

Computers in the New Drug Application Process

Joel F. Studebaker*

IBM Pharmaceutical Industry Center, 300 Executive Drive, West Orange, New Jersey 07052

Received February 12, 1992

Before a company can introduce a new pharmaceutical in the United States, it must receive approval of a New Drug Application (NDA) from the Food and Drug Administration. As a document that may consist of over 100 000 pages, the NDA presents the company's evidence that the new pharmaceutical is safe and effective for treating a particular medical condition. This paper describes four major issues in the NDA process: the role of computers in the FDA review, standards for compound documents and data, substitution of computer-assisted new drug applications (CANDAs) and electronic versions of other documents for paper versions, and management of the heterogeneous collection of material on which an NDA is based.

INTRODUCTION

The process of bringing a new pharmaceutical from the research stage to market requires many years and several phases (Figure 1). After characterization and development of a preparative method, a promising new compound proceeds to preclinical testing: toxicology, bioavailability, and pharmacokinetics. The data generated in this phase provide the basis for the Investigational New Drug Application (IND), which the FDA must review before clinical trials of the compound begin. The three phases of clinical trials involve testing, in humans, for safety and for effectiveness in treating a target medical condition. If the compound still appears to be promising after phase III, the sponsoring company assembles the data generated in the clinical trials and supporting material into the NDA, which the FDA must approve before the compound may appear on the market. After approval of the NDA, the company's monitoring of the drug's effects on patients in the general population constitutes phase IV of the process. An NDA is not a static document. Even after the agency has approved it, the sponsoring company may submit additional information for FDA review when seeking approval for a broader range of therapeutic indications or in other circumstances.

Similar requirements apply to biologics like vaccines, where the Product Licensing Application (PLA) is analogous to the NDA, and to medical devices like catheters, where the Pre-Manufacturing Approval is analogous to the NDA. In addition, companies planning to manufacture biologics must receive approval of an Establishment License Application (ELA) for each manufacturing site. Most pharmaceuticals, biologics, and medical devices are on the market in more than one country, and the regulatory agencies of other nations must approve new ones before they enter overseas markets. Those agencies often have requirements for form and content that are considerably different from those the FDA specifies for documents submitted to it.

COMPUTERS IN NDA REVIEW

Submitting statistical data from clinical trials to the FDA on magnetic tape has been a common feature of the NDA process for many years. Recent efforts in CANDAs have sought a broader use of computers, as noted in the 1988 statement by the FDA commissioner that the agency "believes

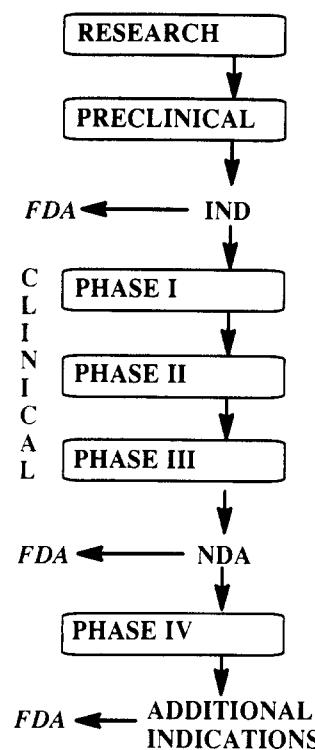


Figure 1. Stages in the development of a new pharmaceutical. The entire process may take up to 15 years, and a very small fraction of the compounds characterized in the discovery phase arrive on the market.

the increased use of computers may improve the efficiency of the drug application review process".¹ A survey report² at the joint meeting of the FDA and the Pharmaceutical Manufacturers Association (PMA) in July 1990 noted that 15 companies had submitted 44 CANDAs of various types to the agency by that date. By November of 1991, the number of CANDAs had reached 65. In each case, however, it was still necessary for the sponsoring company to submit the full paper copy of the NDA, which was the legal copy. Furthermore, the FDA has stated that submitting a CANDA does not affect the place of an NDA in the review queue.

Of the 65 CANDAs, approximately half have made various facilities available to FDA reviewers through Research Data Corporation (RDC).³ The facilities include access to clinical databases on computer systems at RDC or at the sponsoring company, electronic mail communication with the sponsoring company, and SAS datasets for statistical analysis. The system also provides electronic presentation of case report forms

* Address correspondence to his present address: Unified Systems Solutions, 49 Old Bloomfield Ave., Mountain Lakes, NJ 07046.

SAF Pharmaceuticals Inc.		TORIOMENE PELLETS(SOE 33G) VITAL SIGNS & PHYSICAL EXAMINATION						
Visit 3	Prog A	Protocol 462	Investigator Name/Number Laser / 222	Patient I.D. SOC	Patient No. 61	Date of Visit 8/10/88		
VITAL SIGNS								
Hours since last Dose: 1.2		Weight (lbs) 202	Respiratory Rate (min.) 49					
(48-60)		(48-70)	(48-60)		(48-60)			
SUPINE (10 Min.)		Blood Pressure (mm/Hg)	Apical Rate	STANDING (3 min.)		Blood Pressure (mm/Hg)	Apical Rate (mm/Hr)	
READING No 1	4.30/ 60	80	70	READING No 2	4.30/ 60	83	78	
READING No 3	4.25/ 60	70	70				26	
IF LEFT ARM USED EXPLAIN:								
PHYSICAL EXAMINATION								
Normal		Abnormal	Not Examined	Explanation of Abnormalities				
<input type="checkbox"/>		<input checked="" type="checkbox"/>	<input type="checkbox"/>	Rash on neck				
<input type="checkbox"/>		<input type="checkbox"/>	<input checked="" type="checkbox"/>					
Fund/K W Grade -			<input checked="" type="checkbox"/>	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input checked="" type="checkbox"/> 3	<input type="checkbox"/> 4
Heart		<input type="checkbox"/>	<input type="checkbox"/>					
Lungs		<input checked="" type="checkbox"/>	<input type="checkbox"/>					
Abdomen		<input checked="" type="checkbox"/>	<input type="checkbox"/>					
Extremities		<input checked="" type="checkbox"/>	<input type="checkbox"/>					
Skin		<input type="checkbox"/>	<input checked="" type="checkbox"/>	Small freckles				
Neurological		<input checked="" type="checkbox"/>	<input type="checkbox"/>					
Clinical Labs		Date of Sample: 7/17/88		Fasting?	<input type="checkbox"/> No		<input type="checkbox"/> Yes	
Repeat Labs		<input type="checkbox"/> No		<input checked="" type="checkbox"/> Yes				
I HAVE REVIEWED ALL THE CASE RECORD FORMS FOR THIS VISIT Signature of Principal Investigator: <i>Seth Laser</i> Date: <i>3/10/88</i>								

Figure 2. Sample page from a Case Report Form.

