

PHARMACEUTICAL SUPPLY CHAINS: KEY ISSUES AND STRATEGIES FOR OPTIMISATION

Nilay Shah

Centre for Process Systems Engineering, Department of Chemical Engineering,
Imperial College of Science, Technology and Medicine
London SW7 2BY, UK

Abstract

Supply chain optimisation is now a major research theme in process operations and management. A great deal of research has been undertaken on facility location and design, inventory and distribution planning, capacity and production planning and detailed scheduling. Only a small proportion of this work directly addresses the issues faced in the pharmaceutical sector. On the other hand, this sector is very much ready for and in need of sophisticated supply chain optimisation techniques.

At the supply chain design stage, a particular problem faced by this industry is the need to balance future capacity with anticipated demands in the face of the very significant uncertainty that arises out of clinical trials and competitor activity. Efficient capacity utilisation plans and robust infrastructure investment decisions will be important as regulatory pressures increase and margins are eroded. The ability to locate nodes of the supply chain in tax havens and optimise trading and transfer price structures results in interesting degrees of freedom in the supply chain design problem. Prior even to capacity planning comes the problem of pipeline and testing planning, where the selection of products for development and the scheduling of the development tasks requires a careful management of risk and potential rewards.

At the operation stage, it is often difficult to ensure responsiveness. Most pharmaceutical products involve primary active ingredient production (often multi-stage chemical synthesis or bioprocess) and secondary (formulation) production. Both of the stages are characterised by low manufacturing velocities and are hampered by the need for quality assurance activities at several points. It is not unusual for the overall supply chain cycle time to be 300 days. In this environment, supply chain debottlenecking and decoupling strategies together with co-ordinated inventory management are crucial for quick responses to changing market trends. A good understanding of what actually drives the supply chain dynamics is also required. As often as not, erratic dynamics are introduced by business processes rather than by external demand, and may be eliminated by the re-design of internal business processes or supplier/customer relationships.

This paper will consider important issues in supply chain design and operation drawn from the literature and from our collaborative research projects in this area. The main features of the problems will be reviewed as will the literature to date. Some strategies for solution will be identified, as will some future research needs.

Keywords

Pharmaceutical industries, supply chain optimisation, pipeline management.

Introduction

The pharmaceutical industry can be defined as a complex of processes, operations and organisations involved in the discovery, development and manufacture of drugs and medications.

The World Health Organisation (WHO) defines a drug or pharmaceutical preparation as:

any substance or mixture of substances manufactured, sold, offered for sale or represented for use in the diagnosis, treatment, mitigation or prevention of disease, abnormal physical state or the symptoms thereof in man or animal; [and for use in] restoring, correcting or modifying organic functions in man or animal.

This is a very wide definition, and correspondingly, there are number of key players in the pharmaceutical industry, including:

- i. The large, research and development based multinationals with a global presence in branded products, both ethical/prescription and over-the-counter. They tend to have manufacturing sites in many locations.
- ii. The large generic manufacturers, who produce out-of-patent ethical products and over-the-counter products.
- iii. Local manufacturing companies that operate in their home country, producing both generic products and branded products under licence or contract.
- iv. Contract manufacturers, who do not have their own product portfolio, but produce either key intermediates, active ingredients or even final products by providing outsourcing services to other companies.
- v. Drug discovery and biotechnology companies, often relatively new start-ups with no significant manufacturing capacity.

Most of the material in this paper is particularly relevant to the first group. This group dominates the marketplace and, due to the global nature of the enterprises involved, tends to have the most challenging supply chain problems.

The changing circumstances of the industry

In the recent past, the high returns on investment and high turnovers from “blockbuster” products resulted in the following regime (Booth, 1999):

- good R&D productivity, often creating compounds to treat previously untreatable diseases;
- long effective patent lives of these compounds;
- ability of these patents to provide technological barriers to entry;

- a limited number of product substitutes in a given therapeutic area;
- a low price sensitivity; supported by the separation between prescribing and paying responsibilities.

The resulting corporate strategy was to ensure high margins by exploiting the price inelasticity and invest a large proportion of the resultant profits in R&D (approximately 25% of sales), in order to ensure a healthy product pipeline.

The more recent circumstances are much more challenging:

- R&D productivity (in terms of numbers of new chemical entities registered per unit amount of investment) is declining;
- effective patent lives are shortening;
- even while active, patents provide lower barriers to entry;
- there are many product substitutes in many therapeutic areas; either alternative compounds (“me-too drugs”) or off-patent generics;
- the payers of healthcare are exerting strong price pressure and influencing prescribing practices; this means that in order to be approved, new drugs must address new therapeutic areas or have very significant cost or health benefits over existing treatments.

On the one hand, the global marketplace has become more liberalised, exposing products to competition. On the other, governments and other agencies have tended to intervene more as they become concerned at every increasing healthcare costs associated with aging populations. Measures taken include strict controls on the prices of new drugs, more cost-benefit analysis, and encouragement of the use of generic substitutes or alternatives where possible.

A further weakness that will hamper the large players in the area is the historical dependence on “blockbuster” drugs. A recent report by Datamonitor predicts that only 4 of the 19 companies currently selling blockbuster drugs will be able to maintain double digit growth between 2001 and 2008 (Butler, 2002). This dependence is illustrated by the following figures (Butler, 2002):

- Eli Lilly’s net profits dropped by 20% after Prozac came off patent.
- BMS’ patent on Glucophage (with previous annual sales of \$2bn) has expired and its sales are down by 87% in the first three months of 2002 after the launch of generic substitutes.
- Losec represented 34% of AstraZeneca’s sales in 2001 and it might come off patent this year if the

company is unsuccessful in its legal battle – in which case generics are expected on the market in late 2002.

in many cases to propitiate national interest in return for sympathetic pricing”.

Drivers in the pharmaceutical industry

Probably the single most important driver in the pharmaceutical industry is the time-to-market. Companies secure very significant returns in the early life of a successful drug, before any competition. The competition-free life is shortening, typically from 5 to 1-2 years. For example, Bayer’s anti-cholesterol drug Baycol was withdrawn in 2001 due to safety concerns, and the two later entrants Pravachol (from BMS) and Lipitor (from Pfizer) are now the biggest sellers for their companies (Butler, 2002).

The regulatory process tends to be slow and expensive; both these effects must be borne by the industry. Furthermore, the complex chemical compounds involved have more complex manufacturing processes, and the activities of route investigation, process development, scale-up plant design/retrofit, commissioning and qualification are either increasing in duration or proving stubborn to shorten.

