

Chapter 10

Information Resources and Computer Systems

One of the biggest barriers to using available information is that most people do not know how to use existing resources or what systems are available for use.

John S. Wassom
Oak Ridge National Laboratory
March 1985

The best computer programs evolve into large creations. It is rare if possible to imagine a very large computer activity at the outset and build it as such.

Charles S. Tidball
The George Washington University Medical Center
March 4, 1985

CONTENTS

	P+I@
Sources of Research and Testing Data	219
Primary Literature	219
Secondary Literature	219
Unpublished Information	219
Information Centers	220
The Availability of Information	221
Journal Publication Policies	221
Federal Laws Affecting Unpublished Data	221
Barriers to Using Available Information	222
Data Quality and Comparability	222
International Barriers to Sharing Information	223
Retrieving Research and Testing Data	223
Abstracting and Citation Services ~	223
Retrieving Unpublished Information	224
Computer Systems	228
Advantages of Computers	228
Toxicology Data Bank	229
Registry of Toxic Effects of Chemical Substances	229
On-Line Literature	231
Laboratory Animal Data Bank	233
Building Phase, 1975-80	233
Public Accessibility, 1980-81	235
Reasons for the Failure of LADB	235
Lessons Learned From LADB	236
Expanding the LADB Concept: A Computerized Registry of Research and Testing Data	237
Summary and Conclusions	238
Chapter 10 References	239

List of Tables

<i>Table No.</i>	<i>Page</i>
10-1 Growth and Publication Frequency of Literature Related to Genetic Toxicology, Carcinogenicity, Mutagenicity, and Teratogenicity	220
10-2. Examples of Databases Available for Searches of Literature Involving Animal Research and Testing	230
10-3. Examples of On-line Databases of the National Library of Medicine.	231

List of Figures

<i>Figure No.</i>	<i>Page</i>
10-1. A Scientific Abstract and Corresponding Index Entry in BIOSIS	225
IO-2. Sample Bibliographic Entries in Biological <i>Abstracts/RRM</i>	226
10-3. Promotional Material From Commercial Supplier of Full Texts of Scientific Publications	227
10-4. A Typical Substance Entry in the Registry of Toxic Effects of Chemical Substances (RTECS)	232
10-5. A Representative Page of the Eight-Page Data Input Form for the Laboratory Animal Data Bank	234

Information Resources and Computer Systems

Earlier chapters have described the quantity and variety of data generated by using animals in research, testing, and education. To assess fully the alternatives to animal use in these areas, therefore, it is important to consider how the data are shared once they are generated. Anything that in-

creases information exchange reduces the need of other investigators to perform the same experiments. The pivotal role computers can play in that process has recently become an important topic for consideration and is examined in this chapter.

SOURCES OF RESEARCH AND TESTING DATA

Primary Literature

One of the most important ways to make data publicly available is through the "primary literature" in which they are published for the first time and in greatest detail. A significant form of this is the scientific journal, the most up-to-date and ubiquitous of the published sources available. Journal articles that are reviewed by knowledgeable peers before they are accepted for publication are considered especially reliable. Most normally contain a description of the methodology of the experiment, the results obtained, the conclusions drawn by the author or authors, and references to and discussions of related published and unpublished information.

Other primary sources are published reports (e.g., of Government-sponsored research), proceedings of technical meetings, or similar collections of articles. As a rule, reports and proceedings are not as widely available as journal articles. They may or may not have been peer-reviewed,

Secondary Literature

Secondary sources contain information drawn solely from other published material. The most common forms are books, reviews, and reports. (A book that contains original material would not be considered a secondary source.) Handbooks are a useful secondary source for numerical data and for citations to the primary literature in which they were first published. Because secondary sources draw from primary sources, the information they

report can be somewhat dated, as there is a time-lag ranging from months to years between the publication of a primary source and that of any secondary sources that rely on it.

Many reviews and reports are prepared to meet the specific needs of various organizations. Government agencies, such as the Food and Drug Administration and the Environmental Protection Agency (EPA), prepare reports to support regulatory activities. Research institutions, such as the National Institutes of Health and the Chemical Industry Institute of Toxicology (CIIT), prepare reports to announce the results of a particular study. Other organizations, such as the Chemical Manufacturers Association and the World Health Organization, prepare reports to further their programs.

