# Effective decision-making: progressing compounds through clinical development

Christine A. Shillingford and Colin W. Vose

**In today's drug-development environment, companies are under increasing pressure to improve efficiency and maintain returns on investment. Tomorrow's environment is likely to be harsher still, and companies that survive and prosper will be those that are already responding by re-inventing their structures and culture to meet the challenges that lie ahead. In this review, we explore some of the strategies and issues that are central to this process, with particular reference to decision-making in drug development.**

***Christine A. Shillingford**
Phase I Services
Quintiles
Station House
Market Street
Bracknell, UK RG12 1HX
*tel: +44 1344 708 659
fax: +44 1344 708 450
e-mail: christine.shillingford@quintiles.com
**Colin W. Vose**
Preclinical Sciences
Quintiles
Heriot Watt University
Research Park
Edinburgh, UK EH14 4AP

▼ Taking into account the cost of failures, the average drug currently costs US$500 million to develop, and takes 10–12 years from discovery to market[1]. The average dossier now contains ~35 clinical trials, involving >4000 subjects. Although the number of trials per dossier has actually fallen in recent years, there has been no equivalent fall in the total number of subjects as fewer, but larger and more complex, studies become the norm.

Meanwhile, 90% of marketed drugs earn less than US$180 million per annum (p.a.), 70% fail to recover their R&D costs and only ~12% provide a satisfactory return on investment (ROI). There is also decreasing growth within the pharmaceutical industry. Figures for 2000 are expected to show a fall to 8% growth, compared with 10% in 1999, with fewer potential blockbusters in the pipeline. To add to the industry's woes, drugs that are worth US$43 billion p. a. will lose patent protection or market exclusivity between 2000 and 2004 (Ref. 2). The industry has never been under greater pressure, forcing companies to seek more effective decision-making strategies with respect to candidate selection and clinical development.

Currently, this process is inefficient and compounds are subject to the following average attrition rates during development where, from each of the 30,000 compounds synthesized[3]:

- 2000 (6.7%) enter preclinical development;
- 200 (0.67%) enter Phase I trials;
- 40 (0.13%) enter Phase II trials;
- 12 (0.04%) enter Phase III trials;
- 8 (0.027%) are approved;
- 1 (0.003%) makes a satisfactory ROI.

These attrition rates are further complicated by the opportunities to generate more active new chemical entities (NCEs), which are offered by the combination of genomics, rational drug design, combinatorial chemistry and HTS. In many cases, this is already delivering more compounds to the preclinical discovery–development interface, increasing the choice of development candidates. However, the search for increased potency and receptor subtype specificity leads to more complex molecular structures, with the addition of large lipophilic substituents to the primary pharmacophore to achieve these objectives. Many of these compounds could have physicochemical properties (e.g. partition coefficients, molecular weight, molar refractivity and aqueous solubility) that reduce their development potential significantly and, therefore, increase the risk of failure during development. The early screening of these compounds for poor absorption, the potential for extensive first-pass metabolism, and other properties that would make them poor development candidates, will increase the attrition at the late discovery stage. However, attrition should then fall in preclinical development and Phase I trials as better molecules are selected for development.

1359-6446/01/$ – see front matter  PII: S1359-6446(01)01945-6 

Further improvements in modelling the relationships between physicochemical properties of drug molecules and their pharmacokinetics (PK) and receptor binding will guide the synthesis of NCEs with a greater potential for successful development. This extension of rational drug design application and *in silico* modelling of the optimal drug characteristics, should reduce the number of molecules synthesized with physicochemical properties that make them poor development candidates. Nevertheless, at present the cost of this attrition is staggeringly high. However, pharmaceutical companies are hit hardest when drugs fail in late clinical-development because, by the time drugs are completing Phase III trials, >90% of their development costs have been incurred[1]. This is driving a culture of 'zero tolerance' of attrition losses at this stage[4]. The risk of late failure varies across therapeutic areas, being highest for drugs for diseases of the CNS (50%) and lowest for antibiotics (20%; Ref. 3). These differences might reflect the relative reliability and predictive power of the early efficacy data on which go/no-go decisions are based in these areas. They could also be of strategic importance when companies are planning their development portfolios.