An estimate of £200-400m is required to launch a new drug, and an average of 8-12 years elapses from patent filing to first sale (see, e.g., Grabowski, 1997).

There is a general trend for companies to divest excess capacity that came about from having many local manufacturing sites, and move towards a global supply chain management process. This brings with it many complex co-ordination issues and much tighter capacity constraints. Currently, the logistics cost in the sector is relatively high (Booth, 1999).

Research efficiency is declining in the sense that the cost of each new chemical entity (NCE) is increasing. Although growth in investment in R&D has exceeded inflation over the last 30 years, the global trend are as follows: 844 NCEs were registered in 1961-1970, 665 in 1971-1980 and 506 in 1981-1990 (Ballance et al., 1992). This has been one of the main drivers behind the recent series of mergers and acquisitions in the industry.

Historically, most management attention has been paid to drug discovery and sales and marketing (the extreme ends of the supply chain), but now much more attention is being paid to supply chain optimisation as a means of delivering value. According to Booth (1999),

- “there is a welcome move away from viewing the supply chain as merely having to deliver security of supply at minimum cost, to a recognition of its ability to generate both value for the customer and hence to the shareholder”
- “restructuring of the supply chain along regional and global lines will require massive reductions in capacity, which was acquired

The life cycle of a pharmaceutical product

In order to put this paper in the right context, it is important to describe the life-cycle of a drug; it is somewhat different from that of other process industry products.

The *research* or *discovery* phase tends to use thousands of more or less random test compounds against therapeutic targets. It typically takes about 10 years to result in a potential new drug that is registered. From this point onwards patent protection applies.

The potential new drug must then be tested for both safety and efficacy. This involves a variety of trials; early on for toxicity and later on for ability to alleviate symptoms and remove disease. Finally, the process development activity comes up with a chemical or biochemical route to manufacture and an associated manufacturing process. This set of activities typically takes 6-8 years and is usually known as the *development* activity.

Finally, the more familiar processes of manufacturing and distribution follow.

Components of the pharmaceutical industry manufacturing and distribution chain

A typical pharmaceutical supply chain will consist of the one or more of the following nodes:

- i. primary manufacturing (possibly including contractor sites)
- ii. secondary manufacturing (possibly including contractor sites)
- iii. market warehouses/distribution centres
- iv. wholesalers
- v. retailers/hospitals

Primary manufacturing

The primary manufacturing site is responsible for the production of the active ingredient (AI or API). This normally involves either several chemical synthesis and separation stages to build up the complex molecules involved, or fermentation and product recovery and purification in the case of biochemical processes.

The manufacturing process is characterised by long task processing times, often rounded to multiples of shifts. Where multistage processes are operated, considerable inventories are often held between stages. Furthermore, material from an intermediate stage must often pass some form of quality control check before being approved for use downstream in the process. This can introduce additional delays into the system.

The traditional process technology involves batch equipment and flexible pipework. The relatively low production volumes result in multipurpose plants to spread

the capital cost between products. The need to avoid cross-contamination of products and requirements for validated cleaning and changeovers results in long downtimes between products. These have been of order 4 weeks in the past, but the application of techniques similar to the single-minute exchange of die (SMED) methods (see, e.g. Moser et al., 2000) applied to the car industry have reduced these somewhat. These downtimes in turn imply that long campaigns are the norm, otherwise equipment utilisation is too low. It is not unusual for one year's production of a product to be produced in a single campaign, and the material produced being stored until the next campaign in the following year. Since most complex pharmaceuticals are produced through multistage processes, the same often holds true for the stable intermediates (stage products). Needless to say, this mode of operation does not lend itself well to responsiveness, and contributes significantly to some of the poor supply chain metrics exhibited by this industry.

A further source of complexity (and convenience) is the use of contractors to manufacture some or indeed all of the active ingredient stages. This process of outsourcing is a growing one, as research-oriented companies concentrate on the discovery and development activities and rely on third parties' manufacturing competence. This gives rise to extended supply chain co-ordination problems.

Secondary manufacturing

This is concerned with taking the active ingredient produced at the primary site and adding "excipient" inert materials along with further processing and packaging to produce the final products, usually in SKU form. For example, a product that is sold in pill form would undergo:

- i. granulation – with addition of all the incipient materials
- ii. compression – forming the pills
- iii. coating
- iv. quality control
- v. packaging

The secondary manufacturing locations are often geographically separate from the primary manufacturing locations. This is frequently the outcome of tax and transfer price optimisation within the enterprise. There are often many more secondary manufacturing sites than primary ones, serving local or regional markets. Transportation between sites is of the order of one or two weeks if by ship (usually the default mode) and of the order of one or two days if by air.

Wholesalers play a significant role in this sector. They tend to be large and few. About 80% of demand flows through this channel in the UK (with three large players accounting for almost all the demand), with the large part of the remainder going to hospitals. In the US another

intermediary is growing – the managed care organization (MCO) or healthcare maintenance organization (HMO).

Operational issues in the pharmaceutical supply chain

Although the processes will vary between companies, all major pharmaceutical companies will operate ERP systems and follow a business process along the following lines:

- Demand management – in each geographical region, forward forecasts (e.g. 3-24 months) are developed, based on historical data, market intelligence etc. Tenders for manufacture may also be evaluated and possibly accepted at this stage.
- Inventory management and distribution requirements planning – the demands determined are aggregated and imposed on the appropriate warehouse/distribution centre. The impact on finished goods inventory is assessed and if necessary, orders are placed on upstream secondary manufacturing sites.
- Secondary production planning and scheduling – the orders placed on the secondary sites are planned (typically using MRP-II type tools) and then scheduled in detail (typically using APS tools). The impact of production plans on active ingredient (AI) raw material stocks is evaluated and if necessary, orders for AI are placed on the upstream.
- Primary manufacturing campaign planning and AI inventory management. Here, the demands placed by secondary manufacturing are satisfied by careful management of inventory and production planning.

An interesting feature of this process is that the customer-facing end is effectively a "pull" process (driven by orders) but the primary manufacturing stage has long cycle times which make it difficult to ensure end-to-end responsiveness. This means that primary production is effectively a "push" process, driven by medium- and long-term forecasts. Relatively large stocks of AI must be held to ensure good service levels and ensure smooth operation at the interface of these processes. The well-documented "bullwhip" or Forrester effect is often felt at the primary manufacturing site, which is unfortunate since this is the least responsive part of the supply chain as it normally operates in campaign mode. This makes it difficult to exploit short-term opportunities (e.g. shortages of supply of a competitor's product, tenders for national supplies, epidemics, etc.).