(CRFs), which are the multipage forms that physicians and others conducting a clinical trial fill out to report the results of physical examinations and laboratory tests on patients in the trial (Figure 2). Originally developed on midrange computers, the RDC functions are now available on IBM or compatible PCs or on Macintoshes.

According to subjective estimates of the reviewers, this approach has accomplished the important goal of reducing the time required for review. With the database, the SAS datasets, and the CRFs readily available, reviewers have been able to develop answers to questions that they would previously have addressed to the company sponsoring the NDA. When reviewers raise questions about the data or analysis and address them to the sponsor, several months may elapse before an answer arrives. During that time, the reviewer typically shifts attention to other regulatory submissions. Picking up the review again, once the answer is in hand, takes some time. The benefit of the CANDA approach in this regard was evident in the first one that used the RDC approach. Abbott Laboratories submitted it in 1985 for two cardiovascular drugs, and the FDA director who carried out the tertiary review of the CANDAs noted that he was able to answer a series of questions about the submission through the data and facilities on the computer. No questions to the sponsor were necessary. In the more recent review of Ciba-Geigy's Voltaren, company and FDA officials estimated that the CANDA facilities had shortened the review time by 6 months to 2 years.⁴

When it has been necessary for a reviewer to direct a question to the sponsor, the availability of electronic mail, via modems, between reviewers and submitting companies appears to have reduced the lag between the time a reviewer asked a question of a company and the time he or she received an answer.

The other major category of CANDAs is the optical NDA (ONDA), in which sponsoring companies submit the NDA on Write Once, Read Many (WORM) optical disks. ICI Pharmaceuticals submitted the first one in 1988, and Laser Recording Systems has been the primary developer of these submissions. In an ONDA, the entire document is available in (bit map) images, supported by a database that allows query and retrieval by keyword, NDA section, or section page number. Optical disks, with high capacity at low cost and relatively long access times, are well suited to storing the large files of the bit maps themselves.⁵ For efficient searching of the information describing an image, however, the database itself is on a magnetic hard disk, which has a shorter access time. Reviewers can annotate particular pages with comments for later incorporation into their reviews, and the system automatically links those comments to the pages they refer to. Reviewers have found the ability to retrieve images of CRFs covering patients who fit a set of criteria particularly valuable. Images of pages of descriptive text are not as useful, however, because reviewers would prefer to read them on paper rather than on a computer screen.⁶ It is also possible for the sponsoring company to incorporate data in other formats, such as text of database tables, and to include such data in the ONDA database. Through a Microsoft Windows interface, a reviewer may display more than one page at a time, analogous to having two paper volumes open at the same time, but clearly more convenient. Multiple reviewers can study the same ONDA through multiple WORM disk copies or through an optical disk drive on a network, protected by ONDA security in either case.

In Canada, approximately 15 major submissions to date have fitted the definition of Electronic New Drug Submissions, which incorporate a database management system for clinical data and a document management system for the pages of the submission. A joint task force of the Canadian Drug Directorate and the Pharmaceutical Manufacturers Association of Canada (PMAC) has also suggested that sponsoring companies incorporate hypertext links that would, for example, connect the discussion of recommended dose for a particular patient subgroup with the clinical trial data for that subgroup.⁷ Another noteworthy feature of the Canadian review process is that some reviewers of INDs and New Drug Submissions have used the TOPKAT software to estimate the toxicity of new chemical entities, their metabolites, and side products of manufacturing when limited data has been available. From quantitative structure-activity relationships derived from its underlying database, this software provides estimates for carcinogenicity, Ames test mutagenicity, and other properties.

The Bundesgesundheitamt (BGA), the German drug regulatory agency, has been the most active agency in Europe in the CANDA area. Working with a group of pharmaceutical companies that includes Bayer, Boehringer Ingelheim, Hoechst, and E. Merck, the agency has developed a standard for ONDAs called Drug Application Methodology with Optical Storage (DAMOS).⁸ This standard specifies the formats for both the image files and the database describing them.

Features of CANDAs that FDA reviewers regard as essential or extremely useful⁹⁻¹¹ include cut-and-paste of text and graphics into their own word-processing documents, full-text search, data manipulation, and ad hoc database queries. The electronic mail link to the sponsoring companies function has not reached its full potential because many reviewers prefer telephone contact with sponsors. Some CANDA elements are more useful to one type of reviewer than another. Spreadsheets are a good example; clinicians find them useful,

(A)

```

OPEN THE MAIL          000
In-basket for: Robert Hirsh
Press the PF key for the mail item you want, or type AUTH.
or a special mail command and press ENTER <-->
---SENDER---      ----TO----   TYPE    DUE DATE DOCUMENT NO.
PF1  Joel F. Studebaker Hirsh, R.J.      Note    03/26/92 11:04:37
Subject: Structure of o-xylene
PF2  Robert Hirsh      Robert Hirsh      Note    03/26/92 10:45
Subject: Marketing Guide Update Review Profs Note

```

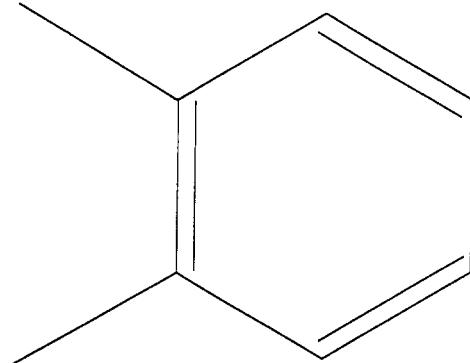
PF9 Help PF10 Next Screen PF11 Previous Screen Screen 1 of 37 PF12 Return

(B)

VIEW THE NOTE E01
From: JFSTUDEB--NYCVMIC2 Date and time 03/26/92 11:04:37
To: RMIRSH --NYCVMIC2 Hirsh, R.J.
From: Joel Studebaker, PhD ..Pharm Center..300 Executive Dr.
W. Orange, NJ 07052..Phone 201-669-6517, 8 223 6517
FAX 223 5695
Subject: Structure of o-xylene
Bob,
The structure of o-xylene is in file 920117C0086.
Regards, Joel

