Why Don’t We Tackle The Pharmaceuticals Industry ?

Industry Overview

The pharmaceutical’s industry has gone through some significant consolidation over the past five years and competition has become increasingly international in nature. The average prescription drug costs $200 million to develop and requires an average of 14 years of R&D plus Government approval time. Time to market is becoming an increasingly important competitive factor and the Government approval delay averages between 6 months to 3 years.

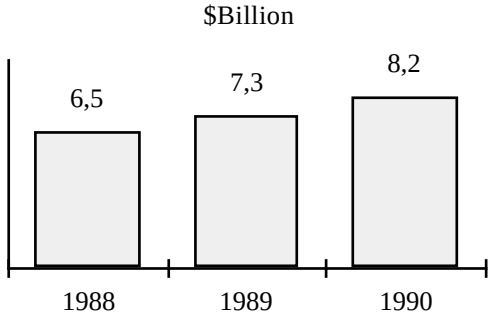


Figure 1. Escalating R&D Investments.

Escalating R&D investments are being accompanied by growing risk. A company can expect to lose 35% of market share, the first year a product goes off patent, 50% the second year. About 17 generics enter the market within a couple of years after a patent expires (patents are granted for 10 years).

The number of patients participating in anti-microbial clinical trial in 1979, averaged 1,500. In 1990 a similar trial involved 10,000 patients. The figure doubles every four to six years. Also, diseases remaining unconquered today such as AIDS, cancer diabetes, arthritis, and Alzheimer’s, are too slow, degenerative diseases, which require many years of clinical trials to prove efficacy of therapies.

Clinical lab tests double every two years. Case report forms per NDA double every one and half years. Eli Lilly’s latest NDA for Pindac (pinacidil) contained 470,000 pages. ICI Pharmaceuticals Group submits four million pages of regulatory filings a year. Glaxo’s approval paper work for Zantac (ranitidine) in 100 nations weighed about 16 tons, which is about 128 acres of forest. The average size of a NDA dossier is of 200,000 to 300,000 pages for a new drug and of 50,000 pages for a generic.

The average pharmaceutical company increased the size of its pre-clinical and clinical databases from two gigabytes to about sixty gigabytes over the past five years; that figure is doubling every two years.

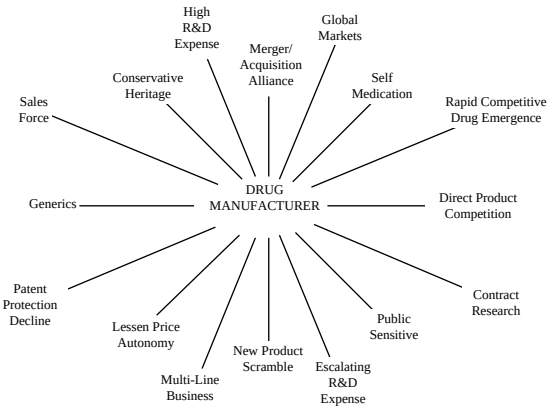


Figure 2. Pharmaceutical Industry Dynamics.

Case Study

Europe Pharmaceuticals Company (EPC) is a $1,2 billion Ireland-based manufacturer of ethical drugs. The industry as a whole has sales in Europe of almost $30 billion, with approximately 125 companies that undertake primary drug development.

EPC currently depends on the antibiotic's market place for 70% of its revenues; $500 million is derived from sales of Elecyclyoneziatran, a general antibiotic that has been on the market for more than five years. Although profits on this drug exceed $200 million, final sales are now flat and other companies are due to release new competitive products within the next two years.

EPC has recognized the internationalization of its industry and although less than 25% of its sales are made outside Europe, it is determined to compete more effectively both in the US and in the Third World.

Pricing strategies have been progressive in the Third World, and sales have increased from $100 million to $150 million. Sales of existing products will not achieve a higher level without changing price policies. This may lead to gray market products appearing in Europe or in the US.

EPC has acquired France First SA, a Paris-based manufacturer of home testing products for such conditions as high blood pressure and pregnancy. In addition France First is nearing completion of a medication designed to address abnormal heart rhythms. Communications with its French subsidiary have been hampered by incompatible office automation environments and language problems.