Another feature of this process is an outcome of its large scale and geographical span. This is the distributed nature of decision-making, which can lead to tensions and sub-optimal decisions. Different nodes are not really aware of upstream nodes' resource constraints, and orders

may be filled in order of receipt, rather than on an economic basis. Of course, centralised planning would not be without its difficulties in this context.

In our experience, the following supply chain performance measures are typical of the industry:

- The stock levels in the whole chain (“pipeline stocks”) typically amount to 30-90% of annual demand in quantity, and there are usually 4-24 weeks’ worth of finished good stocks.
- Stock turns (defined as annual sales/average stock) are typically between 1 and 8.
- Supply chain cycle times (defined as elapsed time between material entering as raw material and leaving as product) are often between 1000-8000 hours.
- The value-added time (time when something happens to material as a percentage of chain cycle time) is of order 0.3-5%.
- Material efficiencies (the amounts of product produced per unit amount of total materials used) are 1-10%.

The relatively high levels of stock are required to buffer the slow supply chain against market dynamics.

Strategic and design issues in the pharmaceutical supply chain

The decisions to be taken at this level include:

- Pipeline and development management – this involves the selection of potential drugs to develop further, and the planning of the development activity.
- Process development – the investigation of manufacturing routes and the generation of manufacturing processes.
- Capacity planning and plant and supply chain network design
- Plant design – the selection and sizing of the major equipment and storage units.

Some of the key issues are:

- Uncertainty in the demands for existing drugs (due to competition, uncertainty in the ability to extend the protected life through new formulations etc.)
- Uncertainty in the pipeline of new drugs – in particular, which ones will be successful in trials, what sort of dosage and treatment regime will be optimal.

- Process development – this is a complex problem, driven by chemistry and yield optimisation. It often results in inefficient processes that are operated much more slowly than the intrinsic rates – giving rise to batch processes and long cycle times responsible for some of the problems seen at the primary production planning stage.
- Capacity planning – the long lead times to make capacity effective mean that decisions often need to be taken at times of high uncertainty. Waiting for the uncertainties to be resolved might delay the time to market by an unacceptable amount.
- Network design – often tax implications take precedence over logistics issues, these result in economic but potentially complicated supply chains.
- Plant design – this tends to be very traditional, with no real change in manufacturing technology for 50 years (the workhorse of the primary manufacturing site is the glass-lined stainless steel batch reactor). There are significant opportunities for intensified, continuous processing.

Overview of some recent work

The recent work in the literature that is relevant here can be categorised under these headings:

- Pipeline and development management
- Capacity planning
- Simultaneous development and capacity planning
- Process development and plant design
- Production planning and scheduling
- Supply chain simulation and dynamics

The areas are reviewed in turn below.

Pipeline and development management

Schmidt and Grossmann (1996) considered the problem of sequencing of testing tasks where unlimited resources are assumed to be available. The key feature of the model that distinguishes it from classical project scheduling is that each task has a probability of failure; this affects the need for successor tasks. They formulated the problem as a continuous-time MILP and solved the problem of maximising the overall expected NPV. If many tests for a product are performed in parallel, the testing activity will be more expensive, as the effect of failures on successor tests are not taken into account. On the other hand, the product may come to market much earlier, resulting in a much better cash flow profile. Conversely, sequential test planning might avoid unnecessary tests and

reduce expense, but result in a later arrival of the product in the marketplace.

In practice, the testing activity tends to be quite resource-constrained, and may also involve outsourcing of some stages. Jain and Grossmann (1999) develop a methodology for the sequencing and scheduling of testing tasks under resource constraints. In this approach, each product has a specified set of testing tasks. Each task is characterised by a duration, cost, precedence constraints, resource requirements and probability of success. A task may be outsourced at a higher cost; in this case no internal resources are required. The income associated with a product is given as a function of the time of launch in the market.

The formulation developed is conservative and always feasible in that the resource constraints are always enforced, regardless of the probability of a task not actually taking place. The cost component is modelled as an expected cost. This ensures that the effect of starting tasks earlier than necessary is modelled; i.e. that later tasks may not actually take place due to the failure of the earlier one. Two alternative formulations are presented, a continuous-time MILP and a graph-based one. The latter was found to be more efficient and was able to cope with problems consisting of 30 tasks.

Blau et al. (2000) consider the problem of risk management at the development stage. As mentioned earlier, the development phase selects candidate drugs and takes them through trials and process development. It is a long, costly and inherently risky process with a large up-front commitment. The aim of this work is to support the process of product selection and test planning while managing risk effectively. The development activities are modelled as a probabilistic activity network, where each activity has a time, precedence relations, resource requirements and probability of success.

Risk is defined as the adverse consequences of exposure to uncertainty, and in this context is usually related to the premature withdrawal of a candidate drug. The risk of a set of decisions must be balanced against the potential reward. In this case, the potential reward is the expected financial returns of drugs that do make it through the development process. The risk/reward ratio can then be used to compare different drug candidates. A screening process removes any obviously unpromising candidates, and then the remainder must be sequenced through the development pipeline. A heuristic approach using simulation with local rules in response to trigger events (e.g. failure of a test) is employed. This aims to process tasks as quickly as possible and although there is no guarantee of not violating resource constraints, these violations are usually not large.

Subramanian et al. (2001) extend this work to take explicit account of the resource requirements of the problem. The problem statement is generalised in that more sources of uncertainty in the problem are considered, and include:

- task processing times
- task resource requirements
- task success probabilities
- task costs
- market returns

They make the point that a single-level mathematical programming problem cannot hope to capture all these features. On the other hand, discrete-event dynamic systems (DEDS) techniques cope well with the stochastic elements, but require local, myopic rules to resolve conflicts or make choices as they arise. They therefore developed an integrated optimisation-simulation framework (SIM-OPT), where a DEDS simulator reverts to an optimisation layer (with different degrees of optimisation) to resolve conflicts or make choices such as task sequencing. The optimisation layer is an MILP which is updated by the latest status of the plant. The results show that using optimisation far outperforms the typical local rules used in classical DEDS. By repetitive simulation, the statistical trends can be tracked and answers to questions about corporate policy (particularly in relation to risk and resourcing) can be obtained.