If a drug is going to fail, it should fail quickly, thus the industry is trying to shift attrition to earlier in development, aiming for zero failure rates in Phase III and post-registration. Companies are beginning to re-align their resources to facilitate this; for instance, industry surveys show that the proportion of time spent by preclinical scientists in discovery screening to select compounds with improved development potential, is increasing. This has created an increased demand for preclinical scientists. However, there is a general difficulty in recruiting well-trained and experienced staff to support the demands of both discovery screening and regulatory preclinical studies. This has led to an increase in the outsourcing of the regulatory toxicology, safety pharmacology, PK and absorption, distribution, metabolism and excretion (ADME) testing[4,5]. Also, companies are embracing new technologies, such as computerized trial simulation, and increasingly sophisticated *in vitro* and, in some cases, *in silico* screening filters to improve the reliability of their data and reduce their drug-failure rate. New technologies are expensive to implement, however, and costs are increasing before benefits, such as fewer late failures and increased ROI, are being enjoyed[3,5]. As a result, development decisions are becoming both more crucial and more complex.

## What is an effective decision-making process?

It can be argued that an effective decision-making process minimizes risk and maximizes benefit. In clinical development, decisions carry risks to both the corporation and the individual decision-maker, whose performance can be evaluated on the basis of the judgements he or she makes. Long clinical development cycles make it unlikely that one person will oversee a drug from Phase I trials to registration, so an individual's performance milestones can be based on successful transition to the next development phase, rather than on successful registration. Thus, Phase I decision-makers might be rewarded for the number of compounds they hand on to Phase II, rather than the number of those compounds that subsequently reach the market: this can create the risk of the 'never mind the quality, feel the width' phenomenon as a driver of decisions. Therefore, it is important to acknowledge that no-go decisions could carry adverse consequences for the individuals who make them because their hand-over targets might not be met. This could happen even when the no-go decision is correct because the accuracy of a no-go decision is impossible to prove and the benefit to the company is not easily measurable. It is only possible to prove that go decisions were correct or incorrect because revenues are tangible and measurable.

However, only one in 250 drugs entering preclinical development eventually reaches the market, so no-go decisions must, for necessity, be taken more frequently than go decisions. There are two types of potential error in the decision-making process: incorrect termination of a successful product and prolonged development of an unsuccessful product.

Currently, 24 out of 25 drugs entering clinical development will fail to reach the market, and if these drugs could be terminated before entering clinical development, significant savings could be made. If one of 25 candidates was chosen for development at random, the chance of picking the correct one would be 4%. If one of the candidates was chosen for termination, the chance of picking one of the right ones at random would be 96%. Because we have a greater chance of being correct when we terminate drugs, we should assume drugs to be failures unless they display clear characteristics for success; for instance, we should set termination criteria much more strictly in early development, to avoid costly prolongation of development programmes for flawed candidates. Pharmacoeconomic targets for ROI should be included in the termination criteria because they translate potential clinical benefits into commercial viability[6]. In an environment where increasing numbers of drugs fail at the registration stage because they cannot demonstrate cost-effectiveness, we have to select drugs that will not only meet treatment objectives but will also make a satisfactory ROI.

The political and psychological pressure against making no-go decisions and the personal risk associated with them

might cause delaying or avoidance behaviour, thereby increasing the time and money wasted on poor development candidates passed down the line. Therefore, to take effective decisions, personal risk or benefit must be uncoupled from the decision-making process to enhance objectivity.

It has been said that, 'Vision creates intent. Culture determines action[7].' When corporate vision and management culture are out of step, perverse incentives can be produced encouraging counter-productive behaviour. If functional groups in the development chain are given inappropriate targets and incentives, their ability to make objective assessments is reduced. Therefore, if the discovery or preclinical development group is rewarded for the quantity of drugs entering clinical development, there is a positive incentive to progress a drug for the sake of achieving the target, because if drugs fail later another functional group will bear the consequences. However, if they are rewarded for the quality of the drugs they select (e.g. for low attrition rates in Phase I), fewer drugs might be selected for development but the attrition rate in clinical development could be lower. As long as functional groups are measured on quantity and not quality, the number of drugs they are expected to propose for clinical development each year will be corrected for current attrition rates, and will inevitably be higher than the number of good-quality candidates they can actually identify. This means that, to meet the targets set, inappropriate or sub-optimal candidates will enter development, only to fail later on. In this way, the current attrition rates become self-perpetuating.

**Table 1. Criteria and data supporting go or no-go decision-making**

| Criterion | Go | No-go |
|---|---|---|
| PD activity at tolerable doses | Reproducible | Absent |
| | Relevant | Variable |
| | Dose-related | Poorly related to dose |
| PD duration | Allows dosing regime acceptable to patients | Requires inconvenient, complex dosing regime |
| PK characteristics | Linear with dose and time | Non-linear with dose or time |
| | Low inter/intra-subject variability | High inter/intra-subject variability |
| PK–PD relationship | Well-defined | Inadequate |
| | Predictable | Unpredictable |
| Safety profile | Predictable | Unpredictable |
| | Wide therapeutic ratio | Narrow therapeutic ratio |
| Bioavailability | Acceptable | Unacceptable |
| | Predictable | Unpredictable |
| | Low inter/intra-subject variability | High inter/intra-subject variability |
| Physicochemical properties | Allows adequate exposure in human studies in an appropriate formulation | Difficult to achieve adequate exposure in human studies in an appropriate formulation |
| Commercial viability | Adequate market size and/or price profile to achieve return on investment | Market too small and/or low-priced to ensure adequate return on investment |

Abbreviations: PD, pharmacodynamic; PK, pharmacokinetic.