Market Opportunities

Most pharmaceutical companies have invested significant resources in attempting to employ image systems technology in a number of different ways[1]. Half a dozen pharmaceutical companies have started test programs to submit their new drug applications to the Government approval department using optical disk technology.

Companies have bought image systems either for their medical library areas, records management functions, or for the purposes of automating the drug development functions during clinical trials. Some have even gone so far as to give an imaging system to the Government approval department in the hope that this would improve the time lags in the drug approval process[2]. The format and layout of submissions is not imposed; in some countries, the content is not imposed either. Existing guidelines and reglementations only defines the topics which should be addressed. This gives considerable latitude for system design; each company desires a custom system (which they view as a discriminator).

Pharmaceutical companies hence present the perfect client profile: they are big, profitable and healthy.

---

1 The US pharmaceuticals industry new investment budget was rated at $400 million for 1990.

2 Current CANDA experiences show that the average approval time of 28 months for the paper applications was cut down to 17 month with the computer assisted applications.

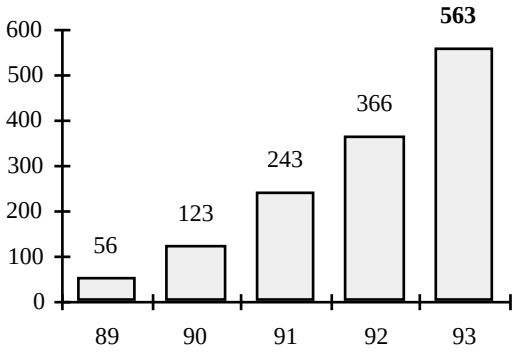


Figure 3. Pharmaceutical image systems sales projection.

Investigational new ethical drug applications (IND)	1,337
New ethical drug applications (NDA)	126
New ethical drug applications approved	87
New biological drug applications approved	113
New generic drug applications approved	265

Table 1. US Pharmaceuticals Industry Activity Volumes for 1990.

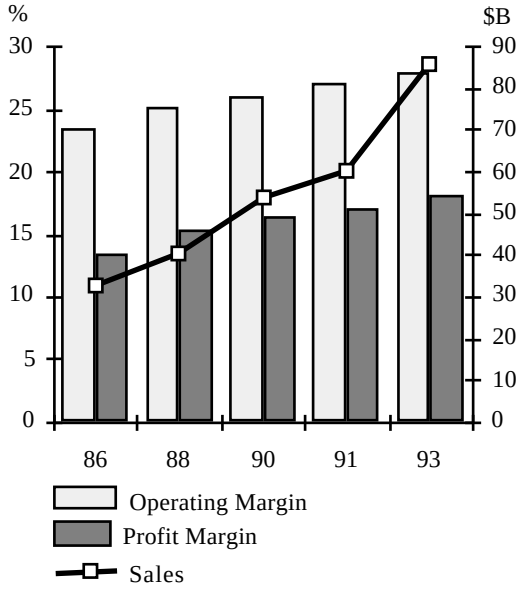


Figure 4. US Pharmaceutical Industry Sales.

Typical Grounds for Image Systems

The average research scientist spends almost two hours per day looking for the appropriate documentation. Since a scientist and his/her associated costs total approximately $200,000 per year, the cost for low productivity is of $40,000 per scientist. For a company of 600 scientists, the overall cost for low productivity would be of $24 million: an image system can reduce this time to less than 10 mn per day per scientist. The cost of an image system can be paid back in short order with increased productivity.

Typical Office Automation Environment

The corporate side of business (e.g, accounting, legal) is often an IBM environment, while the clinical side (e.g. drug development) is primarily Digital. In addition Unix, PC or Macintosh workstations are used by the scientists, Hewlett-Packard systems are used for manufacturing and Wang systems for administrative functions.

The corporate MIS function has an annual budget in the range of $100 million of which the clinical budget for MIS represents over 50% (the manufacturing budget represents approximately 10%).