Capacity planning

The capacity planning under clinical trials uncertainty problem has recently received some attention in the literature. The deterministic problem of allocating new manufacturing capacity to existing or potential sites around the world is described by Papageorgiou et al. (2001). They describe the features particular to the pharmaceutical industry, and emphasise the importance of modelling financial flows as the taxation regimes affect the rewards associated with alternative solutions significantly. Indeed, taxation considerations can easily dominate the location decisions.

The problem of capacity planning under uncertainty was considered by Rotstein et al. (1999). They considered the problem where three products are at the start of clinical trials, and plans for current and future manufacturing capacity are to be made. The key trade-off in the capacity planning decision comes about due to the lead time between deciding to invest in additional manufacturing capacity, and that capacity coming on-stream. Deferring capacity planning decisions until more information is available from trials is obviously a lower-risk strategy, but increases the time to market. As mentioned earlier, this measure is critical. For example, when Tagamet came to the market in the 1970s, it was free from competition for at least 5 years, but now this competition-free time is often only 1-2 years.

Rotstein et al. (1999) use a scenario tree to capture the outcomes of the trials, and use a two-stage stochastic programming with recourse formulation to model the problem. The “here-and-now” decisions related to immediate capacity expansions and the “wait-and-see” decisions depend on trial outcomes and include further

capacity expansions, plant or product abandonment and production and inventory planning. They show how different options can be compared using a number of metrics, including expected NPV, the probability of the NPV being negative, the worst case scenario, the total demand met of all the potential products and the total demand met of the products chosen from the portfolio.

Gatica et al. (2001, 2002a) extend this work to the case where different products are at different stages in their life-cycles, and those that are in trials will complete those trials at different times. This gives rise to a much more complicated scenario structure – it is a multistage stochastic optimisation problem, with each stage reflecting the completion of a clinical trial. In contrast to most of the work reported on trials in the literature, this work uses more than two outcomes (success or failure) for the completion of the trial. Based on typical practices in industry, four outcomes (failure, low, target and high) are used. This means that four scenarios are required per stage, and for a problem with N stages (i.e. N products in trials), there is one scenario in the first stage (reflecting products currently in the market with well-forecasted demands), four scenarios in the second stage (reflecting the four possible outcomes for the first pipeline product to come out of trials), 16 in the third stage (the combinations of outcomes for the two products to complete trials) and so on, until the final stage which has 4^N scenarios. Each scenario has associated with it possible capacity expansions and production and inventory planning variables and constraints, so overall the problem becomes a large scale stochastic programming problem with integer and continuous decisions. It is solved as its deterministic equivalent, a large MILP. This is relatively straightforward to solve for the 4-product case the authors report.

Clearly, the approach is limited by the complexity of the scenario tree, so Gatica et al. (2002b) extend it using a scenario aggregation procedure similar to that of Clay and Grossmann (1997) to enable the solution of larger problems.

Simultaneous development and capacity planning

The research reviewed above deals either with the problem of organizing the development (testing) activities or planning the capacity investments and future production. Maravelias and Grossmann (2001) consider the problems of planning of testing tasks and capacity simultaneously. The aim is to optimize a performance measure (expected NPV) for the process as a whole.

This bridges the gap between the two problems and aims to ensure that the company is ready to produce a product once testing is complete (if the product is successful).

The testing process is modeled as a set of tasks with technological precedence constraints, durations and resource requirements. The tasks have two possible outcomes, success or failure. All tasks other than process development may be outsourced if internal resources do

not suffice. Since the method is to be applied at any time in the company's operation, it takes account of the fact that different products will be at different stages in their life cycle. The testing network is probabilistic in that each test has a probability of being passed. If at any stage a product fails a test, it is abandoned.

The manufacturing process is assumed to have existing capacity as well as potential new capacity. So, overall, the decisions to be taken are:

- The selection of products for testing
- The assignment of resources to testing tasks and any outsourcing decisions
- Task sequencing
- Selection of new plants or expansions of existing ones (including timings of expansions)
- Production planning

However, production only takes place if a product successfully completes its tests. This results in a stochastic problem. Since the uncertainties are discrete, this is well represented by a scenario tree. A large scale MILP results; this is solved by a Lagrangean decomposition scheme.

Risk in pharmaceutical supply chain infrastructure decisions

It is clear that much of the infrastructure-related work (in particular the product selection and development and capacity planning decisions) is subject to considerable risks. These include product failures during trials, product withdrawal during sales due to side effects, uncertainties about final dosage and treatment regimes, competition from similar products etc. These are in addition to the background demand uncertainty associated with the indication.

Most strategic/infrastructural decisions have historically been based on NPV or some form of expected NPV, which in turn utilise weighted average costs of capital or some required return on investment. The problem with the expected NPV measure in this context is that the risks tend to be few, significant and discrete in nature; the expected NPV term is better suited to the situation where uncertainties are many (each of which is relatively "small") and continuously distributed. Keynes succinctly summarises the problem with the expected NPV approach in this context as being based on the assumption that "a certain state of mediocrity is as desirable as an even chance of heaven or hell".

There are a number of approaches to extending the expected NPV metric to deal with risk. Perhaps the most classical of these is the employment of a "mean-variance" type objective function (see, e.g., Mulvey et al., 1997), which is of the form:

$$\max Z = \alpha \text{ mean}(r) - (1-\alpha)\text{var}(r)$$

where r is the reward from the project and α is a parameter to trade off the relative importance of the expected return and its variability.

This form is not particularly well-suited to the pharmaceutical industry for at least three reasons:

- i. the focus in this sector is particularly on downside risk rather than variability of return;
- ii. the distributions of reward associated with a particular project tend to be bimodal; with one mode reflecting failure at some stage and the other success;
- iii. the one-sided nature of the NPV measure means that a large variance is not necessarily a bad thing if the downside risk is low.

Alternative formulations focus on downside risk, and can be of the form of constraints enforcing a maximum probability of the reward being less than a particular figure (see, e.g. Kall and Wallace, 1994), i.e., enforcing:

$$\text{Prob} \{(r) \leq r^0\} \leq \beta$$

where r^0 is a minimum threshold return and β is the maximum allowed probability that the actual reward r is below r^0 . In many cases, the most important sources of uncertainty are discrete (due to product failure). This gives rise to a discrete outcome space, and a constraint can then be imposed on the worst case, i.e. the scenario that leads to the lowest reward.