E N D O F N O T E

PF1 Alternate PFs PF2 File NOTE PF3 Keep PF4 Erase PF5 Forward Note
PF6 Reply PF7 Resend PF8 Print PF9 Help PF10 Next PF11 Previous PF12 Return

(C)

PF1 Turn Off PF3 Print at 100% PF9 Help PF12 Return

Figure 3. Viewing graphics transmitted by electronic mail. The standard E-mail IN Basket (A) notifies a user that a note (B) has arrived. Placing the cursor under the number of the graphics object in the note brings up the object, here a chemical structure (C).

but statistical reviewers prefer SAS. Regarding potential new CANDAs, chemistry reviewers¹² noted that on-line availability of drug masterfiles from suppliers of excipients or packaging materials used with a particular drug would facilitate their review. The review of additional claims, submitted after the initial NDA has been approved, would be more efficient if the reviewer of the additional claims could refer to a computer-based version of the NDA.

Because the FDA review process has some parallels to refereeing papers submitted to academic journals, it is worth recalling an effort to reduce the refereeing time that was based on electronic mail communications between editors, referees, and authors.¹³ One of the major problems with that effort was the rudimentary graphics available for electronic mail systems at the time, but now the systems can handle molecular structures and other graphics (Figure 3). Thus a reduction of refereeing time might accompany increases in the number of papers submitted in machine-readable format to journals.¹⁴ In that vein, *The On-Line Journal of Current Clinical Trials* recently made its debut, featuring access to articles via modem and (Microsoft Windows or DOS) software provided by the publisher, keyword search of all the journal's published papers,

bit map images of figures, and hypertext links to MEDLINE abstracts of referenced articles.¹⁵

In addition to using the RDC approach or the ONDA, various sponsoring companies have submitted independent pilot CANDAs on microcomputers, which have included various combinations of spreadsheets; text in ASCII, Wordperfect, or Microsoft Word format; databases; and other features. The pilot project phase of CANDA development cannot continue indefinitely, however. In a recent paper, Cwiklo and Usdin¹⁶ noted that the demand for space for new systems and for time for the reviewers to learn them will limit the FDA's willingness "to try to use any automated support system that a drug company provides". A speaker from the FDA has also noted that the agency's systems staff can only provide support for a limited number of systems.⁶ A change from the current experimental approach to one incorporating standards is thus a key to further progress.

STANDARDS

Ideally, an integrated standard for the NDA, the Mt. Everest of compound documents, would incorporate standards for text, graphics, bit map images, and tables. For text, the "ASCII" file serves as a standard of sorts when retaining document formatting is not necessary. When formatting is required, Standard Generalized Markup Language (SGML) or possibly the formats of the most popular word-processing packages may serve. For vector graphics, the Hewlett-Packard Graphics Language (HPGL),¹⁷ the Computer Graphics Metafile (CGM),¹⁸ and encapsulated PostScript have become de facto standards.^{19,20} Plotting packages²¹ provide a means of generating HPGL graphic files from data that comes into computers from analytical instruments via RS232 or IEEE488 interfaces or through software controlling analog-to-digital boards. Several drawing packages can export graphics in the Computer Graphics Metafile format.

Bit map images from a scanner, a FAX device, or painting software place greater burdens on storage and communications than do vector graphics representations, particularly at high resolution. Compression methods can reduce the demand to an extent. CCITT group 3 and 4¹⁹ compression for image transmission and TIFF (tagged image file format) are important for image representation.

Because the hardware, operating systems, and software present in pharmaceutical companies are virtually always heterogeneous, often by choice, multiple formats for text, graphics, and bit map images are the result. Thus, converting from many formats to standard ones is a key part of preparing a compound document like the NDA. With text, little conversion is necessary for pure "ASCII" files. If formatting characters are present, however, even the best format converters do not provide complete fidelity between one text standard and another.²² Problems generally occur because a particular format code in one package has no equivalent in another. Conversions between graphics formats can also lead to problems. In one conversion to a mainframe graphics format, for example, the HPGL command to use pen #8 (thick black) became a command to use a line drawn in the background color. Thus the line was invisible.

Several organizations have defined integrated document content standards to handle compound documents. The Computer Assisted Logistics (CALS) initiative defines standards for interchange of text, vector graphics, images, and CAD drawings between the U.S. Department of Defense and its contractors and between contractors.²³ After 1992, the department will require that all documents submitted to it use

