The Dynamics of Research and Development in the Pharmaceutical Industry

Productivity of Traditional versus New Research Technologies

Martin U. Großmann
Industrieseminar der Universität Mannheim
University of Mannheim
Schloss
D-68131 Mannheim
Federal Republic of Germany

Tel.: (+49 621) 292 - 5527
Fax: (+49 621) 292 - 5259
E-Mail: grossmann@mailrum.uni-mannheim.de

ABSTRACT

The process of research and development (R&D) in the pharmaceutical industry has become increasingly unproductive during the last few decades. One reason, among others, for this development is the diminishing level of performance reached by research technologies. In the following study the term 'performance' is limited to an output measurement which is reflected by the number of new drugs launched into the market by which therapeutic improvements can be realized.

The purpose of this study is to analyze the decreasing performance of traditional technologies in order to partly explain the reduction in R&D productivity. Subsequently, the potential impact of new technologies upon research performance will be simulated by using System Dynamics.

Broad-scale random screening is the main technological process traditionally used to discover chemical substances for new drugs. This study reveals that random screening can be adequately modelled by the statistical formula Poisson function. The function is used to calculate the probability of discovering new drugs. Empirical data from the German pharmaceutical industry from the 1950s onwards were put into the formula. The results show that the probability of discovering new drugs has decreased strongly by using random screening. Furthermore, the risk involved in research with random screening can be measured by Poisson distribution functions. It can be seen that risk has risen significantly since the 1950s.

The Poisson formula also provides a formal framework for forecasting the impact of new technologies on the rate of drug discovery. The high potential performance of new biotechnologies, especially genetic engineering, could increase research success rates significantly. A System Dynamics model has been constructed in a prototype version to generate scenarios for future output rates. The high uncertainty in predicting research successes can be estimated by a best, a worst and an intermediate forecast based upon varying assumptions. The software application Vensim has been used for modelling and simulating. The model is partly based on hypothetical data and is, therefore, a first step towards forecasting the impact of genetic engineering on research performance in the pharmaceutical industry.

I. Traditional versus New Research Technologies for Drug Discovery

The paper focusses on a specific category of pharmaceutical R&D called 'research route' (Cox, Millane, Styles 1975). The drug search process in research routes is often based on inadequate knowledge of the disease for which a chemical substance has to be discovered and subsequently developed to a medicine. The causes of diseases and their mechanisms are fairly unknown and thus, limited knowledge about their treatment is present. The objective pursued with research routes is to discover new ethical drugs by which therapeutic advances can be realized.

The action of drugs is often described by a lock-and-key analogy. The drug is seen as the key and the lock, as a drug receptor. The receptor is a molecule, for example on the surface of a cell, through which a drug causes a physiological response (OTA 1993, Weber 1992). This physiological response can either alleviate the symptoms or cure the causes of diseases. A chemical substance, which is efficacious in these respects, is a potential drug.

Drugs can basically be identified in two ways: randomly or rationally. The traditional technology used to discover drugs in research routes is called 'broad-scale random screening' or sometimes just 'random screening' (Walker and Parrish 1988). This means that several thousand chemical substances have to be tested, for example in animals, and screened to explore their therapeutic potential. The receptors, which control the stages of a disease, are unknown. Consequently, the search process for drugs does not rest on knowledge of how the drugs work. In this context, the search process is therefore of low rationality.

As a consequence, substances which bind to a receptor and therefore produce desired therapeutic results are found by accident. In other words, the discovery of drugs by the random screening technology occures at low probability. Thus, in Figure 1 the traditional technology, random screening, is classified as a low rational search process in which successes occur at low probability.

This traditional process can be reversed by rational drug design, searching backwards from a known receptor target (OTA 1993). The underlying theme of such a search technology is a deliberate design of substances to affect target molecules. This is only possible, of course, if any target molecules are known.

By using new biotechnologies, especially genetic engineering, such target receptors can be identified (Weber 1992). Moreover, scientific knowledge for understanding causes and mechanisms of diseases has increased significantly through these new research technologies. Genetic engineering allows, therefore, a shifting of scientific foundations of drug discovery from a low to a higher rational search process. Consequently, the successes will occur with higher probability compared to searching for drugs by random screening.