Eppen et al. (1987) develop a “risk factor” based on expected downside risk. This gives gives a measure of the failure to meet a certain target profit. The risk factor is easiest understood in a discrete scenario context. Here, it is calculated as:

$$RF = \sum_{k:r_k \leq r^0} \text{Pr}_k (r^0 - r_k)$$

where k is the scenario index and Pr_k is the probability of scenario k . An upper bound can then be enforced on RF . By tightening the constraint on expected downside risk, it is possible to bring alternative solutions to the attention of the decision makers. This is significant for problems that contain many different solutions. The higher the probability of occurrence of a reward below r^0 , the higher RF is, and the risk involved in the project is larger. If RF is close to zero, then the risk associated with each investment decision is very low, making the investment more attractive from a risk perspective. This risk model is used in the problem of capacity planning for products at different stages in clinical trials by Gatica et al. (2002a).

Applequist et al. (2000) describe a risk premium approach which sets out to find the right balance between the expected value of a set of decisions and the associated risk (captured by the variance in this case). The expected return of an investment decision is compared to one in the

financial market with a similar variance (e.g. government bonds, large company stocks, small company stocks) or a model that fits the expected return-variance correlation. For any investment decision to be approved, its expected reward for the associated variance should be better than that possible in the market. They tackle the problem of determining capital investments and production plans for a process with uncertain (continuously-distributed) product demands. The problem of evaluating the expected return and its variance is very complicated and the polytope volume integration procedure developed is one of the key contributions.

Bhagwat and Griggs (1995) undertook a study of the riskiness of the industry – this was prompted by the perceived above normal rates of return associated with the industry. They surmised that using a risk premium alone will not suffice as there are also some systematic risks over and above those captured by a market measure like risk premium. They utilise the capital asset pricing model (CAPM) to estimate the systematic risk. Here, the required rate of return from an asset i is

$$K_i = R + \beta_i(E(r_m) - R)$$

Where R is the riskless rate of return, $E(r_m)$ is the expected market rate of return and $E(r_m) - R$ is the risk premium. β_i can be thought of as the systematic risk associated with the i th asset. For a systematic risk to be present, this parameter should have a value large than 1. In their study of the US pharmaceutical industry from 1963-1992, Bhagwat and Griggs found an average value of 1.05, with recent values being higher than in the past.

Booth (1999), on the other hand, questions the use of the CAPM to set the required return as it is merely a hypothesis and there is uncertainty about its central premise, namely that portfolios of investments with higher risks will show higher expected returns.

Myers (1999) notes that the cost of capital must also be related to the stage in the life cycle of the associated product(s).

A promising area of research to augment these is the application of real options theory to this area. Lerwent (1994) has considered the case where the initiation of an R&D activity can be seen as purchasing an option to continue with development until further information is obtained. The pricing of the option in this context is a very complicated and open research issue.

Process development and plant design

The problem of process development in this sector has recently been reviewed by Shah et al. (2000). They contend that the current practice of relying on traditional manufacturing technology means that processes are designed to be operated in potentially ineffective ways. A

hierarchical approach to designing processes that are not constrained by traditional equipment is recommended by the authors.

Linninger et al. (1996) also propose a hierarchical approach, with the emphasis on the use of knowledge bases and material balancing at every development level to choose and assess options. The focus is particularly on synthesising routes and developing processes with low environmental impacts. The integration of these methods into a software environment supporting process development is described by Stephanopoulos et al. (2000).

The problem of multipurpose and multiproduct batch plant design is particularly relevant to this area. There has been a large amount of work from 1970 to the present date (see, e.g. Sparrow et al., 1975; Yeh and Reklaitis, 1987; Papageorgaki and Reklaitis, 1990; Shah and Pantelides, 1991, 1992; Voudouris and Grossmann, 1992; Henning et al., 1994; Barbosa-Povoa and Pantelides, 1999). Certainly there are a number of good solutions to the problem of allocating the best set of equipment to processing tasks assuming campaign operation and batch processing. A set of interesting challenges in process and plant design remain ahead of us: the design of novel, intensified equipment which will allow processes to operate at intrinsic rates, and increase manufacturing velocities by orders of magnitude.

Production planning and scheduling

This has also received a lot of attention over the period from 1970 to the present day (see, e.g., reviews by Reklaitis, 1991 and Shah, 1998). However, most of this is on short-term, order-driven scheduling. Most of the work on campaign planning was done some time ago (e.g. Mauderli and Rippin, 1980; Lazaro et al., 1989; Shah and Pantelides, 1991; Papageorgiou and Pantelides, 1996). However the active ingredient production process, which is effectively the rate-limiting step of the supply chain still tends to operate this way. The optimal planning of campaigns within the context of the performance of the supply chain as a whole (in particular trading off the cost and inconvenience of changeovers with overall supply chain responsiveness and inventory) has not really been studied for this type of industry.

Supply chain simulation and dynamics

The long supply chain and the fact that there are many decision-making agents means that understanding the dynamic behaviour will be very important. We have developed a generic approach to the modelling of supply chain dynamics and applied it to some pharmaceutical processes (Gjerdrum et al., 2000). We have not found any other work reported in the literature on this topic, but a similar simulation study of the food supply chain was undertaken by van der Vorst et al. (2000).

Our work is concerned with the modelling of both the physical processes (primary and secondary manufacturing,

distribution and warehousing) as well as the **business** processes. By the latter, we mean how decisions are taken at the different nodes of the chain, who takes them, what tools/methods are used etc. The aim is to replicate the behaviour of the supply chain in software. This means that the logic of software tools used for decision-making at various nodes (e.g. DRP and MRP methods) are replicated in our simulation tool. It is clear that a purely analytical model cannot capture this information easily. The aim of this approach is to suggest non-invasive improvements to the operation of the supply chain (i.e. neither the physical or IT infrastructure should be modified). Such improvements may come about through changes in parameters (e.g. safety stocks) or business processes (e.g. relationships between agents).

Once a model has been developed, it is validated against historical data and then used to perform a variety of what-if studies. In order to assess future performance, uncertainties need to be taken into account. At the operational level, these include product demands, process yields, processing times, transportation lead times etc. A stochastic simulation approach that samples from the uncertain parameters is a useful way of determining expected future performance as well as confidence limits on future performance measures (e.g. service levels). The results of two such studies are described below.

