optional transitional probability values (e.g. for the stage 'Target Identification'):

Prob (Starting TI|MM=1) = 1

Prob (Completing TI|MM=1) = A1

Therefore, P_TI = P (TI=1|MM=1) = A1*1 = A1

Another example, for the stage Scale-up Synthesis (SS):

Prob (Starting SS|TI=1) = P_TI = A1

Prob (completing SS|TI=1) = B1

Therefore, P_SS = P (SS=1|TI=1) and P (TI=1|MM=1) = B1 * A1 * 1 = B1*A1

An implementation of the model was created using Microsoft Excel and Crystal Ball (Table 1). Monte Carlo simulation is used to dynamically produce alternative scenarios in the spreadsheet model. All uncertain variables or model inputs are modelled using probability distributions. The distribution type of each uncertain variable or model input is defined according to its features.

The Cost and Revenue variables will be a positive number, with no upper limit, such that a given value is more likely to be in the vicinity of the mean than far away, matching the conditions of a lognormal distribution. Thus, cells C_TI, …, C_SR follow a lognormal distribution. The Duration variables will have a finite range, with extreme values less likely to occur. These requirements are met using a triangular distribution, used for cells D_TI, …, D_SR. The transitional probability and optional transitional probability variables are constrained by a finite range. However, in the absence of further information it was decided to initially allow equal opportunity for every value in range to occur, which is modelled by a uniform distribution. Therefore, cells P_TI, …, P_SR use uniform distributions.

The required inputs for a typical sample compound are shown in Table 2. The uncertainty values for each uncertain variable are also entered into the model.

Presentation of output for users of the model

The form of the output of the information produced by the model consists of summary tables (that show the time to launch, total costs, total revenue, NPV cost, NPV revenue and NPV profits, as well as the probability of success for a set of compounds) and bubble charts (that provide a more graphical and easy to interpret view of the information) as a means of presenting the results of the simulations to the target user base (typically research managers) within the pharmaceutical industry. In the bubble charts, the diameter of the bubbles is proportional to the probability

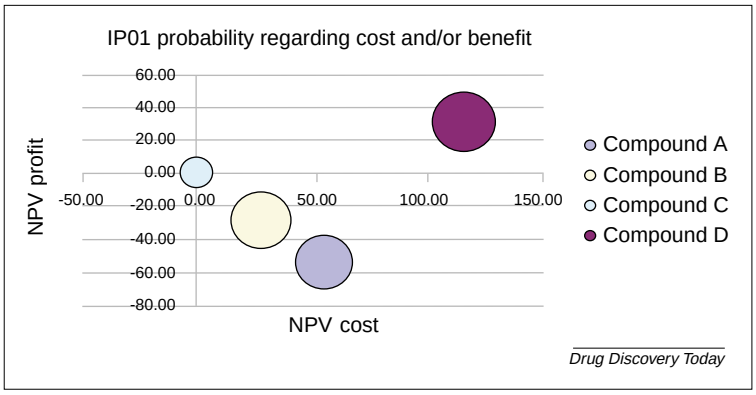

FIGURE 2

Bubble chart produced by the model comparing the balance of benefit to cost for four preferred compounds. The bubble diameter represents the probability of successful completion. The larger the bubble, the higher the probability of success for a given compound.

of successfully completing the development process. The larger the bubble, the higher the probability of success for a given compound.

Figure 2 shows the balance of benefit to cost for four preferred compounds. The position of the bubbles indicates the NPV profit compared to the NPV cost for a set of compounds. Taken together, the size of the bubble and its position summarize the main components of risk, enabling a visual comparison between different candidate compounds for selection. Generally, the higher the NPV profit, the more attractive a compound will be; however, this needs to be balanced against the NPV cost and the probability of success (indicated by the diameter of the bubbles). The bubble charts can be used to assess the impact of variations in input variables visually, for example, lengthening timescales will change the profit to cost balance.

The analyses provided by the model allow managers to compare the probable outcomes for different compounds and help to inform decisions that attempt to maximize profits, while containing costs and reducing risks. The model allows variations of input that can be used to assess the robustness of decisions in terms of how variations in the process (for example, lengthening timescales or increases in the costs of particular stages) are likely to impact the expected profits.