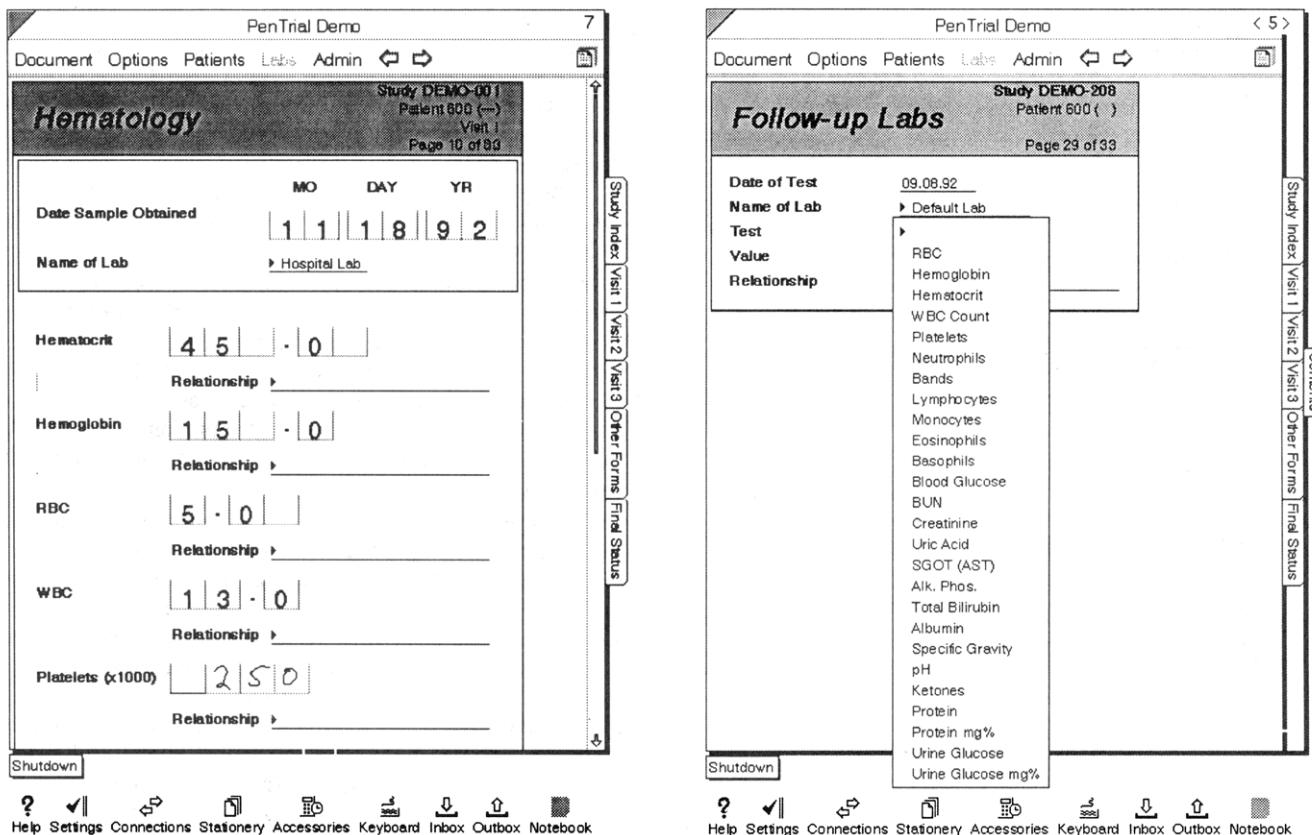


Figure 4. Pages from CRFs on a pen-based computer. (Left) Initial patient visit screen: The user writes values with the stylus in the boxes, and the system converts the handwritten characters to its own font by character recognition. (Right) For subsequent visits, the user selects an individual test from the list with the stylus and completes the name of the lab and the test value. The system then converts and analyzes the data. For both screens, tapping the stylus on the page icon at the upper right provides checking for completeness and values outside normal ranges. (Courtesy of Penergy, Inc., Redwood City, CA.)

these standards. The Air Transport Association and Association of American Publishers have defined standards for manuals on airliner maintenance and for book publishing, respectively. IBM, Digital Equipment, and the International Standards Organization have also defined content standards.¹⁹

Within the FDA, The Center for Biologics Evaluation and Review recently issued a guide for firms submitting PLAs or ELAs.²⁴ The guide includes a statement that the center is "...not yet prepared to suggest a standard format...", but it does state that Wordperfect for DOS and the Macintosh OS, a specific directory structure, PostScript, HPGL for vector graphics, and TIFF for images may be useful. For numeric or tabular data, ASCII files, the DIF interchange format from spreadsheets, and the dataset format of the SAS statistical software are acceptable. More recently, the FDA announced its intention to issue CANDA guidelines before the end of 1992.²⁵

In Canada, a joint task force of the PMAC and the Bureau of Human Prescription Drugs intends to develop standards for regulatory submissions.⁷ Among the elements proposed are hypertext linking of text, a cut-and-paste facility, and a clinical database that a reviewer may query through a menu, with a natural language query, or by filling in a representation of a CRF on the screen. SGML is also part of the proposed standards.

The FDA and the Pharmaceutical Manufacturers' Association (PMA) have introduced a standard for exchanging data between companies submitting NDAs and the biopharmaceutics division of the FDA. This standard, the Field Interchange Specification (FIS), defines a format for submitting data on patient demographics and measurements of plasma and urine concentrations of the prospective new

pharmaceutical and its metabolites. The definition is sufficiently general so that one can use any of several programs, on various types of hardware, to create it. The PMA has developed one program that submitters can use to create files conforming to the standard and a complementary program that reviewers in the biopharmaceutics division may use to query the data and to rearrange it to make comparisons. The developers of this standard emphasize, however, that data in the FIS format is not a substitute for the corresponding section of the paper NDA.

Up to this point, however, neither the FDA nor the Pharmaceutical Manufacturers' Association (PMA) has adopted standards as comprehensive as CALS. In the paper Standards for CANDAs; Do CALS Standards Point the Way?¹⁶ Cwiklo and Ussdin concluded that, though some tailoring of those standards for the pharmaceutical industry would be necessary, those standards are a good starting point. Independent of hardware or operating system, they have allowed the Department of Defense to avoid "support[ing] different hardware/software configurations for every weapons system it purchases". Adopting appropriate elements of CALS would allow the FDA and the industry to avoid the analogous problem with every new regulatory submission. The challenge is reaping the benefits of standards while still providing the flexibility to take advantage of new technical developments.

ELECTRONIC DOCUMENTS

Now that the FDA and the industry have considerable experience with CANDAs and electronic versions of other regulatory submissions, they must decide whether the electronic ones can become the official documents rather than

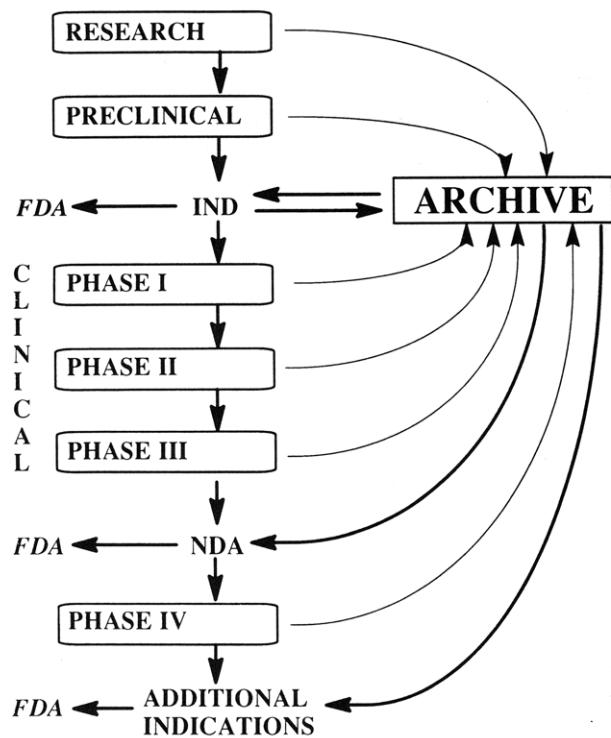


Figure 5. Information flow during development. The archives are largely paper documents at this point, but the industry is incorporating computer-based archives.

supplements to the paper ones. Reporting of adverse drug reaction (ADR) data to the FDA is one area in which electronic

submission is now an accepted practice.²⁶ When a drug product is on the market, pharmaceutical firms must submit FDA 1639 forms to report any "untoward effect that occurs in the course of use of a drug in professional practice" that is not listed on the drug label and that meets the definition of a serious ADR. Before electronic submission, firms would often print out paper copies of electronically-generated 1639 forms and submit them to the FDA, which then reentered the data into computers.