In the first study, a peculiar dynamic behaviour was seen in the market warehouse. Although the background demand for the product was very stable, the manufacturer's warehouse experienced highly fluctuating demands, and needed to hold considerable inventories to buffer against this. Upon some investigation, the reason related to a pricing cycle which caused wholesalers to try to anticipate price increases and request large pre-emptive orders. Of course, once these are received, the wholesaler will not order material again for some time. We used singular-value decomposition techniques to extract the historical dynamics and used them to generate forward forecasts. We compared the future supply chain performance using this model against a model that used collaborative planning between manufacturer and wholesaler. The key metric was the amount of finished goods safety stock cover the manufacturer required to meet a certain customer service level. The results may be compared below:

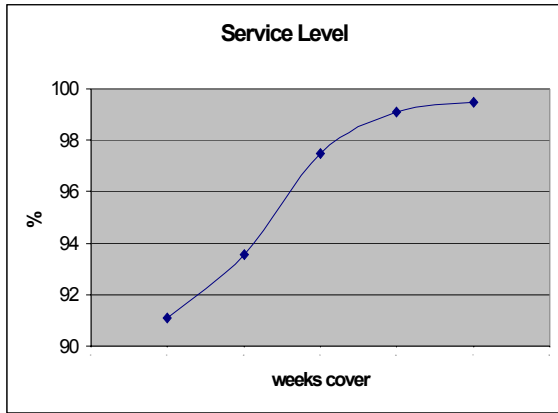


Figure 1. Variation of service level with finished goods safety stock for the non-collaborative case.



Figure 2. Variation of service level with finished goods safety stock for the collaborative case.

To achieve the target service level, the finished goods stock can be approximately halved in the collaborative case. It is clear that significant benefits are possible through an alternative way of running the supply chain. Conservative estimates would put these at \$30m in one-off inventory savings and \$3.6m p.a. savings.

Another study considered the effects not of the production or inventory aspects, but the quality control (QC) procedures. As mentioned earlier, there is a prevalence of QC activities in the industry, although they are not really necessary at all the points where currently

used. These steps account for significant dead time in the process; often of the order of one to two weeks, when all the intervening processes are considered. We developed a model of a process which has five primary synthesis stages and two secondary manufacturing sites. The as-is process has QC activities at the end of each primary stage, and for the final product. The modified process has a QC step for the AI and a QC step for the final product. The results for one of the products are compared below.

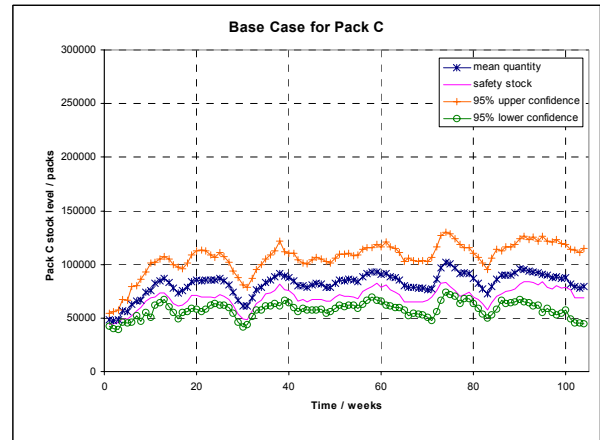


Figure 3. Time profile of expected inventory of finished goods, including confidence intervals, for the case with quality control at only two points.

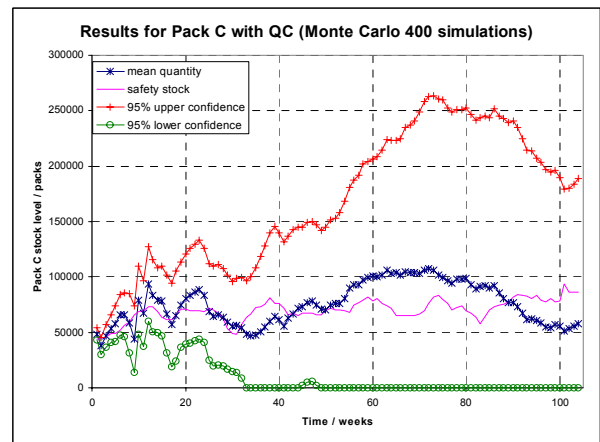


Figure 4. Time profile of expected inventory of finished goods, including confidence intervals, for the case with quality control at all stages.

In figure 3, the forward prediction of finished goods inventory of pack C is quite smooth. The lower confidence (95%) limit on the profile is still positive, giving

confidence that stock-outs are very unlikely. In figure 4, the case with QC at all stages, there is much less certainty in the inventory (the variance grows significantly with time) and the lower confidence level goes to zero. In terms of customer service performance measures, over a two-year period, the average service level in the case with QC is 91% and the probability of a stockout in any week is 5%. On the other hand, in the low QC case, these figures are 100% and 0% respectively. Clearly, as process development and design advance, the comfort provided by QC at so many stages in the supply chain will not be required and the dynamic behaviour will improve markedly.

Overall, the simulation based approach is very promising for studying the large and complex supply chains involved. Other studies include when to plan shutdowns in manufacturing sites and strategies for new product introduction (Gjerdrum et al., 2001)

Future challenges

The future challenges in this area are broad and complex, and will provide fertile ground for research. They can be categorised under three headings:

- Improvements to existing processes
- Improvements to the strategic decision-making process
- Future scenarios

Improvements to existing processes and operations

The supply chain includes many agents, often with different objectives. Their internal dynamics tend to exaggerate the external market dynamics and result in detriments in performance. This is an area where collaborative forecasting, planning and inventory management will be very useful. Here, the different agents in the supply chain will co-ordinate activities across the chain. One of the main reasons for the current, more distributed practice is the large scale of the operations, both in terms of activities and geographical span. Multisite planning and scheduling tools are required to support a collaborative planning activity. The supply chain generally contains larger amounts of inventory than might be necessary if a more co-ordinated approach is followed, with the right supporting tools.

Probably the most important metric to track and try to improve is that of the overall supply chain cycle time. As mentioned earlier, figures of 100-300 days are common. Our experience in the industry is that efforts to reduce this are very effective, but they have been applied in a piecemeal fashion to certain products, while the measure deteriorates for others. A systematic analysis of the components of the cycle time should be undertaken for a wide range of products. Clearly, large scale simulation

tools that take account of both physical and business processes would be useful in this context.

Improvements to the strategic decision-making process

We can consider this in a bottom-up fashion. The current nature of the process technology is one the main supply chain bottlenecks. There is a need for more agile equipment which will shorten process cycle times by an order of magnitude and require minimal time for cleaning and changeover. This will avoid long campaigns and should lead to “pull” based active ingredient manufacturing, and therefore more responsive supply chains. The underlying processes will have to change as well, with the focus being on designing processes that operate at intrinsic rates (e.g. being limited by reaction kinetics) rather than being limited by equipment performance (e.g. heat transfer, mass transfer or mixing characteristics) of traditional equipment. In principle, the low tonnages involved should lead to much less capital intensive plants if this is achieved. Generally speaking, significant improvements to manufacturing technology have not been of the highest priority in this field to date.