Conclusions

This paper proposes a framework for the quantitative risk modelling of the main risk factors in the development of new compounds for medicinal use. Using Bayesian networks sampled by Monte Carlo methods, the framework underpins a scenario analysis tool to compare risk–benefit ratios for candidate compounds, at an early stage following drug discovery but before the substantial investment that is required to bring a new pharmaceutical drug to market. Simulations using artificial data were used to demonstrate the underlying principles and to illustrate the practice of this risk assessment method. The immediate future work is to evaluate the accuracy of the model estimates, using more realistic data.

Further developments of this framework will also focus on the interface with the decision maker to enable qualitative data entry, which is then automatically quantified in the form of parametric probability density functions, and to refine the visualization of the cost–benefit ratios and other outcomes of interest. In addition, Bayesian methods support a detailed sensitivity analysis to ascertain the effects of changes in parameter values as well as distributional assumptions and, finally, to identify the sensitivity to each stage of target thresholds (e.g. by pre-specifying a minimum probability of achieving a set benefit ratio). This will alert the user to the particular components of the model that are most important to estimate as accurately as possible, if necessary by reference to historical data.

Acknowledgements

We acknowledge helpful discussions with J. Akomode of Liverpool John Moores University during the early stages of the research and I. Smith formerly of Liverpool John Moores University, currently an independent consultant on preclinical safety.

References

1 Bognor, W. (1996) *Drugs to Market: Creating Value and Advantage in the Pharmaceutical Industry*, Elsevier Science
2 Cocchetto, D. and Nardi, R. (1992) *Managing the Clinical Drug Development Process*, Marcel Dekker
3 Bonate, P. (2000) Clinical trial simulation in drug development. *Pharm. Res.* 17, 252–256
4 Spiegelhalter, D. *et al.* (1993) Applying Bayesian ideas in drug development and clinical trials. *Stat. Med.* 12, 1501–1511
5 Goggin, T. *et al.* (2003) Modelling and simulation of clinical trials: an industry perspective. *In Simulation for Designing Clinical Trials: A Pharmocokinetic – Pharmacodynamic Modelling Perspective* (Kimko, H., and Duffull, S. eds), pp. 227–244, Marcel Dekker
6 Rooney, K. *et al.* (2001) Modelling and simulation in clinical drug development. *Drug Discov. Today* 6, 802–806
7 Spiegelhalter, D. *et al.* (2004) *Bayesian Approaches to Clinical Trials and Healthcare Evaluation*, John Wiley and Sons
8 Blau, G. *et al.* (2000) Risk management in the development of new products in highly regulated industries. *Comput. Chem. Eng.* 24, 659–664
9 DiMasi, J. *et al.* (2003) The price of innovation: new estimates of drug development costs. *J. Health Econ.* 22, 151–185
10 Brummel, R. (2002) Process chemistry: creativity in large scale: lessons from scale-up. In *From Bench to Pilot Plant: Process Research in the Pharmaceutical Industry* (Nafissi, M. *et al.* eds), American Chemical Society
11 Hutchinson, D. (1997) *How drugs are developed: a practical guide to chemical research*, Brookwood Medical Publications
12 Cole, G. (1993) *Pharmaceutical Production Facilities: Design and Applications*, Ellis Horwood
13 Sheiner, L. and Wakefield, J. (1999) Population modelling in drug development. *Stat. Methods Med. Res.* 8, 183–193
14 Greener, M. (2001) *A Healthy Business: A Guide to the Global Pharmaceutical Industry*, Informa Publishing Group
15 Walley, T. *et al.* (2004) *Pharmocoeconomics*, Churchill Livingstone
16 Bingham, M. *et al.* (2001) Modelling choice behaviour for new pharmaceutical products. *Value Health* 4, 32–44
17 Miller, P. (2005) *Role of Pharmacoeconomic Analysis in R&D Decision Making: When, Where, How?* (Vol 23, No 1), pp. 1–12
18 Congdon, P. (2001) *Bayesian Statistical Modelling*, John Wiley and Sons
19 Emblemsvag, J. (2003) *Life Cycle Costing: Using Activity-Based Costing and Monte Carlo Methods to Manage future Costs and Risks*, John Wiley and Sons
20 Jackel, P. (2002) *Monte Carlo Methods in Finance*, John Wiley and Sons