Competitive Outlook

Many companies have developed a sales strategy for electronic CANDAs (Computer-Assisted New Drug Applications). Many of which are system integrators as shown in Table 2.

CANDA offerings follow one of two approaches: a ‘day zero’ or a ‘day n’ approach.

The ‘day zero’ approach manages the entire life cycle of a NDA dossier, from the first day of discovery research, until the day the regulatory affairs department submits it to the Government(s) for approval. A data repository is opened and made available throughout the contributing departments of the enterprise. The NDA dossier is

created on-line and in real-time. In theory, this is the good global response for technical information management. However, this approach is very constraigning since it dictates the customer's computing environment, imposes his working procedures and does not integrate his existing specialized applications nor his backlog data and documents. Further, the 'day zero' approach supposes that all partners have the same tools. Finally, this approach does not consider all the vital external information sources (e.g. specialized revues and bulletins) and databases (e.g. MEDLINE) which are frequently used by R&D staff.

I believe that 'day zero' CANDAs can only be supplied by the major computer manufacturers, of which Digital Equipment Corp.[1] is, probably, the only one which can win a significant market share.

The 'day n' approach appears as the most practical solution. The NDA dossier is compiled once the project is ready for submission to Regulatory Departments. While 'day zero' CANDA systems can be considered as data repository management systems which manipulate diverse data types including compound documents, 'day n' CANDA systems are real and specialized imaging systems. Real systems because they offer the basic functionality of capturing, storing and retrieving mechanical and digital documents; specialized systems because they offer assembly and printing functions which allow to cut and paste all of parts of documents, and compile them into a NDA dossier.

I believe that Integrated Automation Corp. has a real and significant market opportunity: we have the base technology to build such systems, the approach that best suits customers who look for *partners* and *custom* systems and the opportunity, through RPR, to mature and stabilise a standard offering. The only major handicap might be that of sufficient credibility in a market where we still do not have any reference installations; but this handicap could be turned into an advantage by developping partnerships with companies, such as Interleaf, who have a complementary know-hows aznd a large installed base in the pharmaceutical industry.

AGS Information Services	Grumman Data Systems
Andersen Consulting	Hewlett-Packard Co.
AT Kearny	IBM[2] Corp.
Cap Gemini	Laser Recording Systems
CIMI Corp.	LaserData, Inc.
Communications Network Architects Inc.	Plexus Software, Inc.[3]
Computer Task Group	Science Applications International Group
Coopers & Lybrand	Summit Software Corp.
Digital Equipment Corp.	TAB Products Co.
FileNet Corp.	ViewStar Corp.
	Wang Laboratories, Inc.

Table 2. The Competition Tenors.

Outline of a CANDA Offering

Base Functionality

The functional requirements for a CANDA system can be classified in three categories:

◊ Standard imaging functions;
◊ Standard NDA assembly & printing functions,
◊ Custom document and data interchange functions.

Standard imaging functions

These are the vital functions that are always present:

◊ Document and data capture;
◊ Optical storage
◊ Information retrieval and display

1 Digital Equipment Corp. packages a 'day zero' CANDA solution named *Computer-Integrated New-Drug Applications* (CINDA).

2 IBM has a setup a special 'SolutionHouse' infrastructure called IBM Pharmaceutical Industries Group

3 Plexus Software, Inc. is not an actor as such in this market but is the prime supplier of imaging technology to Hewlett Packard Co. et Grumman Data Systems.

Document and data capture must allow to capture mechanical and electronic documents. It is important to capture facsimile transmissions since they are intensively used and to be able to manipulate foreign files using both a native format and an interchange format (e.g. PostScript). It is equally important to provide automatic data recognition and processing through (e.g. optical mark sense, forms recognition, forms segmentation, and flexible data extraction, computer-aided key-entry and editing) since intensive indexing and bibliographic data is to be input into existing databases.

Optical storage must also encompass backup and security procedures including the possibility to share storage farms across several sites.

Information retrieval and display must offer standard imaging features and should include an annotation facility (revision control is not necessary). It should also be possible to automatically load the native format of foreign files when the correct environment is present.