Processes should be designed with a much greater level of mechanistic understanding and controlled tightly if reductions in quality control activities are to be possible. The impact of this has been demonstrated earlier.

The integration of development management and capacity and production planning will be very important. Currently, capacity issues are often not considered at the development stage. Booth (1999) lists three undesirable outcomes of not co-ordinating these activities:

- i. shortages of materials for pre-clinical studies;
- ii. shortages of materials for clinical trials;
- iii. delays in time-to market – which is not only dependent on having material available, but also on generating demand at the late stages of development.

The approach of Maravelias and Grossmann (2001) is a very promising platform to treat the overall pipeline/capacity planning problem, but it needs to be integrated with a sophisticated treatment of risk (e.g. through the use of real options theory) and economics (e.g. taking account of local taxation regimes, transfer pricing, duty drawbacks etc.), both of which have a very significant impact on investment decisions. Furthermore, the modelling of testing and trials needs to be extended to account for variations in standard outcomes, e.g. very successful trials resulting in short-circuiting of the approvals process in the case of life-saving drugs.

Generally speaking, the development of integrated models of the life-cycle, from discovery through to consumption would greatly facilitate strategic decision-making.

Future scenarios

Companies have recently moved away from product diversification and locally-adapted products. The common packaging and labelling standards in the European Union, for example, have supported this. However, there has been much development in the field of “pharmacoeconomics” which might generate pressures to reverse this trend. This discipline helps to make choices in treatment options by consideration of costs and outcomes (clinical, economic and humanistic). An important outcome of this type of analysis will be the insistence on local solutions to local problems (Thwaites and Townsend, 1998).

Another trend, somewhat further down the line, will come out of genetic research and which will identify target sub-populations for different treatment regimes (the so-called “designer drugs”). These two drivers will give rise to considerable product and supply chain complexity. The current manufacturing processes and supply chains are not well configured to cope with this. Primary manufacturing tends to operate in campaigns; and secondary manufacturing batchsizes are typically 1-4 million tablets. Clearly, the manufacturing processes and supply chains will have to be re-designed if product customisation is to increase in line with these trends. The manufacturers will have to track supply chain performance measures very carefully in order to understand the cost-to-serve for a diverse customer base.

An emerging area is that of rapid response vaccines and other treatments arising out of possible emergencies (e.g. bioterrorism or very fast developing epidemics). Again the traditional supply chain (particularly for vaccines) is very slow and unresponsive. If national governments are to implement emergency preparedness programmes, the entire infrastructure must be well designed and tested through simulation. Decisions such as where to manufacture, in what quantities, where to hold stocks, where people should report etc. need to be taken in a robust fashion. The issue of how well the supply chain measures up to emergency preparedness concerns is being raised in various fora at the present (see, e.g., Anon, 2002).

Companies are investing in the development of crops that are designed in some way to produce pharmaceuticals; this will give rise to new research activities in process (in particular recovery steps) innovation, novel equipment design (e.g. supercritical separations) and of course in a new type of supply chain to optimise.

There are also some changes afoot in the industry structure as well – the growth in medicinal chemistry, biotechnology and genomics will spread IP around and result in looser, virtual enterprises of joint ventures, alliances etc. Including as well the general tendency towards outsourcing of manufacturing, this will give rise to complex extended supply chain co-ordination. Lessons may be learnt from the automotive, PC and consumer

electronics world, where such supply chains already operate.

Conclusions

The pharmaceutical supply chain used to be seen as a tool to supply products to market in an effective way, where the emphasis was on security of supply. Recent changes in the operating environment mean that companies are revisiting the components of their supply chains and identifying ways of extracting additional benefits from them.

In this sector in particular, the supply chain of interest is not simply the physical processes of conversion and distribution of materials. Equally important is the “value-chain” perspective of managing the innovation and development processes through to capacity and production planning. There are still several exciting research challenges in this value chain, many of which the process engineering/process systems engineering community are well placed to address.

Acknowledgements

The author is very grateful to Gabriel Gatica, Jonatan Gjerdrum and Lazaros Papageorgiou for their contributions to this work.

References

- Anon (2002). Summary of the executive session on emergency preparedness and the pharmaceutical supply chain. *Am. J. Health-System Pharmacy*, **59** (3), 247-253.
- Applequist, G.E., J.F. Pekny and G.V. Reklaitis (2000). Risk and uncertainty in managing chemical manufacturing supply chains. *Comput. Chem. Eng.*, **24**, 2211-2222.
- Ballance, R. J. Pogany and H. Forstner (1992). *The World's pharmaceutical industries*. Edward Elgar (for UNIDO), UK.
- Barbosa-Povoa, A.P.F.D. and C.C. Pantelides (1999). Design of multipurpose production facilities: A RTN decomposition-based algorithm. *Comput. Chem. Eng.*, **23**, S7-S10.
- Bhagwat, Y. and F.T. Griggs (1995). Analysis of riskiness of pharmaceutical industry firms. *J. Res. Pharm. Econ.*, **6**, 65-76.
- Blau, G., B. Mehta, S. Bose, J. Pekny, G. Sinclair, K. Keunker and P. Bunch (2000). Risk management in the development of new products in highly regulated industries. *Comput. Chem. Eng.*, **24**, 1005-1011.
- Booth, R. (1999). *The global supply chain. FT healthcare management report*. Financial Times Business Ltd, London.
- Butler, R. (2002). The end of the blockbuster? *Chemistry & Industry*, **9**, 9-10.
- Clay, R.L. and I.E. Grossmann (1997). A disaggregation algorithm for the optimization of stochastic planning models. *Comput. Chem. Eng.*, **21** 751-774.
- Eppen, G.D., R.K. Martin, and L. Schrage (1987). A scenario approach to capacity planning. *Op. Res.*, **37**, 517-527.