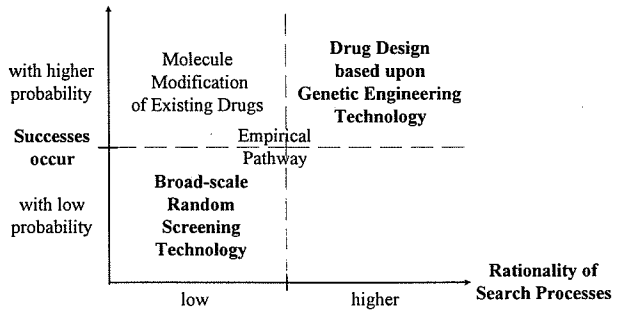

Figure 1: Classification of traditional and new research technologies

In Figure 1, the term 'molecule modification' means that chemical structures of existing drugs will be varied. The result of this trial-and-error approach is often a generic product or a slightly improved medicine. Since the search process is based upon existing drugs, the successes occur with higher probability compared to random screening. Its rationality is similarly limited.

Discovering drugs on an empirical pathway is based upon random screening. During screening, knowledge of the disease can be accumulated by trial-and error-learning and this helps to establish an understanding of drug actions. Empiricism increases the rationality of the search process and the probability of success slightly in comparison to random screening (Figure 1). Neither the molecule modification nor the empirical pathway are subjects of this study. The paper is focussed on research technologies, in Figure 1 bold, which are allocated to opposite classifications.

II. Modelling Drug Search Processes by the Use of Poisson Functions

A traditional technology used in discovering new drugs is that of broad-scale screening. As described above, a high number of chemical substances have to be tested to find active molecules which are useful in treating diseases. This technology for discovering drugs is also called 'random screening'.

As the term 'random' indicates, the efficacy of molecules is discovered by accident. (Webster and Swain 1991; Spilker 1989; Walker and Parrish 1988; Cox, Millane, Styles 1975). The results of testing chemical substances cannot be determined in advance. Thus, and based upon the opinion of the above cited experts, screening tests can be interpreted as random experiments. The following thoughts are based upon this view.

After testing a chemical substance, two results are possible: the compound is efficacious and therefore has a therapeutic potential or it is not of value in treating a disease. The testing of chemical substances via broad-scale screening can therefore be interpreted as a number of Bernoulli trials. To analyze this search process statistically, it is necessary to specify some statistical terms. Figure 2 shows the results.

Statistical term:	**Statistical unit**	→ described by an	**Characteristic**	→ with the following	**Attributes**
Equivalent in the drug search process:	**Chemical substance**	→	**Efficacious in therapy**	→	**Yes (X)** No

Figure 2: Specifying statistical terms

A random variable X is for the following study defined as the number of successful trials that occur during the screening process. This is equivalent to the number of new drugs, which are efficacious in the therapy (see Figure 2). The random variable only takes on non-negative integer values which dictates a discrete probability distribution. Thus, a binomial distribution can be used to model the probability distribution of X.

The Poisson distribution is the limit of the binomial distribution. To be able to use a Poisson function, four conditions have to be met (Cook and Russell 1985):

1. The number of Bernoulli trials (n) is large.
2. The probability of an event (p) within Δt is small.
3. The events occur independently.

4. The probability of an event is constant during Δt.

The conditions for using the binomial distribution are described in the numbers 3 and 4. The additional assumptions for using the Poisson function can be seen in numbers 1 and 2.

In Table 1, the conditions for using Poisson functions are compared with the situation in the research section by using the broad-scale random screening technology.

	Assumptions for using the Poisson Model	Broad-Scale Random Screening
1.	$n \geq 20$	$n \geq 1{,}000$ (in average n = 12,000)
2.	$p \leq 0.25$ The mean λ is the expected number of successes in n Bernoulli trials. The λ should be ≤ 5. $\lambda = np$	$p = 0.00008$ (for n = 12,000)
3.	The events occur independently.	The discovery of any successful substance does not make it more, or less, likely that another successful substance will be discovered. Thus, the successful trials occur independently of one another.
4.	p is constant.	The conditions during testing are held constant. It can therefore be assumed, that p remains constant.

Table 1: Comparing Poisson assumptions with the random drug search process

The comparisons in Table 1 show that the assumptions 1 to 4 hold true in reality. Thus, the Poisson function can be used to model a traditional drug search process in the form of random screening. The four conditions can be accepted as adequately reflecting some aspects of the search process.