Standard NDA assembly & printing functions

The NDA assembly and printing functions cover areas which require special application software. These are generic and applicable to all customers; they include:

◊ Document creation,
◊ Dossier definition
◊ Dossier assembly
◊ Dossier printing

Document creation is a word processing utility in which raster images can be pasted in order to create new documents which will be stored[1].

Dossier definition functions must allow to create NDA dossier outline templates indicating the topics which should be covered and their order. Such templates only deal with the logical structure of the dossier; there are no layout rules.

Dossier assembly functions allow to manage the assembly of an NDA dossier according to a given template. The assembly consists in associating document references to each logical level of the template. The assembly must be able to support cross-referencing (though this could be accomodated by annotations). Workflow support is probably a necessity. Dossier printing involves the construction of a hard copy document with continuous page and paragraph numbering, date stamping and standard headers and footers.

Custom document and data interchange functions

Document and data interchange functions allow the seemless integration of the CANDA system in the existing automation environment. They are necessarily custom and allow to receive and send both documents and data.

Base Technology

The above listed functionality relies on the following base technologies:

◊ High speed A3/A4 scanning;
◊ Group III facsimile capture;
◊ Postscript to Group IV conversion;
◊ Optical character recognition (Omnifont/handprint),
◊ Mark sense detection,
◊ Form and document segmentation,
◊ Multi-layered raster/text imaging,
◊ Workflow management,
◊ High speed A3/A4 printing.
… (not exhaustive)

SGML and DSSSL[2] are contenders to solve the assembly and printing problems. Their standard features allow in particular:

a) documents to be tagged to indicate their position in the dossier,
b) automatic renumbering of the dossier outline (volumes, parts, section, chapters, etc.)
c) definition of headers and footers,

---

1 Interleaf is a perfect candidate. Tests conducted at MC2 have shown that it is possible to automate an interface with Interleaf to offer such a cut & paste solution for raster images.

2 DSSSL is not yet a standard but can be considered as a *final draft*.

d) automatic page numbering,
e) insertion of cross-references
f) insertion of (last minute) free-text
g) insertion of (encapsulated) foreign files.

We currently have no experience of SGML or DSSSL. This is the main issue (along with the sales strategy).

Partie I	**Résumé du dossier**
Partie I A	Renseignements administratifs
Partie I B	Résumé des caractéristiques du produit
Partie I C	Rapports d'experts
Partie II	**Documentation clinique, pharmaceutique et biologique**
Partie II A	Composition
Partie II B	Méthode de préparation
Partie II C	Contrôle des matières premières
Partie II D	Contrôle des produits intermédiares
Partie II E	Contrôle du produit fini
Partie II F	Stabilité
Partie II G	Autres informations
Partie III	**Documentation toxicologique et pharmacologique**
Partie III A	Toxicité par administration unique
Partie III B	Toxicité par administration réitérée
Partie III C	Etudes de reproduction
Partie III D	Potentiel mutagène
Partie III E	Potentiel carcinogène/oncogène
Partie III F	Pharmacodynamie
Partie III G	Pharmacocinétique
Partie III H	Tolérance locale
Partie III Q	Autres informations
Partie IV	**Documentation clinique**
Partie IV A	Pharmacologie humaine
Partie IV B	Documentation clinique
Partie IV Q	Autres informations
Partie V	**Renseignements particuliers**
Partie V A	Présentation
Partie V B	Echantillons
Partie V C	Autorisation(s) de fabrication
Partie V D	Autorisation(s) de mise sur le marché
Partie V E	Déclaration d'expertise ou dispense
Partie V F	Fiche destinée aux centres régionaux de pharmacovigilance et aux centres antipoisons

Figure 5. Top level outline of a NDA dossier for the European Economic Community.

Conclusion

I firmly believe we can adopt a Docuvision III-like approach and deliver Docuvision CANDA. The main technical issue is the SGML/DSSSL integration on which I would like to work. Further, 15 man/days are sufficient to deliver a draft sales kit composed of:

◊ an offering description
◊ a template study and solution proposal,
◊ a module and component catalog (à la DV III).