- Gatica, G., N. Shah and L.G. Papageorgiou (2001). Capacity planning under clinical trials uncertainty for the pharmaceutical industries. *Proc. ESCAPE-11, Kongdig, Denmark*, 865-870.
- Gatica, G., L.G. Papageorgiou and N. Shah (2002a). Capacity planning under uncertainty for the pharmaceutical industry. *Submitted to Trans. IChemE. Part A*.
- Gatica, G., L.G. Papageorgiou and N. Shah (2002b). An aggregation approach for capacity planning under uncertainty for the pharmaceutical industry. *Submitted to FOCAP0-2003*.
- Gjerdrum, J., Q.W.Z. Jalisi, L.G. Papageorgiou and N. Shah (2000). Dynamic simulation of physical and business processes for supply chain improvement. *Proc. 5th Annual Conf. Ind. Eng. Theory, Applications and Practice, Hsinchu, Taiwan*.
- Gjerdrum, J., N. Shah and L.G. Papageorgiou (2001). New Product Introduction Supply Chain Planning. In Olli-Pekka Hilmola (editor): *Contemporary Research Issues in New Product Introduction. Acta Wasaensia*, **77**, 25-49.
- Grabowski, H. (1997). The effect of pharmacoeconomics on company research and development decisions. *Pharmacoeconomics*, **11**, 389-397.
- Henning, G.P., N.B. Camussi and J. Cerda (1994). Design and planning of multipurpose plants involving nonlinear processing networks. *Comput. Chem. Eng.*, **18**, 129-152.
- Jain, V. and I.E. Grossmann (1999). Resource-constrained scheduling of tests in new product development. *Ind. Eng. Chem. Res.*, **38**, 3013-3026.
- Kall, P., and S.W. Wallace (1994). *Stochastic Programming*. Wiley, New York.
- Lazaro, M., A. Espuna and L. Puigjaner (1989). A comprehensive approach to production planning in multipurpose batch plants. *Comput. Chem. Engng.*, **13**, 1031-1047.
- Lerwent, J. (1994). The new pharmaceutical paradigm: scientific management at Merck. *Harvard Business Review*, **Jan-Feb**, 88-89.
- Linninger, A.A., S. Ali and G. Stephanopoulos (1996). Knowledge-based validation and waste management of batch pharmaceutical process designs. *Comput. Chem. Eng.*, **20**, S1431-S1436.
- Maravelias, C.T. and I.E. Grossmann (2001). Simultaneous planning for new product development and batch manufacturing facilities. *Ind. Eng. Chem. Res.*, **40**, 6147-6164.
- Mauderli, A. and D.W.T. Rippin (1979). Production planning and scheduling for multipurpose batch chemical plants. *Comput. Chem. Eng.*, **3**, 199-206.
- Mauderli, A. and D.W.T. Rippin (1980). Scheduling production in multipurpose batch plants : the Batchman program, *Chem. Eng. Progress*, **4**, 37-45.
- Moser, M. G. Calderari, P. Morini (2000). Cleaning validation of a multipurpose plant for active pharmaceutical ingredient bulk production. *Chimia*, **54**, 731-733.
- Mulvey, J.M., D.P. Rosenbaum and B. Shetty (1997). Strategic financial risk management and operations research. *EJOR*, **97**, 1-16.
- Myers, S. (1999). Measuring pharmaceutical risk and the cost of capital. In: *Risk and return in the pharmaceutical industry*; Eds. J. Sussex and N. Marchant, OHE., London, 59-76.
- Papageorgaki, S. and G.V. Reklaitis (1990). Optimal design of multipurpose batch plants, *Ind. Eng. Chem. Res.*, **29**, 2054-2062.
- Papageorgiou, L.G. and C.C. Pantelides (1996). Optimal campaign planning/scheduling of multipurpose batch/semicontinuous tasks. 1. Mathematical formulation. *Ind. Eng. Chem. Res.*, **35**, 488-509.
- Papageorgiou, G.E. Rotstein and N. Shah (2001). Strategic supply chain optimization for the pharmaceutical industries. *Ind. Eng. Chem. Res.*, **40**, 275-286.
- Reklaitis, G.V. (1991). Perspectives on scheduling and planning of process operations. *Proc. PSE-1991*, Montebello, Canada.
- Rotstein, G.E., L.G. Papageorgiou, N. Shah, D.C. Murphy and R. Mustafa (1999). A product portfolio approach in the pharmaceutical industry. *Comput. Chem. Eng.*, **23**, S883-S886.
- Schmidt, C.W. and I.E. Grossmann (1996). Optimization models for the scheduling of testing tasks in new product development. *Ind. Eng. Chem. Res.*, **35**, 3498-3510.
- Shah, N. and C.C. Pantelides (1991). Optimal long-term campaign planning and design of batch plants. *Ind. Eng. Chem. Res.*, **30**, 2308-2321.
- Shah, N. and C.C. Pantelides (1992). Design of multipurpose batch plants with uncertain production requirements. *Ind. Eng. Chem. Res.*, **31**, 1325-1337.
- Shah, N. (1998). Single- and multisite planning and scheduling: current status and future challenges. *AIChE Symp. Ser.* **320(94)**, 75-90.
- Shah, N., N.J. Samsatli, M. Sharif, J.N. Borland and L.G. Papageorgiou (2000). Modelling and optimisation for pharmaceutical and fine chemical process development. *AIChE Symp. Ser.* **323(96)** 31-45.
- Sparrow, R.E., G.J. Forder and D.W.T. Rippin (1975). The choice of equipment sizes for multiproduct batch plants: heuristics versus branch-and-bound. *Ind. Eng. Chem. Process Des. Dev.*, **14**, 197-203.
- Stephanopoulos, G., S. Ali, A. Linninger and E. Salomone (2000). Batch process development: challenging traditional approaches, *AIChE Symp. Ser.* **323(96)** 46-57.
- Subramanian, D., J.F. Pekny and G.V. Reklaitis (2001). A simulation-optimization framework for Research and Development Pipeline management. *AIChE J.*, **47**, 2226-2242.
- Thwaites R. and R.J. Townsend (1998). Pharmacoeconomics in the new millenium. *Pharmacoeconomics*, **13**, 175-180.
- van der Vorst, J.G.A.J, A.J.M. Beulens and P. van Beek (2000). Modelling and simulating multi-echelon food systems. *EJOR*, **122**, 354-366.
- Voudouris, V.T. and I.E. Grossmann (1992). Mixed-integer linear-programming reformulations for batch process design with discrete equipment sizes. *Ind. Eng. Chem. Res.*, **31**, 1315-1325.
- Yeh, N.C. and G.V. Reklaitis (1987). Synthesis and sizing of batch/semicontinuous processes: single product plants. *Comput. Chem. Eng.*, **11**, 639-654.