The random variable X is therefore Poisson distributed. The probability that X will take on any value of k is given by the Poisson function:

$$P(X = k) = \frac{\lambda^{k} e^{-\lambda}}{k!}$$

for k = 0, 1, 2, 3, ...
0 elsewhere

where e = the base of natural logarithmus, with a value approximately equal to 2.71828.
λ = mean number of new drugs.
k = number of new drugs.

If λ is known, the values for the probability of k can be calculated. Table 2 gives λ for the German pharmaceutical industry from 1958 to 1993 (Riebel 1972; BPI 1985; Bäumler 1993). The empirical values for λ are the arithmetic means of successful trials calculated on all diseases, on which research has been undertaken during the respective years. It is an industry-wide figure.

Year	λ out of n = 10,000
1958	5
1967	2
1978	1.4
1985	1
1993	0.8
1993	(0.1)

Table 2: Average technological successes by using random screening

By using the random screening technology, five chemical substances out of 10,000 were identified, on average, as having therapeutic improvements in 1958. This technological success rate fell to 0.8 in 1993. Furthermore, the success rate 0.1 in the year 1993 is related to a specific disease. It is the rate for discovering a cytostatica used to treat cancer. The technological success rate is, on average, 0.1 cytostatica out of 10.000 chemical substances.

The probability of success for k and the mean λ is shown in Table 3. The Poisson function is used to calculate the probabilities for k equals 0 to k equals 10. This has been done for the mean of the respective year. The values have been rounded off, so they will not add up exactly to equal one in the last column. The Poisson frequency distribution of each year can be seen in Diagram 1 which shows the usual way of expressing Poisson functions.

New Drugs	1958	1967	1978	1985	1993	1993
k	$\lambda = 5$	$\lambda = 2$	$\lambda = 1{,}4$	$\lambda = 1$	$\lambda = 0{,}8$	$\lambda = 0{,}1$
0	0,01	0,14	0,25	0,37	0,45	0,90
1	0,03	0,27	0,35	0,37	0,36	0,09
2	0,08	0,27	0,24	0,18	0,14	0,00
3	0,14	0,18	0,11	0,06	0,04	0,00
4	0,18	0,09	0,04	0,02	0,01	0,00
5	0,18	0,04	0,01	0,00	0,00	0,00
6	0,15	0,01	0,00	0,00	0,00	0,00
7	0,10	0,00	0,00	0,00	0,00	0,00
8	0,07	0,00	0,00	0,00	0,00	0,00
9	0,04	0,00	0,00	0,00	0,00	0,00
10	0,02	0,00	0,00	0,00	0,00	0,00

Table 3: Probabilities for values of k at different means

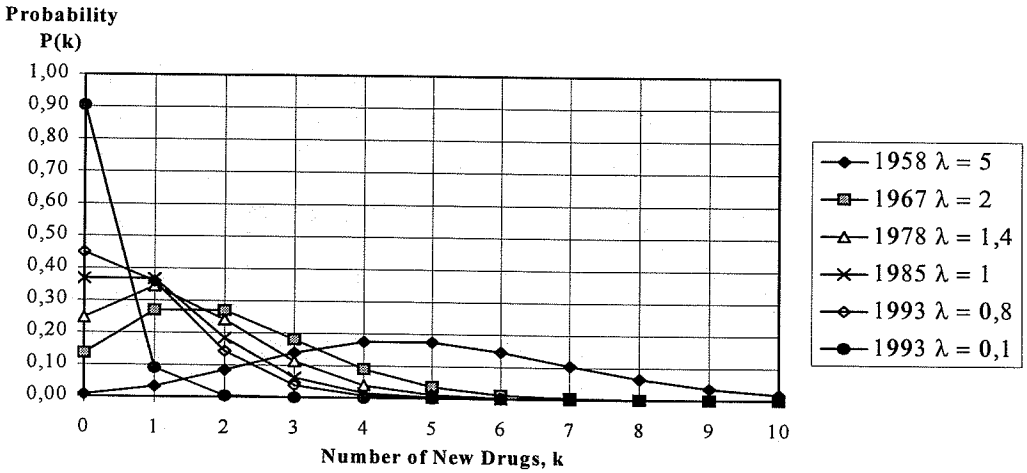

Diagram 1: Poisson Distribution at different means

As Diagram 1 indicates, the Poisson distributions are asymmetrical with respect to their means. The curves have a long tail to the right and therefore they are skewed to the right. Two trends can be seen. Firstly, the current situation of pharmaceutical research is shown by the

curves with small means, e.g. λ is 0.1 or 0.8. The probabilities of the curves associated with small values of k, e.g. k = 0, are the largest. The probabilities decrease rapidly as k increases.