In parallel Sales & Marketing should concentrate on:

◊ the decision process (steps, actors, time scale and purchase influencing factors),
◊ the purchase motivations (tied to the industry and the offering),
◊ the marketing strategy (technical/symbolic value, level of innovation/development, price magnitude)

As might appear in this document, I am highly motivated by this topic and have been since the end of last year when I finally and clearly understood the problem RPR is tackling. This was early November 1991; date at which RPR told us that, if they did the system, they would do it with us (we've heard a lot of that recently!). Since then the account has practicaly been abandonned and I have seen no prospection (at least on the Old Continent) for new pharmaceutical accounts. Hence the problem is not that the market is already saturated, nor that we don't have sufficient credibility, nor that we don't have an offering, etc.…it simply is that we have no marketing strategy and have not identified the sales effort we are willing to place on this market opportunity. I sincerely hope that this small contribution will help make things change.

Abstract

The pharmaceutical industry appears as a possible marketing niche both for IA and MC2. Pierre Mathieu and Dominique Maillet have started an investigation for MC2. Having participated in the RPR proposal, I am keen on the idea and believe that we currently have a sound technical baseline solution for the automated assembly and printing of NDA dossiers. The purpose of this document is to highlight my comprehension of the new drugs approval business case, to briefly described the functionality and technology foundations I suggest as guidelines for an IA/MC2 solution, and to show where we stand.

1. Pharmaceutical business case

Structure Logique d'un Dossier d'Autorisation de Mise sur le Marché

Partie I : Résumé du dossier

Partie I A: Renseignements administratifs
Partie I B: Résumé des caractéristiques du produit
Partie I C: Rapports d'experts

Partie II : Documentation chimique, pharmaceutique, biologique et biotechnologique

Partie II A: Composition
Partie II B: Méthode de préparation
Partie II C: Contrôle des matières premières
Partie II D: Contrôle des produits intermédiaires
Partie II E: Contrôle du produit fini
Partie II F: Stabilité
Partie II G: Autres informations

Partie III : Documentation toxicologique et pharmacologique

Partie III A: Toxicité par administration unique
Partie III B: Toxicité par administration réitérée
Partie III C: Etudes de reproduction
Partie III D: Potentiel mutagène
Partie III E: Potentiel carinogène/oncogène
Partie III F: Pharmacodynamie
Partie III G: Pharmacocinétique
Partie III H: Tolérance locale
Partie III Q: Autres informations