In response to this, a development model comprising three crucial decision-points has been proposed[8]:

(1) Preclinical to exploratory clinical pharmacology in healthy volunteers or mildly ill patients (Phase I/IIa);
(2) Phase IIa to exploratory clinical pharmacology in ill patients (Phase IIb); and
(3) Phase IIb to confirmatory clinical development in the intended patient population (Phase III).

Table 1 shows examples of some criteria of value in supporting efficient decision-making at each stage and examples of data that could lead to go or no-go decisions. To decide the relative importance of the selected criteria at each stage, the desired optimal-product profile for therapeutic and commercial success needs to be clearly identified. This requires close collaboration between the commercial and R&D teams.

Other considerations when formulating an early development decision-making strategy include the target disease (i.e. chronic or acute, mild or life-threatening) and the nature of the product (e.g. first-in-class or fast-follower). These will also influence the relative importance of the criteria selected to support decision-making.

It has been estimated that, of all drugs failing in late-clinical development (excluding antibiotics), 46% fail for lack of efficacy and 27% for lack of safety[9]. In addition, increasing numbers of drugs fail at registration because they do not demonstrate cost-effectiveness. The predictive power of early phase studies for late-stage outcomes still remains

poor and the process by which drugs are evaluated in early development is inefficient.

The problems associated with decision-making at the crucial point where a drug proceeds to human dosing include the predictivity of animal results to man because:

- preclinical profile could be affected by lack of optimal formulation or low bioavailability;
- validated animal models might not exist;
- metabolism in man could differ from animal species used;
- choice of indication might be uncertain and this can affect risk or benefit assessment; and
- lag-time to market could mean that market conditions change, affecting projected ROI.

Similarly, problems associated with decisions to progress to full development include:

- lack of understanding of dose–PK–pharmacodynamics relationship in the disease state;
- relevance of any biological markers used in Phase I and/or Phase IIa to clinical outcomes;
- uncertainty about therapeutic dose range;
- uncertainty as to what constitutes a clinically relevant effect;
- relevance of unexpected findings in Phase I and/or Phase IIa; and
- relative ROI compared with other competing programmes.

## The application of Likelihood Function to decision-making

It has been further proposed that in making these decisions, Likelihood Function (LF) should be used to assess drug candidates, instead of significance levels in individual studies. The LF summarizes the evidence that supports competing hypotheses of treatment effect on the basis of observed data at any point in time (e.g. null hypothesis versus an alternative hypothesis). The alternative hypothesis could represent the benefit of the drug in relation to the clinically relevant treatment effect that was anticipated at the start of its development. The evidence supporting one hypothesis versus the other is the ratio of their likelihoods. This approach enables data to be amalgamated from a variety of investigations as exploratory development progresses, and could be useful in integrating efficacy data in a way that is of more value in Proof of Concept (POC) decision-making. This would support a two-stage drug development model (exploratory and confirmatory)[10] and provide a means of integrating exploratory development data to improve the reliability of decisions to move compounds into full (confirmatory) development.

In applying this method to the decision to enter full development, the following are prerequisites:

- there should be an understanding of the effective dose range;
- the treatment effect seen in Phase IIa should meet or exceed that which is considered clinically relevant; and
- the sponsor should have a draft-annotated label with all relevant issues addressed, or addressable by the confirmatory Phase III trials.

If one accepts this paradigm, then Phase IIb could, in certain cases, be considered part of exploratory development, if characterizing the dose response in the target population is one of the POC parameters. However, the overall objective of Phase IIb is to enable optimal dose selection for the confirmatory trials, the primary hypothesis of which should address the primary indication for the drug.

It is important to note that the LF function has only been used to address efficacy and not safety or risk/benefit assessment. Nonetheless, in its current form, LF does appear to offer a more meaningful assessment of the drug's therapeutic benefit and might contribute to a reduction in late-phase drug failures caused by lack of efficacy. However, sometimes safety might be the most important factor driving decisions and it would be interesting to know if the LF technique is applicable to these data. Such an application could be of value in reducing the 27% late-phase failure attributable to safety and/or toxicity problems.