In pharmaceutical manufacturing, considerable effort has gone into the development of electronic batch record systems. Production operators may fill out forms displayed on computer screens to create these records, which describe the conditions for the manufacture of a particular batch of a pharmaceutical and the data from monitoring the production process.²⁷ At this time, however, paper copies of the records are the only ones with regulatory status.

The development of pen-based computers raises the possibility that they might become an alternative to paper forms for data entry at the sites of clinical trials. Data entered in the fields on a particular page of the pen-based CRF (Figure 4) may be linked directly to fields in the clinical database, as is true today with data entry software that uses keyboard-based computers. The pen system can check the data entered to be sure that it is reasonable [blood glucose should be greater than 0] and consistent with other data [(cholesterol in low density lipoprotein (LDL)] plus (cholesterol in high density lipoprotein (HDL)) should be less than total cholesterol].

Other examples of electronic documents, such as the electronic laboratory notebook,²⁸ are likely to find places in

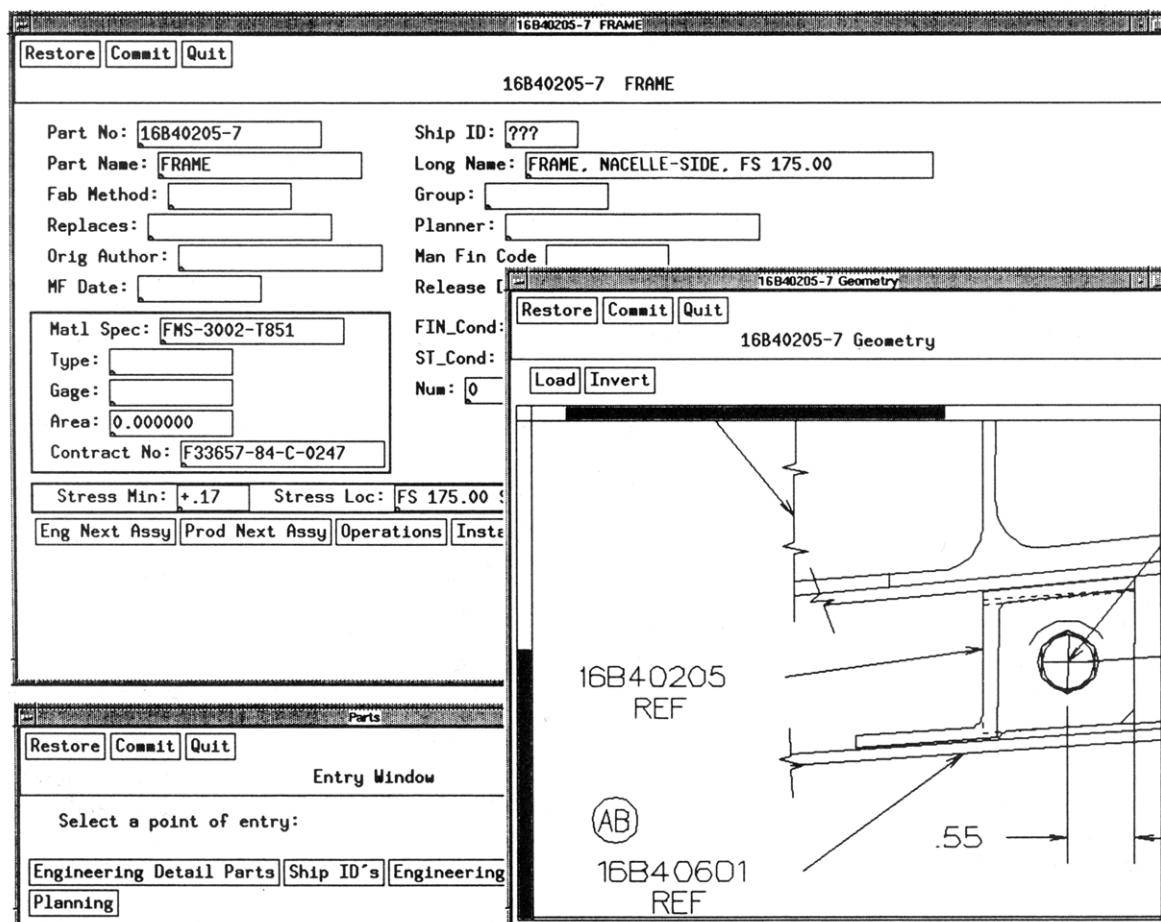


Figure 6. Screen displaying objects from a CAD object database. The part description in the window at the upper left includes toolbars for available functions ("Eng Next Assy," "Prod Next Assy," ...). When a user selects "Geometry" from the list, the system displays the engineering drawing at the lower right (and, in this case, hides the "Geometry" toolbar). (Courtesy of ONTOS, Inc., Burlington, MA.)