The second trend was between the 1950s until the 1970s, with a mean over 1. These curves show, that the largest probabilities of successful trials become associated with values of k equal to one and higher. As k increases, the probabilities decrease slowly in contrast to the first trend described. The largest probabilities of values for k are shown in Table 4, taken from Table 3.

Year	λ	Largest probability of values for k
1978	1.4	0,35 for k = 1
1967	2	0,27 for k = 1,2
1958	5	0,18 for k = 4,5

Table 4: The largest probabilities of technological successes from the 1950s until the 1970s

The λ was 1 in the mid 1980s. This curve divides the two trends described above. Its probability for k equals 0 and 1 was constant at 0.37. The line for λ = 1 proceeds horizontally between these values of k in Diagramm 1. This can be seen in Table 2 as well, where the probability for k equals 0 and 1 remains at 0.37.

III. Measuring Technological Risk by Poisson Distribution Functions

The term 'risk' is related to the research technology 'random screening'. To analyze risk, two components have to be considered (Eckert 1985):

1. The target of research k*, which should be reached by random screening. This is a specific value for the random variable X which is defined as the number of new drugs.
2. The Poisson distribution function: $F(k^*) = P(k < k^*)$.

The Poisson distribution function is defined here as the cummulative probability calculated with probability values for k less than k*. The risk involved in random screening can be expressed by means of these distributon functions. The risk is the sum of all probabilities for values of k less than a specific target k*. Consequently, the technological risk is measured by the probability of **not** reaching the target, i.e. the probability of failure. Diagram 2 shows the distribution functions calculated with values from Table 3.

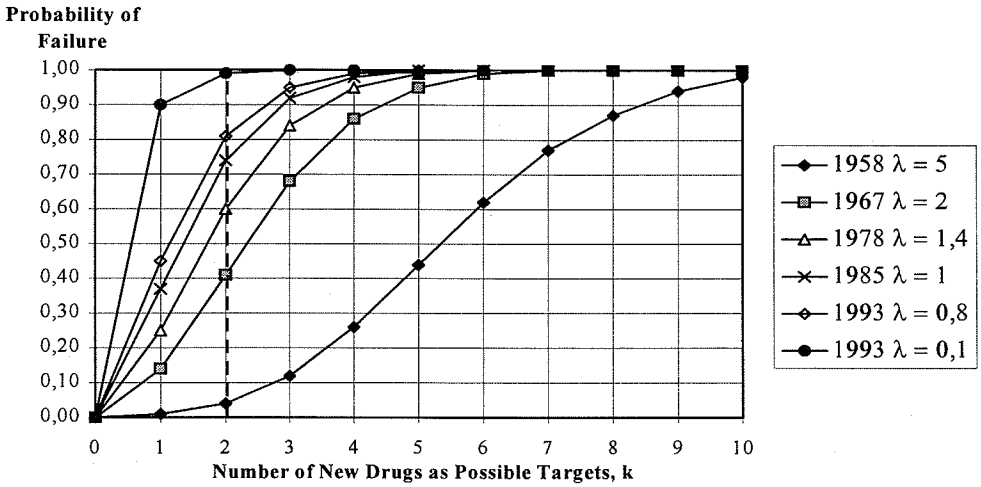

Diagram 2: Technological risk of random screening

The distribution functions in Diagram 2 give the risk profile related to k at the mean λ of a specific year. The research target per year can be set at k* equals 2, which could be the target for an pharmaceutical company, under the first 20 world-wide in terms of annual sales. For the lowest curve in the year 1958, the probability of not reaching the target of 2 was 0.04 or 4 percent in 10,000 trials. By 1993, this probability had increased to 0.81 or 81 percent.

The development of technological risk by using random screening at k* equals 2 can be seen as well in Table 5. The numbers show how enormously technological risk has increased industry-wide by using random screening.