The level of late-phase drug failures caused by safety or toxicity problems observed by the Centre for Medicines Research (CMR; Ref. 9) is in broad agreement with the International Life Sciences Institute's Health & Environmental Science Institute (ILSI-HESI) conclusions[11]. They conducted a survey of 131 compounds for which human toxicities (HTs) had been identified in Phase III or post-registration, and reviewed the preclinical toxicology data to determine if these HTs had been identified in toxicology testing. They concluded that, in 31% of cases, animal results had failed to predict the HTs, even though adequate challenge to the species (maximum tolerated dose) had been achieved in 88% of cases and testing had exceeded one month in 73% of cases. Perhaps more significantly, 69% of HTs were predictable from animal toxicology testing and almost all (95%) in tests of one month's duration or less. They also found that for the compounds studied, the predictivity of data from dogs and primates was similar and in both cases was significantly better than that from rats.

One interpretation of these findings is that greater weight should be given to toxicity findings than at present, in non-rodents at one month, and that failure to

detect toxicities in these tests should be viewed with suspicion and the compound subjected to more critical evaluation. As newer and more sophisticated methods of toxicity testing become available, including new toxicity biomarkers resulting from proteomics research, the criteria for preclinical safety screening can become more stringent, allowing problem compounds to fail earlier. However, until corporate cultures support it, truly comprehensive and critical appraisal of available drug candidates could be inhibited by fear of the consequences of terminating them, resulting in wasted time, money and opportunity.

### Who should decide what?

Given all of the above, objectivity and critical appraisal of each development candidate is vital to any company's long-term health. Who, though, can be objective? This requires both the ability to critically appraise the data against the criteria set for go or no-go decision-making, a lack of personal investment in the decision and a culture that values project assassins as much as project champions.

As long as decision-makers continue to reside within functional groups with the incentive to produce quantity over quality, they will guard this role closely because it enables them to control the achievement of their performance targets. However, there is a strong argument for independent assessment of development candidates if quality is to be achieved.

It has also been suggested that project teams should present not one, but several development options for each product at each resource allocation or go or no-go decision point. In one case (SmithKline Beecham)[12], the following options had to be considered for each of the 20 development candidates in the portfolio:

- current option;
- buy-up option (higher investment);
- buy-down option (lower investment); and
- minimal option (abandon project but preserve as much of the value earned, to date, as possible).

With the aid of a facilitator, the project team adopted the current plan (to develop the product for one or more indications agreed previously) and brainstormed the remaining three scenarios. Values were then calculated for each option according to an agreed paradigm, in a process called 'calibration'. This process was found to generate an improved understanding of those elements of development that added most value to each product and project teams were able to focus their development work more productively.

The four options from each project team were then subjected to peer review to test the assumptions and improve the quality of the project alternatives. The refined options contained the following information:

- the nature of the plan;
- the investment required;
- the value of the product; and
- the anticipated ROI.

These were then presented to the senior management group that would ultimately make the investment decisions. At this stage, this was for discussion only (not evaluation) to assist the project teams in further improving the quality of their plans.

A consistent methodology – decision analysis – was used for the assessment of each project alternative. Six requirements were set for the presentation of a credible plan:

- standardized information set for every plan;
- reliable sources for all information presented, including expert opinions in areas of uncertainty;
- clear documentation of all sources of information, including meetings with experts;
- peer review to ensure consistency of methodology across project teams;
- valuations were compared with those performed by external industry observers and market analysts to ensure realistic values; and
- the impact of each variable on the expected value of the project was to be identified to ensure clear understanding of key value-drivers.

To maintain objectivity, independent analysts, who had no vested interest in any of the projects under evaluation, processed the valuation information. This approach also ensured consistency in the valuation methodology applied to all projects. Further peer and senior management reviews enabled the implications of the valuations to be discussed and, where necessary, the underlying assumptions to be modified. In addition, review by senior management enabled cross-project considerations to be made; for example, would the new product cannibalize an existing one?

Once a consensus had been reached on the valuations and the basis on which they had been made, the decision-making process could begin. At this stage, however, if a project was dropped because a product's chances of approval were low, nobody could challenge the assumptions on which that low probability had been based. Importantly, this prevented 'intuitive' challenges by senior managers from carrying more weight than carefully researched plans.