Partie IV : Documentation clinique

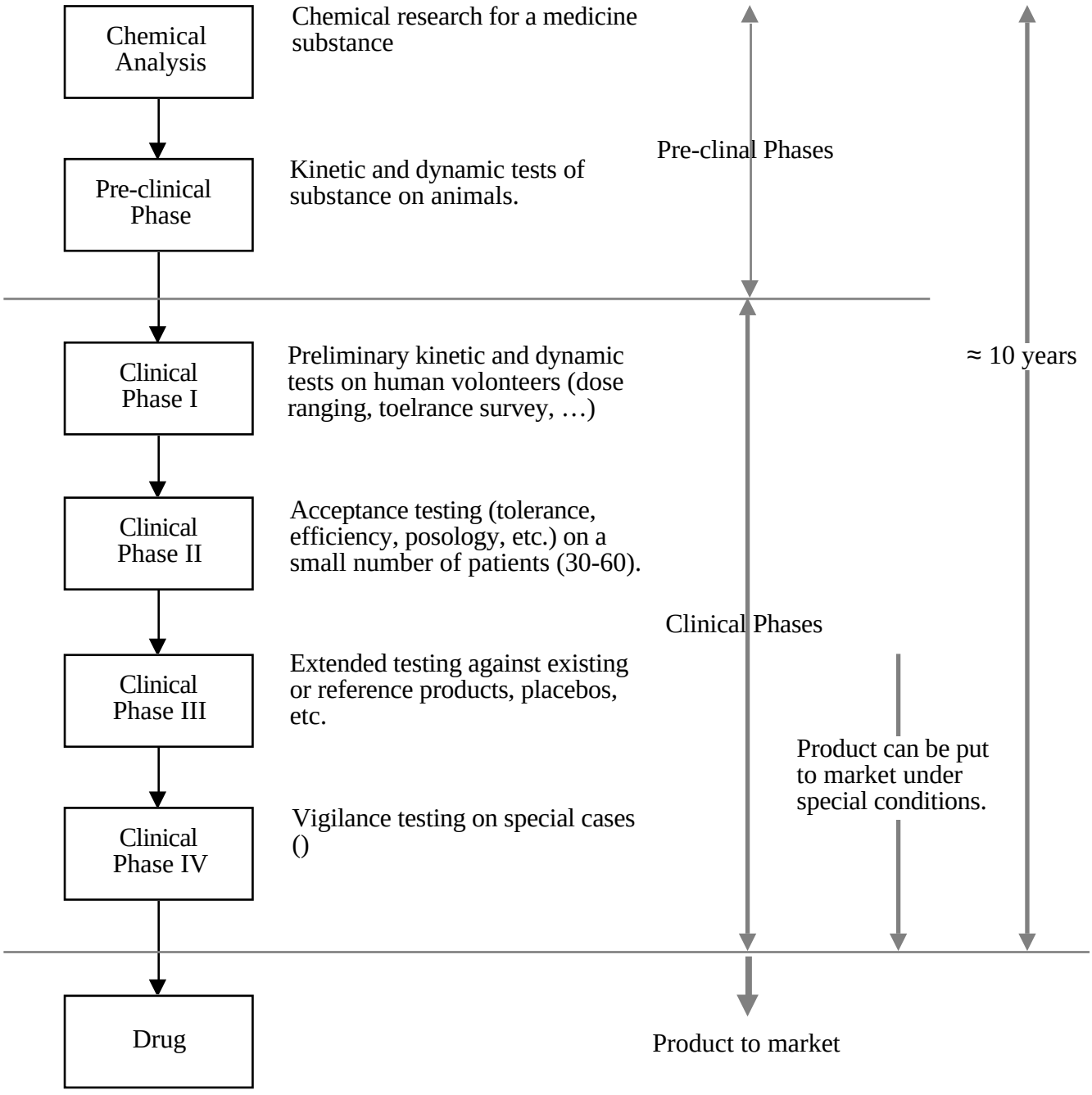
Partie IV A: Pharmacologie humaine
Partie IV B: Documentation clinique
Partie IV C: Autres informations

Partie V : Renseignements particuliers

Partie V A: Présentation
Partie V B: Echantillons
Partie V C: Autorisation(s) de fabrication
Partie V D: Autorisation(s) de mise sur le marché
Partie V E: Déclaration d'expertise ou dispense
Partie V F: Fiche destinée aux centres régionaux pharmacovigilance et aux centres antipoisons