Year	Technological Risk Probability of Failure
1958	4 %
1967	41 %
1978	60 %
1985	74 %
1993	81 %
1993	(99 %)

Table 5: Increasing technological risk

The discussion of Diagrams 1 and 2 reveals that the probability of research success by using random screening has fallen enormously since the 1950s. The performance of this traditional technology is widely "exhausted" in searching for efficacious drugs to treat complex diseases like cancer and AIDS. Drugs for less complex diseases have largely been discovered. The scientific knowledge and medical experience gained in past research activities is inadequate and seems to be resulting in a bottleneck of future drug discovery.

IV. A Prototype System Dynamics Model for Forecasting the Impact of New Technologies on Research Productivity

The scientific knowledge gained with new research technologies, especially through genetic engineering, could help to overcome the bottleneck described above. The impact of genetic engineering related to the process of screening is at least twofold:

1. Genetic engineering helps to design and synthesize new chemical substances rationally. Thus, it may be more likely, as it is the case by random selection of substances, that designed molecules have main biological properties necessary to treat diseases efficaciously (OTA 1993, 120). However, it is assumed that screening tests are still random experiments.
2. The efficacy and safety of substances is traditionally discovered in the laboratory by using disease models, for example animals. These have often had a low predictability for man. Consequently, it is likely that chemical substances "fell through the screen", which could have had a therapeutic improvement in humans. Genetic engineering may provide disease models with a higher predictability for man (Weber 1992, 45-48).

As a consequence of points 1 and 2, the probability p of discovering a chemical substance having a therapeutic improvement increases through genetic engineering. For the use of Poisson functions it is assumed that $\lambda = np$. Thus, by an increasing p, λ will increase whilst the number of trials n is kept constant.

Forecasting the impact of genetic engineering means, that λ has to be predicted. An exact measurement of future λs is not possible, since the point in time of major scientific breakthroughs and their effect on drug discovery is unknown. The high uncertainty in predicting λ can be reduced by a best, a worst and an intermediate scenario. Table 6 shows how the probability of research successes could increase. The numbers are the assumed values at which λ increases annually starting from 1994 onwards. The data are hypothetical and symbolize the growth of scientific knowledge beneficial to discovering drugs.

Best	Intermediate	Worst
0.2	0.1	0.05

Table 6: Forecasting the increasing probability of research successes

The feedback loop of the prototype System Dynamics model is displayed in Figure 3. The model's structure is identified in the real world and this could be the situation for a pharmaceutical company under the first 20s in terms of annual sales.

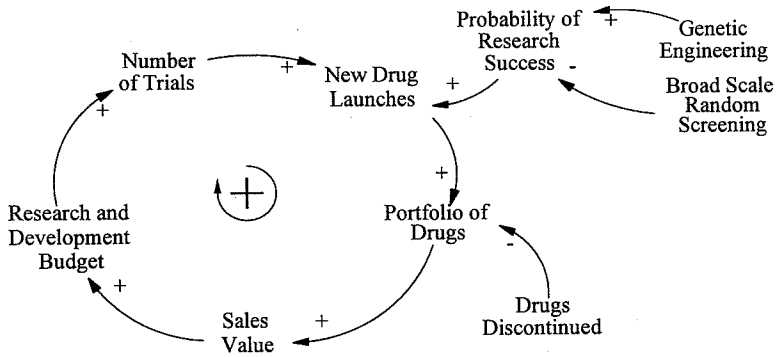

Figure 3: Core R&D loop

By the variable 'Research Intensity' in Figure 4 it is assumed, that the company allocates a percentage of its sales income to R&D. The R&D budget is divided by the average expenditure for 10,000 trials. The result is the number of 10,000 trials which can be carried out in the laboratories. This number multiplied by the probability of success λ creates an annual flow of new drugs which adds to the portfolio of medicines. In reality, the R&D expenditure associated with a new drug and the number of trials in the laboratories is incurred over the 10 - 15 years prior to launch. The length of this process is not captured in the model described here. The same simplifying aggregation was used by Hobbs and Deane (1994).