Unpublished Information

Unpublished information about recent, planned, and ongoing research and testing can be of even greater interest than older, published information. The timelag between submission or acceptance of data for publication and their actual publication is often a handicap to those waiting to learn of experimental results. Time lost while waiting to obtain another investigator's published research results can cost a laboratory its claim to priority in obtaining research results. In testing, proprietary interests create pressure to obtain information as quickly as possible,

One of the oldest sources of unpublished information is networking—that is, the use of personal

contacts. Networking is affected by the economic factors discussed in chapter 11, such as the proprietary value of testing data and the incentives to make it public. Membership in scientific and professional societies and attendance at professional meetings facilitates this form of information exchange. Recent test results are often presented at meetings of professional societies, and valuable information about work in progress is exchanged by participants.

Unpublished data may not be written in report form, which makes it difficult to share the information. Although the data are stored in some kind of organized fashion, the way one person organizes information may not be useful to someone else. Thus, even if it is possible to determine that unpublished useful research or testing data do exist, it is often difficult to share them.

In addition to unpublished material, a separate category of information that is fairly inaccessible includes many Government reports, research institute reports, and obscure journals. This information falls into a grey area—"published" in a literal sense, but not in a practical one.

Information Centers

Because of the large volume of published and unpublished information that is generated, spe-

cial services called "information centers" have been set up to collect, organize, and disseminate it. An information center, to be comprehensive, must have a fairly narrow scope. These centers are a good vehicle for sharing unpublished information, although they do not have the resources to seek it out.

The most well known information center with holdings of research and testing data is the International Agency for Research on Cancer, in Lyons, France. The United Nations maintains several collections of published and unpublished data on chemicals potentially of international interest, e.g., through the International Program on Chemical Safety and the International Registry for Potentially Toxic Chemicals in Geneva, Switzerland. These agencies have a much broader scope than a typical information center, although they carry out many of the same functions. The Oak Ridge National Laboratory in Oak Ridge, TN, has individual information centers for environmental carcinogens, teratogens, and mutagens. Statistics on the volume and rate of growth of publications in the areas for which Oak Ridge has holdings are given in table 10-1.

Table 10.1.—Growth and Publication Frequency of Literature Related to Genetic Toxicology, Carcinogenicity, and Teratogenicity

Subject	Papers published per year ^a	Increase in papers published per year ^b	Publication sources providing information
Genetic toxicology.	4,000-5,000	200-300	3,400
In vivo animal carcinogenicity studies	1,500-2,000	50-100	1,000
In vitro cell transformation studies	400-500	25-50	500
Teratogenicity.	2,000-2,800	100-150	3,500

^aFigures would be substantially greater if augmented with unpublished or inaccessible published material.

^bNumbers shown are projected increases based on trends cataloged from the literature for the period 1979-84.

^cIncludes journals, books, symposium proceedings, government reports, and abstracts.

SOURCE: J.S. Wassom, Director, Environmental Mutagen, Carcinogen, and Teratogen Information Program, Oak Ridge National Laboratory, Oak Ridge, TN, personal communication, November 1985.

THE AVAILABILITY OF INFORMATION

One of the most important incentives to publish, both for people and for organizations, is to establish a professional reputation. Although “publish or perish” is an enduring part of academic tradition, in nonacademic research and testing sectors there is often little incentive to publish. As a rule, industry is more concerned with the protection of proprietary information and the conservation of financial resources than with publishing.

Federal organizations are likewise more interested in carrying out missions required by law than in the publication of research and testing data (unless that is their mission). As a result, many agencies' reports are never sent to the National Technical Information Service (NTIS) for distribution and cataloging, or too little time is spent indexing them in a fashion that facilitates easy retrieval of the information.

Journal publication Policies

Because of the importance of journals as a source of testing data, their publication policies are crucial to the effective exchange of information. Some journal policies (e.g., the limitations on the length of an article and the amount of detail it contains) are related to high printing and distribution costs. Others, such as an unwillingness to publish results that have already been disclosed publicly, are a result of the stiff competition that exists among journals.