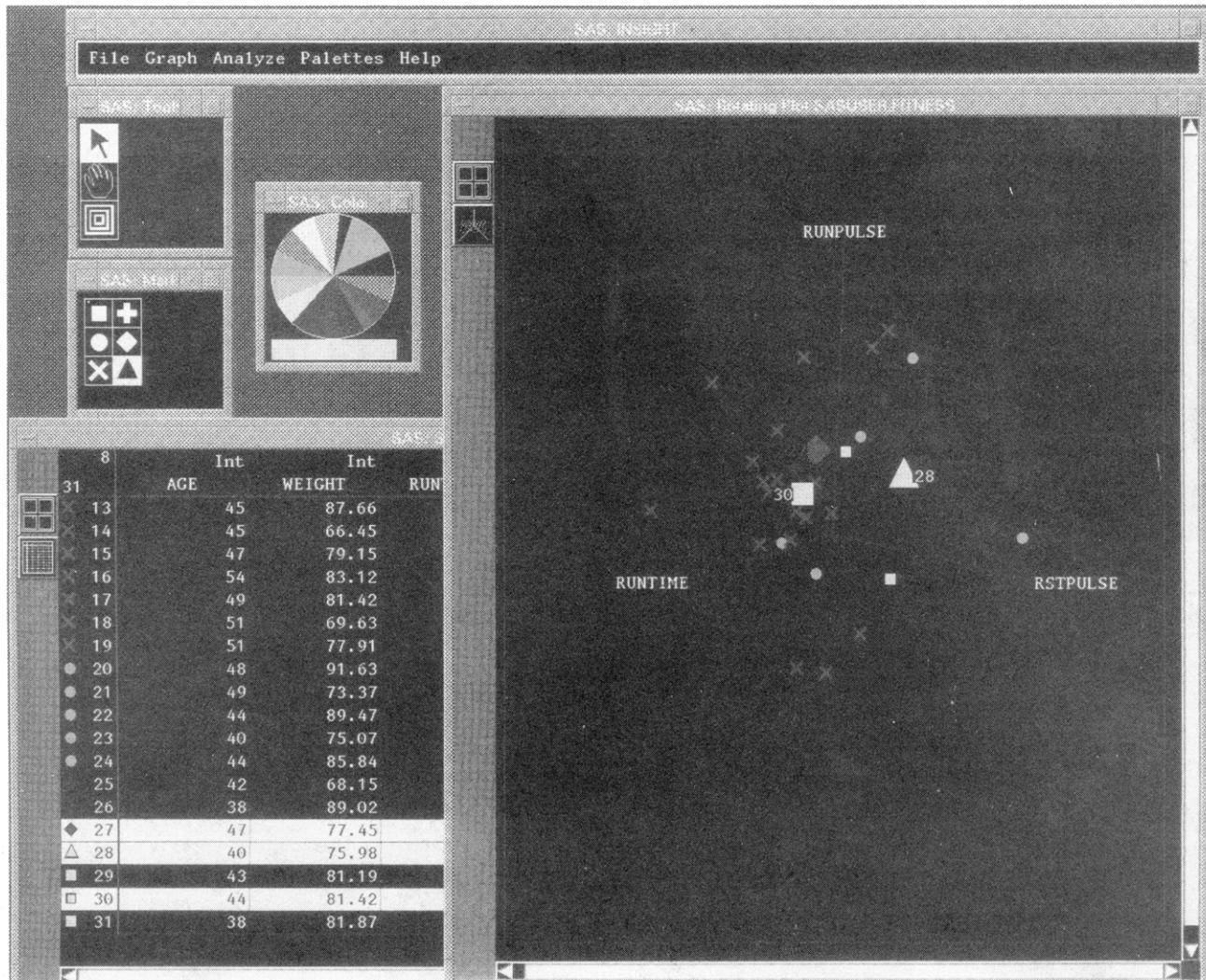


Figure 7. Presentation of the same data as a table and a three-dimensional graph. Selecting a particular point on the graph (# 30, near the origin, for example) automatically highlights, with reverse video, the data for that point in the table (Courtesy of SAS Institute, Cary, NC.)

the pharmaceutical industry and, thereby, press the question of their legal status. A consortium that includes representatives from companies in pharmaceuticals, chemicals, and other industries has formed to address the question "how do we verify that a particular piece of electronic information/data was generated by a particular person at a particular time?"²⁹ The consortium, which is open to all corporations, intends to present the guidelines developed on its initiative to the American Intellectual Property Association "with a view to seeking their assistance in obtaining endorsement from the U.S. Patent and Trademark office and Courts.". Once mechanisms are in place to authenticate the person and the time for creation of electronic records,³⁰ extending the procedures now used with form 1639 for ADRs to CRFs and other documents would be advantageous. The FDA recently issued an advance notice of proposed rulemaking that stated that the agency is studying electronic signatures "within the general context of three categories of current and future paperless (electronic) records: (1) Records maintained by industry which are subject to FDA inspection (e.g., batch production records for drug products, low acid canned foods, or infant formulas); (2) records submitted to FDA for review and approval, usually as part of research or marketing applications (e.g., new drug applications and food additive petitions); and (3) FDA's own records (e.g., sample collection reports) and notifications to industry (e.g., electronic mail)".³¹

DOCUMENT MANAGEMENT

As the pharmaceutical industry proceeds through the transition from manual NDA preparation procedures to electronic ones, it will benefit from the increased efficiency that accompanies generic computer publishing functions like automatic generation of cross references. There appears to be greater potential benefit, however, in the application of computers to management of the archive of reports, tables, and other materials that form the basis of regulatory documents. Once preparation of a particular document begins, it is critical that the preparers be able to query this heterogeneous archive and retrieve materials that fit the criteria of the query. Because the archive receives materials from many sources over a number of years (Figure 5), query and retrieval is a formidable challenge for an archive based on paper copies.

Databases for unstructured, natural language text and images are now available commercially, and pharmaceutical companies are applying them to managing components of those types in archives. Full-text retrieval systems³² provide the user with the ability to search a text document directly. As a result, one can initiate queries not possible with methods based on searching keyword/term indices with Boolean AND, OR, or NOT operators. For example, one can find all the instances in which the word "renin" occurs less than nine words away from "inhibitor" in a particular set of documents. In