Decisions were then made with the single, clear objective of arriving at the highest overall portfolio value, with the final ROI evaluations made by a neutral team of

analysts, who again had no vested interest in any of the projects under evaluation. The various possible configurations of the project mix were tested for stability under a variety of different conditions, for balance across therapeutic areas, and so on. This process reduced controversy in the resource-allocation process and removed the bias towards go decisions caused by project champions. In addition, it led to a reversal of the company's investment strategy, which had been that of development budget reduction. They found that by retaining the current level of investment, a portfolio of 30% more value than the old one could be generated. Furthermore, the marginal return on additional investment increased from 5:1 to 15:1 and the company increased its development spending by 50%.

## Measuring effectiveness

Clinical development decisions centre on one pivotal question: What shall we invest and where shall we invest it? The accuracy of these decisions bears directly on the success of new treatments for human disease and the ROI made by the company, which is in turn reflected in stockholder value.

To maximize ROI, it is necessary to channel resources into the rapid development and registration of candidates with the best potential for success. This means that there has to be an incentive to terminate poor drugs early. If we turn that around, we have to provide a disincentive to make incorrect go decisions, which can be identified by failure of the compound later in development and which might be a more accurate performance measure at each stage. Hence, the performance of the preclinical decision-makers could be measured not by the number of compounds they pass to clinical pharmacology but by the percentage of those compounds that fail in Phase I and/or Phase IIa trials. Similarly, the performance of decision-makers at the crucial point where drugs are proposed for full development could be measured by the failure rate in Phase III trials and/or Registration. The performance of the whole development group is then reflected in portfolio value and ROI.

The overall objective of this type of process is, as described earlier, to increase the quality and value of the portfolio and thereby increase the ROI. However, a valuable by-product of this approach can be to focus attention at all stages of discovery and development on having well-defined criteria for decision-making, especially for making go decisions. In practice it is likely that not all of the criteria established for go or no-go decisions will be met at each of the key milestones. The type of process described above encourages clearer definition of the relative importance of the criteria selected, and the proportion that must be fulfilled to enable a go decision to be made.

## Conclusions

The problems described in this review are being experienced by all sections of the pharmaceutical industry. The extent of these problems and the detailed issues involved will be different between companies, across therapeutic areas and between classical medicinal chemistry, new chemical entities, biotechnology products and gene therapy. Inevitably, the approaches described will not be universally applicable, but they could provide opportunities to refine the response to the intense challenges the industry is facing and will continue to face in the 21st century. Companies should continue to create, apply, refine and evaluate new ways to improve decision-making in development-candidate selection. The reporting of such investigations would be a valued contribution to the continuous improvement of the development process.

## References


1 PriceWaterhouseCoopers Pharmaceutical Consulting Team (1998) *Pharma 2005: An Industrial Revolution in R&D.* (http://www.pwcglobal/pharma/)
2 Outlook 2000 Report, Tufts Center for the Study of Drug Development (2000) (http://www.tufts.edu/med/research/csdd/recntpub.html#drug)
3 Kuhlman, J. (1997) Drug research: from the idea to the product. *Int. J. Clin. Pharm. Ther.* 35, 541–552
4 Centre for Medicines Research (2000) *Strategic Statistics (3): Pharmaceutical Investment and Output in 1998.* CMR International R&D Briefing, 26 (January) (http://www.cmr.org)
5 Centre for Medicines Research (2000) *The Changing Roles of Non-clinical Safety Evaluation at the Discovery/Development Interface: Report of an Industry Discussion Briefing.* CMR International R&D Briefing, 29 (May) (http://www.cmr.org)
6 Clemens, K. *et al.* (1993) Strategic use of pharmacoeconomic research in early drug development and global pricing. *PharmacoEconomics* 4, 315–322
7 Olsen, E. (1999) When culture undermines vision. Boston Consulting Group (http://www.bcg.com/publications/search_view_perspectives.asp?pubID=301)
8 Peck, C.C. and Woosley, R.L. (1997) Clinical development may soon be accomplished in less than 3 years: Will FDA and the pharmaceutical industry be ready? *Testimony for the Hearing Record: Reauthorization of the Prescription Drug User Fee Act and FDA Reform 4/23/97* (http://www.georgetown.edu/research/cdds/hearing.html)
9 Peck, R. (1997) Strategies in selecting an agent for clinical trials. *Ann. New York Acad. Sci.* 823, 319–327
10 Huster, W.J. and Enas, G.G. (1998) A framework for establishing clear decision criteria for the assessment of drug efficacy. *Stat. Med.* 17, 1829–1838
11 Olson, H. *et al.* (1998) The predictivity of the toxicity of pharmaceuticals in humans from animal data – an interim report. *Toxicol. Lett.* 102–103, 535–538
12 Sharpe, P. and Keelin, T. (1998) How SmithKline Beecham makes better resource allocation decisions. *Harvard Business Review* 76, 45–50