One of the most frustrating publication policies from the standpoint of avoiding duplicative research and testing is that most journals (and therefore secondary sources) rarely publish negative results. It is natural that people would be more interested in knowing, for example, which chemicals have been found to be hazardous than which chemicals have not. As a consequence, a certain number of experimental protocols are repeated because the negative results of earlier experiments were not published. This policy is not likely to change without dramatic alterations in the stance of journal publishers, the policies of professional societies, and, indeed, the tradition of scholarly publication in academia. One notable exception

to this is the journal *Mutation Research*, which in 1977 made it a policy to also publish negative results.

Federal Laws Affecting Unpublished Data

One method available to the Federal Government for collecting testing data is to require them, either through registration requirements such as those under the Federal Insecticide, Fungicide, and Rodenticide Act (Public Law 92-516, as amended by Public Laws 94-140 and 95-396), or through reporting rules such as those promulgated under the Toxic Substances Control Act (TSCA) (Public Law 94-469). Section 8(d) of TSCA requires manufacturers and processors to submit citations or copies of health and safety studies they have sponsored, or about which they are aware, for specified chemicals. As of June 1984, EPA had received over 6,000 such submissions, about half of which were health-effects studies. For the specified chemicals, when regulatory notices were published in the Federal Register, about one-quarter of the citations were to data received under Section 8(d) (10).

Some unpublished data are given to Government agencies voluntarily, either through personal contacts or in response to publicity that the government is working on a particular problem. Much of the data concern adverse effects, but some concern negative results as well.

Unlike most countries, the United States has a policy of making information held by the Government as available as possible, consistent with protecting its proprietary value. Key laws in implementing this policy are the Administrative Procedures Act (Public Law 79-404, as amended by Public Law 89-554), which encompasses the Freedom of Information Act (Public Law 90-23, as amended by Public Laws 93-502, 94-409, and 95-454). This act makes all information held by the executive branch of the Federal Government available to anyone who asks for it, unless the information is specifically exempted or is protected under another law. The person requesting the information is frequently required to pay search and duplication costs, but the burden is on the Government to show why information should be withheld.

Under these laws, the public also has access to collections of published and unpublished nonproprietary data gathered to support administrative actions such as rulemaking. This '(public docket' contains all reports, literature, memos, letters, and other information considered in taking the action.

Once information has been obtained by the Federal Government, it may be shared within and among Government agencies. Often such sharing is very informal, and with informality comes unpredictability and oversights. Various committees have been set up to facilitate intragovernmental networking, such as the Interagency Regulatory Liaison Group of the late 1970s, the Interagency Risk Management Council, and the Interagency Toxic Substances Data Committee. These efforts increase the amount of information available to solve particular problems. They also reduce duplicative information requests made of industry.

In 1980, an interagency Toxic Substances Strategy Committee examined the sharing of information, focusing principally on the data held by Federal agencies (20). The Committee noted there were then more than 200 independent data systems, mostly incompatible. Barriers to sharing information included diverse methods of identifying chemicals and differing reliability and review of the databases. The Committee noted that coordination of Federal agencies' chemical data systems could reduce duplication of information gathering, minimize delay, and, to some extent, decrease uncertainties in decisionmaking. The benefits of such coordination would likely extend beyond the Federal Government to State and local governments, industry, labor, public interest groups, academic institutions, international organizations, and foreign governments.

BARRIERS TO USING AVAILABLE INFORMATION

Data Quality and Comparability

Before data are to be used, the user must be confident of their quality. This judgment is based on a variety of facts and inferences. People will frequently take into account the professional reputation of the investigator or the investigator's industrial, academic, or professional affiliation or organization. If the person has no reputation, good or bad, many scientists will not rely on that investigator's data. This phenomenon is most acute with investigations carried out in foreign countries and published overseas (14). Further, many scientists will not (and perhaps should not) trust results that have not been peer-reviewed. Lastly, some organizations tend not to trust any data that they have not generated.