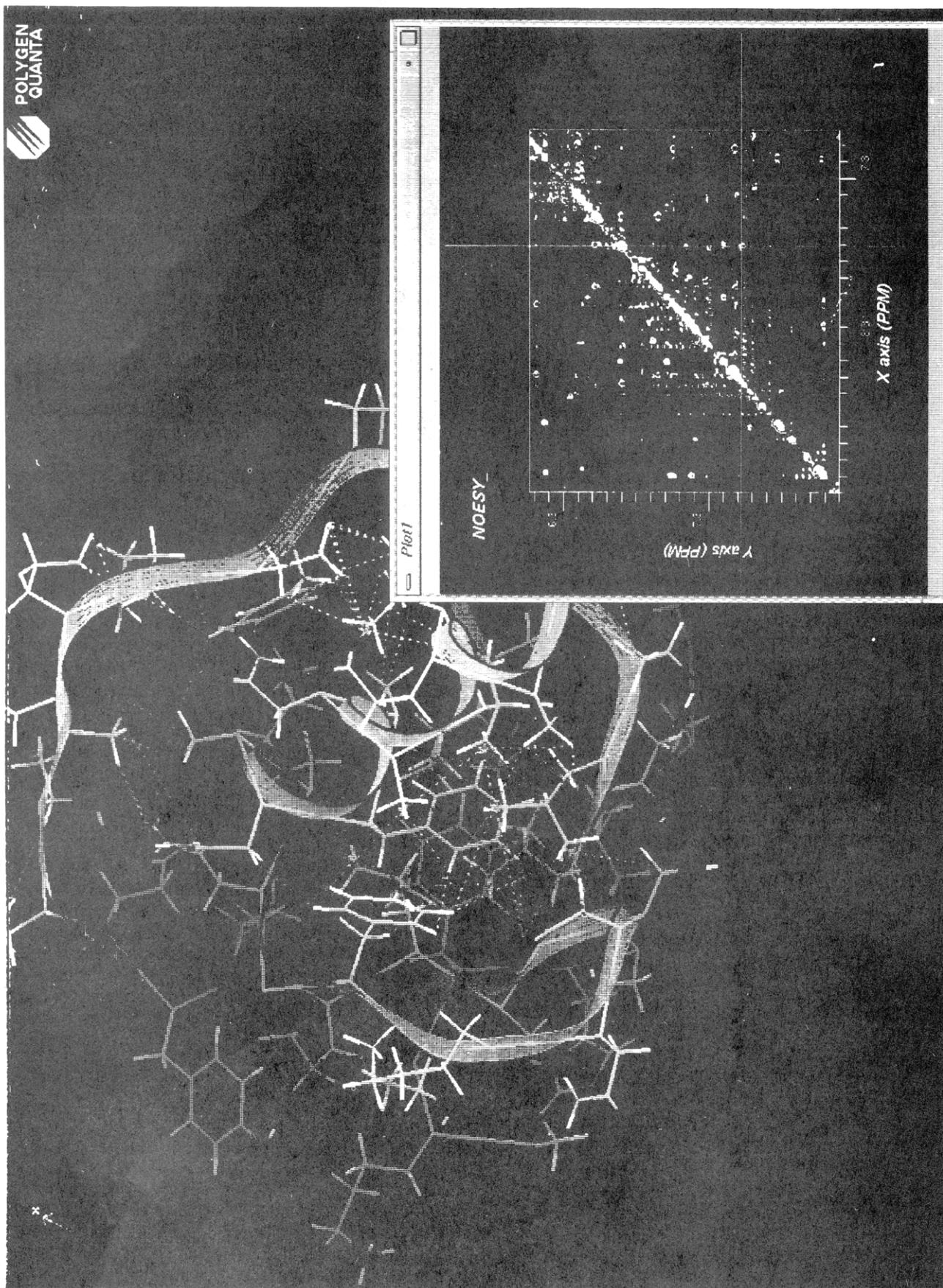


Figure 8. Displaying nuclear Overhauser effect (NOE) data (lower right window) and a chemical structure (upper left window). Selecting a particular subset of the NOE data marks, in the structure, the two protons produce the data and display the proton-proton distance calculated from that data. (Courtesy of Polygen-MSI, Inc., Waltham, MA.)

addition, full-text systems can apply weights to the various terms in the query and then rank the documents satisfying a given query in order of importance. In the case of image databases,⁵ it is necessary for a person entering a new image in the database to enter the descriptive data into an associated relational database. At this time, optical character recognition of image data is not sufficiently accurate to be acceptable for entering CRF information directly into a clinical database.

Though relational databases cannot handle image and free form text data directly, extensions to them can provide that capability.³³ In addition, some full-text retrieval systems provide facilities for searching Structured Query Language (SQL) databases. Potentially, the most powerful tool for providing a single query and retrieval system for text, images, data, and other elements is the object database. For pharmaceutical records management, important characteristics of the object database are that handling data in heterogeneous formats is inherent to it and that, between objects, it can provide relationships analogous to hypertext links. Computer-Aided Design (CAD)³⁴ and Very Large Scale Integrated Circuit (VLSI)³⁵ design systems have used these relationships (Figure 6). An example of such a relationship is available in SAS Insight, which links graphs and tables of the same data (Figure 7). Object databases should also meet the requirement that Warr,²⁰ quoting Ward,³⁶ placed on the ideal record management system: "it is simply not enough to be able to identify the existence of a document if one cannot display it, edit it, and transmit it, securely and accurately". Encapsulated functions that are part of the definition of a particular class of objects may provide facilities display, editing (if authorized), and transmission. Passing control to a standard software package, as in object embedding for Microsoft Windows, will usually make more sense than developing new functions to provide these facilities. For example, one would use a plotting package to convert raw ASCII data from analytical instruments into X-Y plots.²¹ Object databases for CAD and VLSI systems provide some of the functions, such as version control, required by a records/document management system suitable for managing an NDA archive, but it would be necessary to develop other essential functions.

For chemists, it is worth noting that the object database may also be a means of managing diverse formats of laboratory data: spectra, raw ASCII data from chromatography, images, text, and chemical structures. Relationships between objects should also be useful for this application, as in that between protons in a molecule and the region of the nuclear Overhauser effect data assigned to those protons (Figure 8).

CONCLUSION

Most people who have prepared CANDAs at the sponsoring companies have stated that they would submit additional ones in the future. All of the FDA reviewers interviewed in the PMA/FDA survey expressed an interest in seeing more CANDA submissions.² Though variables other than the impact of CANDA technology affect approval times, the people working in the field should take encouragement from statistics on the median time for review of NDAs. It has decreased from 30 months in 1989 to 26 months in 1990 and further to 22 months in 1991. In addition, changes that the sponsoring companies have made in internal procedures to accommodate CANDA preparation have improved the efficiency of many of those procedures.

ACKNOWLEDGMENT

I thank Hugh Peteet, Lofton Harris, Frederick Moehle, Agnes Farris, Ping Tsang, Margaret Stevens, Staley Wenger,

Scott Os, Douglas May, and Phillip Loftus for advice during preparation of this manuscript. K. K. Likhiani kindly provided information on the electronic records effort (ref 15). An early form of this paper was presented at the American Chemical Society meeting in Miami, FL, September 1989.