1.1. Drug R&D Cycle

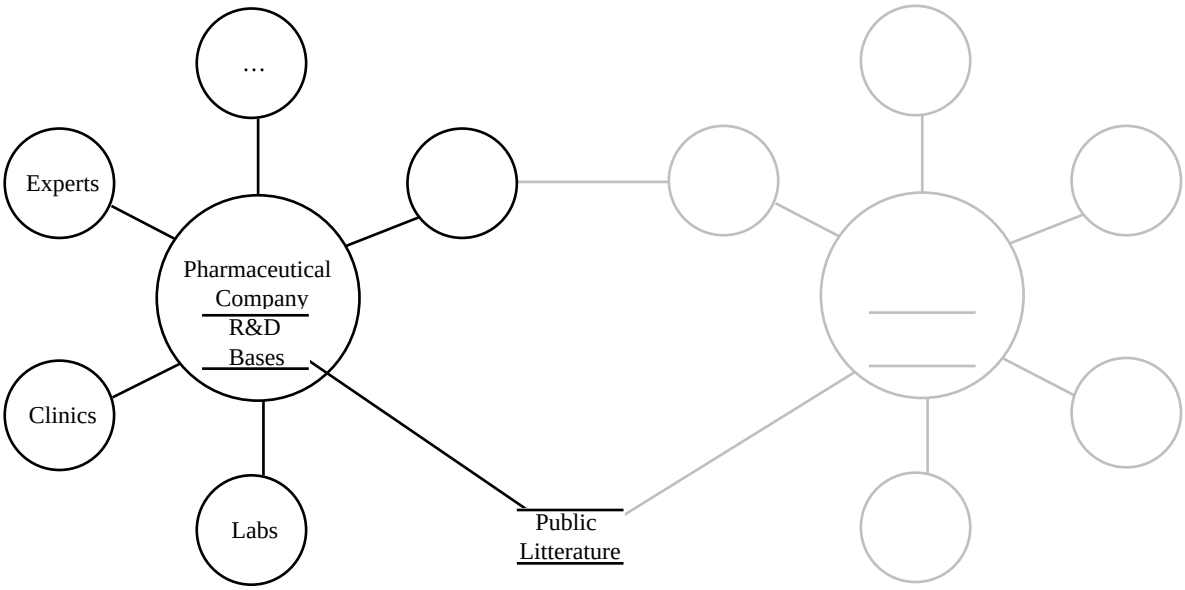
Before a pharmaceutical product comes to market it goes through a long R&D cycle detailed below. The result of the R&D cycle is a NDA dossier which, if it is approved by the Regulatory Offices to which it is submitted (FDA/USA, DPHM/France, CSP/EEC, etc.), brings the drug to the marketplace along with a 10 year copyright.



1.2. Key points in the marketing of drugs

A pharmaceutical company's success and profitability hinges on its ability to get new drugs approved. This approval process demands the genration of huge volumes of information and documents which must be gathered, managed and maintained.

The automation of NDA dossiers is therefore a critical issue. Many solution makers have identified the market opportunity; all offerings I have seen so far, are far from matching the RPR context..



1.3. NDA dossier assembly by Regulatory Departements

2. Building an IA/MC2 proposal

2.1. Main Functional Issues

The functional requirements and issues for the automated assembly of NDA dossiers can be classified as follows:

- ◊ standard imaging system functions;
- ◊ record management functions;
- ◊ NDA dossier assembly and printing functions;
- ◊ NDA dossier workflow functions.

2.1.1. Standard Imaging System Functions

The standard imaging functions are those we currently offer in our systems: digital and non-digital document capture, storage, retrieval, and distribution.

Document Capture

Document capture must allow to convert paper and eletronic files as well as electronic transmissions (eg. facsimiles). There are two major areas of concern: the (semi)automated capture of bibliographic and indexing data and the automated environment.

Lots of bibliographic and indexing information needs to be associated with the captured documents. For backlog documents, this information is mostly available in existing databases. For frontfile documents, a means of automating the capture of this information is needed. RPR uses BASIS+ to implement a full-text retrieval system, which covers all technical fields, associated with many index tables. This approach seems to be specific to RPR (explained by their concern to quickly achieve, after the political merger of Rhône-Poulenc and Rorer, a common information base). Segmentation and OCR technologies are required. The use of packages such as BASIS+ should be put under detailed scrutiny[1].

The RPR automation systems is are a typical example of a heterogeneous environment: IBM and DEC mainframes, Unix, PC and Macintosh workstations, SNA and Ethernet networks, etc., resulting in a wide range of different products. For document creation purposes, we counted approximately 15 different packages. Hence a critical issue for the capture and exchange of electronic files: we cannot, for n formats, build n(n-1) translators! The solution we retained was based on three formats: a raster format for imaging, the specific native format of the originating system for editing and updating, and an interchange format (PostScript). The RPR environment may be an isolated case; the problem remains though, since the captured documents originate from many different companies.

Document Storage and Retrieval

This area was not sufficiently investigated with RPR by lack of quantitative information. A major concern here is the assembly and printing of the NDA dossiers: 600 000 pages to be assembled and printed in a day! This seemed critical at first; however further investigation showed that prestaging techniques can be applied based on the knowledge of the logical structure of the dossier and of the level of completion of its various parts.