It is important to assess the quality of data. Thus, even though numerical databases are convenient because they contain data in summary form, often there is no way to determine from the information contained there how reliable the data are (unless they were peer-reviewed before being put into the system). This problem has been addressed by the National Bureau of Standards (NBS), the Chem-

ical Manufacturers' Association, and others. A workshop held in 1982 (16) recommended that computerized databases (discussed at length later in this chapter) include the following "data quality indicators" that would allow the user to determine reliability for specific needs:

- the method(s) used to obtain the data,
- the extent to which the data have been evaluated,
- the source of the data, and
- some indication of the accuracy of the data.

An important part of evaluating data is comparing them with data obtained using similar methods—that is, validating the data. In deciding, for example, to rely on a particular test protocol, it is necessary to be confident not only that the test is a useful model of the effect of interest, but also that the results can be trusted, even though they are unexpected. For many investigators, validation involves repeating at least a portion of an experimental protocol in their own laboratories. They might also compare the results with those generated by other procedures with which they are more familiar.

International Barriers to Sharing Information

Animal research and testing is conducted in many countries (as described in ch. 16). The importance of communicating scientific information among nations has been recognized in the United Nations, in the Organization for Economic Cooperation and Development (OECD), and in regional and bilateral forums. Although much has been done to facilitate this, many barriers must still be overcome.

International communications cost more and take longer than domestic communications. Moreover, there are fewer international personal acquaintances on whom to rely for information than there are on a national level. Communication problems are exacerbated by institutional differences. It is difficult for industry-to-industry communications to occur, for example, when one industry is privately owned and another is government-owned, because governments typically deal through diplomatic channels.

Political animosities hinder information exchange. Defense-related information is affected the most, but all information sharing must suffer in such a climate. Even political differences cause problems in sharing information. It is difficult for agencies within the U.S. Government to obtain information from governments that have close working relationships with their industries, such as Ja-

pan, particularly when any information received would be subject to Freedom of Information Act requests in the United States.

Language differences are a large problem, both in the use of written materials and in personal communications. Translation and interpreting are expensive, particularly in the United States, where the number of people who speak more than one language has been decreasing. English translation costs for the four principal languages of science (French, German, Russian, and Japanese) range from \$40 to \$88 Per thousand words. An estimated \$4 billion to \$5 billion would be required to translate the current foreign-language holdings of the National Library of Medicine (NLM), for example, with an ongoing yearly translation cost of \$150 million (9). Duplicative translations are avoided through the clearinghouse effort of the John Crerar Library in Chicago, IL. Translations donated by a variety of sources on a broad spectrum of topics are made available to others.

Common protocols can also facilitate the international exchange of, for example, testing data. OECD members decided in 1981 that health-effects data generated according to OECD test guidelines should be mutually acceptable in all member countries, regardless of where the testing was done (see app. A) (17). Although this decision has not been fully implemented, OECD test guidelines are readily available and are receiving considerable use.

RETRIEVING RESEARCH AND TESTING DATA

The ways data are obtained and the amount sought are functions of the resources available for searching, how the data are to be used, the likelihood that the information exists at all, and how reliable the information is likely to be. Many methods for finding information are available, and most of them overlap to some extent.

Abstracting and Citation Services

In research and testing, several hundred thousand scientific articles in thousands of journals are published each year in the primary literature (6). Abstracting and indexing services and biblio-

graphic services play a vital role in making these accessible to those who need them. (An index based on references cited, or citations, permits the user to follow the literature into the future to locate pertinent articles. For example, a user with a 1981 article in hand who is seeking related, more recent publications can consult a citation index to identify 1985 publications that referenced the 1981 article.) Because animals are used for a variety of research purposes (see chs. 5 and 6), however, and because testing is interdisciplinary (see chs. 7 and 8), information may be indexed in the fields of chemistry, biology, pharmacology, medicine, and so on.

Abstracting and indexing services and bibliographic services have existed since the 17th century and have grown in number and size as published literature has expanded. The first major services for scientific information were published by professional societies (e.g., *Chemical Abstracts*). Some were sponsored by the Federal Government (e.g., Air Pollution *Abstracts* and *AGRICOLA*) or by commercial enterprises (e.g., *Current Contents* and *Environmental Abstracts*) (8). Some, such as the Chemical Information System, originated in Government and were later converted into commercial enterprises (12).