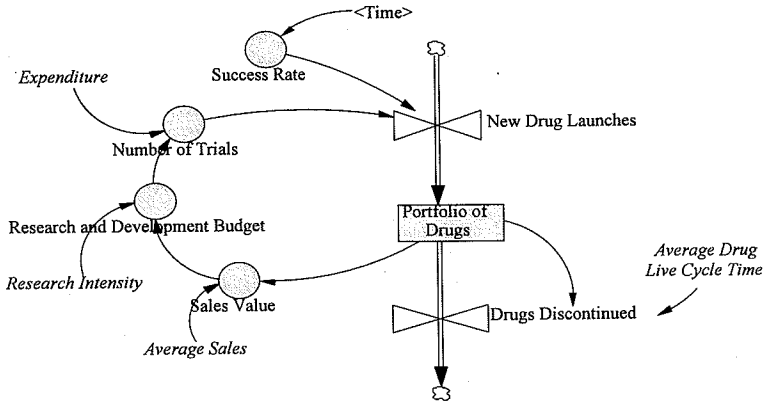

Figure 4: Core stock and flow of R&D

There are different tests in order to proof if a model can be accepted as a valid representation of the real world (Saeed 1994). In the model, empirical and partly statistical values are used for the parameters. Thus, it can be stated that the model is valid in this respect (Milling 1974). Table 6 shows the values of parameters used.

Parameters	Values
Average Drug Live Cycle Time	25 years (Grabowski and Vernon 1990)
Average Sales	$ 50m p.a. (Hobbs and Deane 1994)
Research Intensity	15,52 percent (BPI 1994)
Expenditure	$ 194m p.a. (estimated out of OTA 1993)
Initial number of drugs in the portfolio	50 drugs

Table 6: Parameters used in the model

Diagram 3 presents the scenarios for new drug launches after a simulation run until the year 2005. The scenarios start at the empirical value 0.8 in the year 1993.

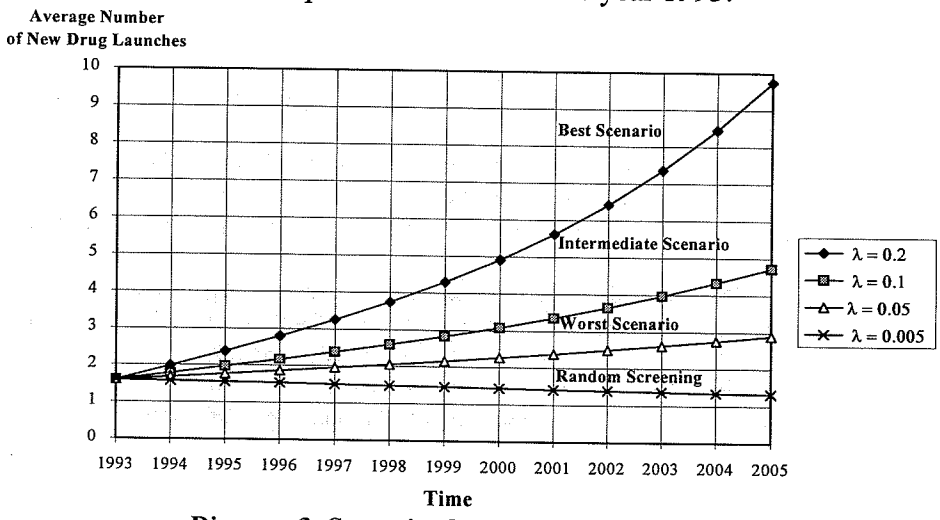

Diagram 3: Scenarios for new drug launches

The random screening scenario shows the output of R&D by using this traditional technology. It is assumed that the success rate of random screening falls 0.005 p.a. starting from the empirical value 0.8 in the year 1993. Diagram 4 shows the sales prospects under the conditions of Table 6.