Distribution

The working procedures of Regulatory staff was only partly with RPR. We know what needs to be done, but we do not know how users do their job. Both clerical and knowledged workers are involved in the assembly of a NDA dossier. Further, specialised staff is needed: a toxicology specialist may be asked to make a synthesis of several patient reports or a clinical expert may be needed to verify an apparently wrong dosage. This is why a distribution scheme with workflow support is most probably a vital necessity.

---

1 MC2 has acquired in depth knowledge of BASIS+ through the ERUDIT project currently in acceptance phase.

2.1.2. Record management functions

The record management functions enable the NDA dossier assembly system:

- ◊ to automate the extraction or constitution of the bibliographic data of stored documents;
- ◊ to be coupled with existing bibliographic databases which are either company specific or commonly used by the pharmaceutical industry (eg. Medline).

Because of the wide range of investigation areas (e.g. toxicology, pharmacology, drug disposition, pharmacokinetics, bioavailability, microbiology, regulation, etc.) involved in the definition of a drug, lots of information is needed for each document in order to determine its contents. The resulting problem is a) how to gather this information and, b) how to allow its manipulation in relation with an imaging system. RPR is in the process of developing a generic system (WISDOM) under BASIS+; this is to our advantage since it greatly simplifies the interface. But the RPR approach is, I believe, an isolated case due to the merger. I believe the general case is the necessity for the interconnection with many different specialised systems and repositories.

2.1.3. NDA dossier assembly and printing functions

The automated assembly and printing of NDA dossiers calls for special functions in the following areas :

- ◊ assembly work flow management,
- ◊ dossier creation,
- ◊ document creation.

NDA dossier assembly involves concurrent team work based (as I understand) on technical fields. Workflow support has already been adressed above.

NDA dossiers have a logical layout which is entirely dictated by the

2.1.4. Document creation

As part of the NDA

2.2. Technology foundations

The processing of NDA dossiers bears many similarities with the processing of patent applications as designed in the ELFOS study. The major ELFOS concepts can be applied: subject group taxonomy, dossier/document structure, mailboxes, distribution monitoring, etc. I believe ELFOS is the appropriate base and would be

- ◊ the ELFOS DocEntry functions are all relevant,

- ◊ the subject group taxonomy can implement the NDA dossier structure

- ◊ extensive bibliographic data capture is required,

◊ distribution and workflow◊ In both casesSlight modifications of the ELFOS entity/relation and data flow diagrams would suffice to make me confident in presenting them to a pharmaceutical company.

The budgetary estimate for the RPR/DMS proposal was in the range of $2M. I do not really know how to qualify this variable: from a technical standpoint, this figure seems too low, whereas it appeared to be the upper psychological barrier for RPR.

3. Conclusion

I hope this document convinces you that the technical solution, or at least the base-line solution, is identified. We have a major distinguishing factor with the major competitors I know of through the product-like function offerings while allowing to fully integrate into a possibly highly complex automated environment. Yet this is not sufficient to have sufficient credibility with pharmaceutical prospects; the various interactions we have had so far with RPR have essentially been on technical grounds not giving us the necessary insight on the institutional, cultural and economic environment of this industry. We know little to nothing on the pharmaceutical market network. We knowas little on the existing solutions and on the competition. And, most important, we have no measurement of the market size.

How then can we develop a strategy ? determine the mix ? identify the segments ? MC2 has urged its sales staff to start prospecting and has named Perre Mathieu as the market manager and Dominique Maillet as the market consultant. Even though it is obvious, I wish to emphasize the key aspects on which we should concentrate first; these are:

◊ the decision process (the steps, actors, time scale and the influencing factors of a purchase);

◊ the purchase motivations (tied to the industry and to the offering);

◊ the marketing strategy (the technical and symbolic characteristics of the product, the short to long term approach, the level of innovation and development and the price magnitude)

I am not seeking theory. I believe that we can take advantage of the summer to further investigate the market in parallel with the constitution of a prospect list.. The market knowledge we will acquire associated with the existing technical base-line solution will lead to the construction of the set of tools needed for prospection and sales:

◊ an offering description;
◊ a template study and solution proposal,
◊ a price catalog.