The largest abstracting and indexing service for biological and biomedical research is BIOSIS, the Biosciences Information Service. In 1985, its coverage extended to 440,000 items from over 9,000 sources worldwide. The file accumulated to date contains over 6 million items, the largest biological file in the English language. Items covered include abstracts and citations for journal articles and other serial publications, and citations to reports, reviews, and scientific meetings (6).

A typical abstract of a journal article and an illustration of how it is indexed by BIOSIS appear in figure 10-1. Information like this is contained in the semimonthly publication *Biological Abstracts*. Another publication, *Biological Abstracts/RRM*, contains bibliographic entries for research reports, reviews, meetings, and books (see fig. 10-2). BIOSIS also offers several computer-based services that provide citations tailored to the customer's information needs. All of these resources are regularly used by scientists. As the figures illustrate, however, it is often difficult to tell from a title, or even from an abstract, whether a particular article would satisfy a reader's needs.

Once a citation has been obtained, it is easy to acquire the full text of a research report. Most libraries have the necessary services available, or the inquirer can write to the author and ask for a reprint. In addition, some commercial vendors offer to supply by mail the full text of virtually any article (see fig. 10-3).

A recent comparison of databases for literature on 10 pesticides illustrates the problem of overlap (15). Eight databases had to be searched in order to get 90 percent of all data relevant to a par-

ticular regulatory decision. The share of citations produced by these databases that were not relevant ranged from 11 to 27 percent. Used together, the four most consistently relevant databases—TOXLINE, CAB Abstracts, BIOSIS, and *Chemical Abstracts*—produced 25 to 91 percent of all relevant citations, with an average of 69 percent.

These statistics illustrate the fragmentation that may accompany a literature search. Although the number of databases that need to be searched may be small for some fields, questions of an interdisciplinary nature require substantial resources for a complete literature search.

Retrieving Unpublished Information

Citation services are available for some unpublished data and testing in progress. Federal databases and publications include the *Bioassay Status Report* and Tox-Tips of the National Toxicology Program (NTP), the *EPA Chemical Activity Status Report*, the *Current Research Database* of the National Institute for Occupational Safety and Health, *NTIS's Federal Research in Progress*, and the *Smithsonian Science Information Exchange* (no longer active). There are also many small databases used to keep track of specialized data, such as information used in the implementation of a specific law.

Similar citation services to unpublished data or ongoing testing exist on an international level. The International Agency for Research on Cancer, which has substantial U.S. support, coordinates the sharing of information about current carcinogenicity testing in laboratories around the world and publishes an information bulletin, *Survey of Chemicals Being Tested for Carcinogenic Activity*. The International Program on Chemical Safety of the United Nations Environment Program (UNEP) is establishing a database for *Chemicals Currently Being Tested for Toxicological Effects*. This database is designed for long-term or otherwise expensive studies other than those on carcinogenicity. Participants in both programs include governments, industry, academia, and research institutes. In addition, Infoterra, a service of UNEP, publishes a directory through which experts in numerous subject areas can be located. Assistance is also provided by national representatives. The U.N.'s International Registry of Potentially Toxic Chemi-

Figure IO-I.—A Scientific Abstract and Corresponding Index Entry in BIOSIS

ABSTRACT FORMAT

	TOXICOLOGY-	Major Heading
	ENVIRONMENTAL AND INDUSTRIAL	~ Subheading
	Authors	Author Address
Reforsnco N u m b e r Abstract	- 23330. CARSONS, JOANNE N and JOHN O. GOULDEN (Arch= Oceanogr. Inst., Phila., Pa. 19103, USA.) The effects of chlorine pollution on growth and respiration rates of larval lobsters (<i>Homarus americanus</i>). BIOL RES 11(12): 1433-1438. 1985. The length, dry weight and standard respiration rate of larval lobsters (<i>H. americanus</i>) were measured following 20 days Immersion In coastal waters surrounding a power plant. Significantly lower increases in dry weight (P<. 05) and significant reductions in stan- dard respiration rates (P<. 01) were measured in exposed organisms when compared to control organisms. Water samples taken from the immersion site contained high concentrations of free Cl.	Article Title Journal Volume, (Issue), Page, Year