REFERENCES AND NOTES

- (1) Young, F. E. Submission of Drug Applications to the Food and Drug Administration Using Computer Technology. *Fed. Regist.* 1988, 53, FR 35912.
- (2) Albrecht, S. CANDA Project Evaluation Report—Survey Results. Presented at the Joint PMA/FDA Meeting on CANDA Review, Washington, DC, 1990. The transcript is available as an audiotape (no. 572-90) from Audio Transcripts, Alexandria, VA.
- (3) Harris, L. H.; Washburn, C.; Moehle, F. L. W. Future CANDA Technology. In *CANDA: A Regulatory, Technology, and Strategy Report*; Mathieu, M. P., Ed.; Parexel: Waltham, MA, 1992.
- (4) CANDA Activities: An Industry-Based Survey. In *CANDA: A Regulatory, Technology, and Strategy Report*; Mathieu, M. P., Ed.; Parexel: Waltham, MA, 1992.
- (5) Turtur, A.; Prampolini, F.; Fantini, M.; Guarda, R.; Imperto, M. A. IDB: An Image Database System. *IBM J. Res. Dev.* 1991, 35, 94–101.
- (6) Lipicky, R. L. CANDA Project Evaluation Report—FDA Senior Staff Perspectives. Presented at the Joint PMA/FDA Meeting on CANDA Review, Washington, DC, 1990.
- (7) Sutherland, D. J. B.; Levy, M. International Activities: Regulatory and Sponsor Perspectives. Presented at the Joint PMA/FDA Meeting on CANDA Review, Washington, DC, 1990.
- (8) Witzel, W. G. Germany Beams Ahead with Optical Disk Technology. *Pharm. Exec.* 1991, 11, 96–100.
- (9) Dobbs, J. H. Computer Assisted New Drug Applications: A Review. *J. Clin. Res. Pharmacopipid.* 1990, 4, 227–234.
- (10) Dobbs, J. H. The CANDA: an Overview. In *CANDA: A Regulatory, Technology, and Strategy Report*; Mathieu, M. P., Ed.; Parexel: Waltham, MA, 1992.
- (11) Berninger, C. CANDA Task Group Reports—Chemistry. Presented at the Joint PMA/FDA Meeting on CANDA Review, Washington, DC, 1990.
- (12) I am indebted to the members of the Chemists' CANDA Review Committee of the FDA for these points.
- (13) Dodd, W. P. Convergent Publication, or the Hybrid Journal: Paper & Telecommunications. *Electron. Pub.* 1990, 3, 47–59.
- (14) Brogan, M. C.; Garson, L. R. Requirements for and Challenges Associated with Submission of Machine-Readable Manuscripts. *J. Chem. Inf. Comput. Sci.* 1990, 30, 271–277.
- (15) Editorial Office; Information for Authors (editorial). *On-Line J. Curr. Clin. Trials* [Serial On Line] 1992; JUL1: 1992 (DOC #3): [3459 words; 98 paragraphs].
- (16) Cwiklo, W. E.; Usdin, B. T. Standards for CANDAs; DoCALS Standards Point the Way? *Drug Inf. J.* 1990, 24 (3), 531–536.
- (17) 7580A and 7585A Drafting Plotter Interfacing and Programming Manual: Option 001 and Option 002 using HP-GL Instructions; Hewlett-Packard Co.: San Diego, 1982.
- (18) Henderson, L.; Mumford, A. *The Computer Graphics Metafile*; Butterworth-Heinemann: Stoneham, MA, 1990.
- (19) Town, W. G. Integration of Microcomputer and Mainframe Information Systems: *J. Chem. Inf. Comput. Sci.* 1991, 31, 176–180.
- (20) Warr, W. A. Some Observations on Piecemeal Electronic Publishing Solutions in the Pharmaceutical Industry. *J. Chem. Inf. Comput. Sci.* 1991, 31, 181–186.
- (21) For example, SigmaPlot, available from the ACS.
- (22) Poor, A. Text Converting Software. *PC Mag.* 1989, 8 (8), 141–155.
- (23) Moehle, F. L. W. Future CANDA Technology. In *CANDA: A Regulatory, Technology, and Strategy Report*; Mathieu, M. P., Ed.; Parexel: Waltham, MA, 1992.
- (24) Center for Biologics Evaluation & Research, FDA. *Points to Consider: Computer Assisted Submissions for License Applications*, 1990.
- (25) Computer-assisted NDA Guidelines are Being Prepared by FDA for Release. *FDA Reports: The Pink Sheet*, June 24, 1991.
- (26) Anello, C. Electronic Submission of Periodic Reports. *Drug Inf. J.* 1989, 23 (2), 221–225.
- (27) Enrico, J. J. Electronic Batch Recording. *Pharm. Technol.* 1989, 13, 48–58.
- (28) Shipman, F. M., III; Chaney, R. J.; Gorry, G. A. Distributed hypertext for collaborative research: the Virtual Notebook System. *SIGCHI Bull.* (USA) spec. issue. 1989, Nov, P129-35 (proceedings of Hypertext '89 Pittsburgh, PA, Nov 5–8, 1989).
- (29) Likhiani, K. K. Authentication of Electronic Records for Legal Defensibility—Part II. Presented at the Scientific Computing and Automation Conference, Philadelphia, 1990.
- (30) Haber, S.; Stornetta, W. S. How to time stamp a digital document. *J. Cryptol.* 1991, 3, 99–111.
- (31) Electronic Identification/Signatures; Electronic Records; Request for Information and Comment. *Fed. Regist.* 1992, 57, FR 32185 [cited in *PMA News*. 1992, 34, 30].

- (32) Salton, S. Developments in Automatic Text Retrieval. *Science* 1991, 253, 974–979.
- (33) Codd, E. F. Extending the Data Base Relational Model to Capture More Meaning. *ACM Trans.* 1979, 4, 397–434.
- (34) Katz, R. H.; Chang, E. Managing Change in a Computer-Aided Design Database. In *Readings in Object-Oriented Database Systems*; Zdonick, S. B.; Maier, D., Eds.; Morgan-Kaufmann: San Mateo, 1990.
- (35) Afsarmanesh, H.; Knapp, D.; McLeod, D.; Parker, A. An Extensible Object-Oriented Approach to Databases for VLSI/CAD. In *Readings in Object-Oriented Database Systems*; Zdonick, S. B.; Maier, D., Eds.; Morgan-Kaufmann: San Mateo, 1990.
- (36) Ward, S. E. Text Retrieval—Its Place in Information Technology. In *Text Retrieval the State of the Art*; Gillman, P., Ed.; Taylor Graham: London, 1990; p 5; cited in ref 20.