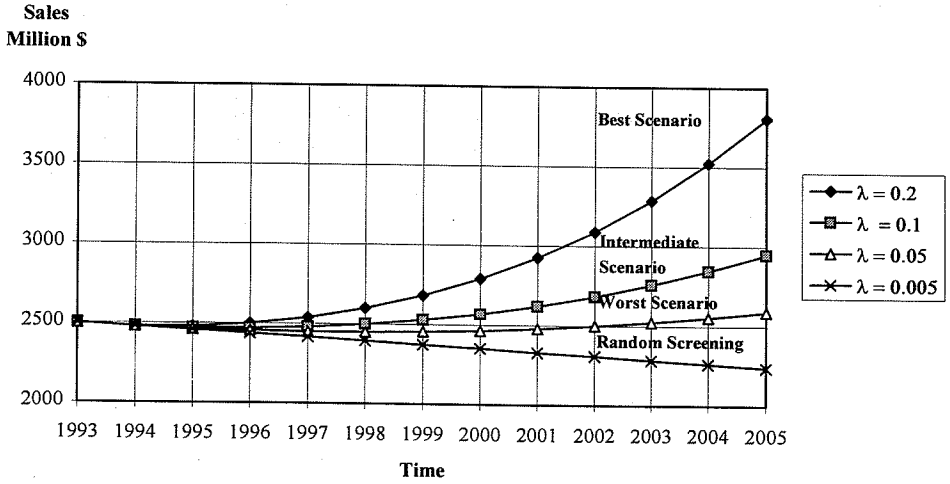

Diagram 4: Scenarios for annual sales

The impact of genetic engineering on research output can basically revealed by this prototype model for a pharmaceutical company of the top 20. The model can be anjusted to specific diseases. In addition, the input side of R&D, the expenditure for 10,000 trials, is kept constant during simulation. In further studies, causes of expenditure have to be identified and predicted to be able to gain an overall picture of future R&D performance by using genetic engineering.

References

Bäumler, E. 1993. Im Fadenkreuz von Forschung und Entwicklung. *Blick durch die Wirtschaft.* Jg. 36. Nr. 207. 26. Oktober 1993: 7.

BPI (Hrsg.). 1994. Bundesverband der Pharmazeutischen Industrie. Pharma Daten '94. 24. überarbeitete Auflage. Frankfurt am Main: 36.

BPI (Hrsg.). 1985. Bundesverband der Pharmazeutischen Industrie. Arzneimittelforschung in Deutschland. Erfolge der Vergangenheit - Stand - künftige Entwicklung. 4. Auflage. Frankfurt am Main: 6-7.

Cook, T.M., and R.A. **Russell**. 1985. Introduction to Management Science. third edition. New Jersey: Prentice-Hall, Inc.: 377-378.

Cox, J.S.G., B.V. **Millane**, and A.E.J. **Styles**. 1975. A planning model of pharmaceutical research and development. *R&D Management.* 3(5): 219-227.

Eckert, D. 1985. Risikostrukturen industrieller Forschung und Entwicklung. Theoretische und empirische Ansatzpunkte einer Risikoanalyse technologischer Innovationen. Berlin: Erich Schmidt Verlag: 37-38.

Grabowski, H., and J. **Vernon**. 1990. A new look at the returns and risks to pharmaceutical R&D. *Management Science.* 7(36): 804-821. esp. 809.

Hobbs, G.D., and B.R. **Deane**. 1994. Modelling the constraints on the global pharmaceutical industry. *Conference Proceedings (Industry)* of the 1994 International System Dynamics Conference in Stirling. Stirling: 45-56.

Milling, P. 1974. Der technische Fortschritt beim Produktionsprozeß. Ein dynamisches Modell für innovative Industrieunternehmen. Wiesbaden: Betriebswirtschaftlicher Verlag Dr. Th. Gabler: 212-214.

OTA. 1993. U.S. Congress. Office of Technology Assessment. Pharmaceutical R&D: Costs, Risks and Rewards. Washington, D.C.: U.S. Government Printing Office: 105-134.

Riebel, P. 1972. Kosten und Preise bei verbundener Produktion, Substitutionskonkurrenz und verbundener Nachfrage. 2. Auflage. Opladen: Westdeutscher Verlag: 79

Saeed, K. 1994. Development Planning and Policy Design. A system dynamics approach. Aldershot et al.: Avebury: 22-23.

Spilker, B. 1989. Multinational Drug Companies: Issues in Drug Discovery and Development. New York: Raven Press: 29.

Walker, S.R., J.A. **Parrish**. 1988. Innovation and New Drug Development. In Trends and Changes in Drug Research and Development. Walker, B.C., and S.R. Walker. eds.: 1-28.

Walker, B.C., and S.R. **Walker**. eds. 1988. Trends and Changes in Drug Research and Development. Proceedings of the Society for Drug Research. 20th Anniversary Meeting. Dordrecht, Boston, London: Kluwer Academic Publishers.

Weber, M. 1992. Gentechnik in der Pharmaindustrie. 2. überarbeitete Auflage. Basel: 43.

Webster, A., V. **Swain**. 1991. The Pharmaceutical Industry: Towards a New Innovation Environment. *Technology Analysis & Strategic Management* 3(2): 127-142.