BIOSIS' INDEXING SYSTEM

AUTHOR INDEX (Personal or Corporate Names)				
Author Index	NAME	REF. NO.	NAME	REF. NO.
(Personal or Corporate Names)	CARSONS J N	23330+	ELL A W	26787
	CASEY N	29606	FINEMAN C	26884
	DAVIES R	24001	GOULDEN J O	23330
BIOSYSTEMATIC INDEX (Taxonomic Categories)				
Biosystematic index	ARTHROPODA.....		HIGHER TAXONOMIC CATEGORY	
(Broad Taxonomic Categories)	Crustacea.....		LOWER TAXONOMIC CATEGORIES	
	Malacostraca.....			
	Environmental and Industrial Toxicology.....		MAJOR CONCEPT	
	23330	23572	25352.....	REFERENCE NUMBERS
GENERIC INDEX (Genus-species Names)				
Generic index	GENUS-SPECIES	MAJOR CONCEPT	REF. NO.	
(Genus-species Names)	HOMARUS-AMERICANUS	TOXIC INDUS	23330	
		WILDLIFE AQU	24063	
	MICROCERUS-BERONI	CRUSTAC SYST	19145's	
SUBJECT INDEX (Specific Words)				
Alphabetic POSITION				
Subject index	SUBJECT CONTEXT	KEYWORD	REF. NO.	
(Specific Words)	GHT/ THE EFFECTS OF CHLORINE POLLUTION ON GROWTH		23330	
	TOBACTER/ EFFECT OF	SUBSTITUTION ON THE	26575	
	LORINE POLLUTION ON GROWTH AND RESPIRATION RATES		23330	
	GERMINATION RADICAL	BARRIER TEMPERATURE	27304	

SOURCE: *The 1985 8/OS/S Information Catalog* (Philadelphia, PA: Biosciences Information Service, 1985).

Figure 10=2. -Sample Bibliographic Entries in Biological Abstracts/RRM
EXAMPLES OF BIBLIOGRAPHIC ENTRIES IN BA/RRM:

CONTENT SUMMARY FORMAT

TOXICOLOGY ← Major Heading
 ENVIRONMENTAL AND INDUSTRIAL ← Subheading

Reference Number → 13003. BADASHKEYE, A.G., T.S. GALL, E.V. EFIMOVA, D. G. KNOORE, A. V. LEBEDEV and S. D. MYSINA. (Inst. Organic Chemistry, Siberian Div. USSR Acad. Sci., Novosibirsk, 630090, USSR.) FEBS (FED EUR BIOCHEM SOC) LETT 155(2): 263-266. ← Author Address

Bibliographic Detail → 1985. *Biological effects of chlorination on coastal waters.* / HOMARUS AMERICANUS, LARVAE, JUVENILE MORTALITY, RESPIRATION DECLINE, POWER PLANT, POLLUTION ← Subject Terms

Subject Concepts → CON: Animal Ecology/Oceanography/
 Air, Water, & Soil Pollution/
 Comparative & Experimental Crustacea.
 TAX: Malacostraca

BOOK SYNOPSIS FORMAT

Main Entry → 8720. OGLEVEY, O. C. and M. STIGGS (Ed.). (Lab. Biochem., Swiss Fed. Inst. Tech., Zurich, Switz.) Springer-Verlag; New York, N.Y. USA; Berlin, West Germany. Illus. Paper. ISBN 0-387-06842-1 ISBN 3-540-06842-1. x+175 p. 1984 (recd. 1985). **Membrane biochemistry: Bacterial membrane transport.** This manual presents a series of 17 experiments on membrane transport and membrane bioenergetics in bacteria, covering a large spectrum of modern techniques in the field of membrane biochemistry. Tables, mathematical illustrations and diagrams complement the text. The work is intended for graduate and advanced undergraduate students, but should also be of value to researchers in the field.

Book Synopsis →

Book Chapters → 8721. SPONSON, HAROLD. (Lab. Biochem., Swiss Fed. Inst. Technol., Zurich, Switz.) 164-170. **Characteristics of ionophores using artificial lipid membranes.** / BLACK LIPID MEMBRANE, ANIONIC-CATIONIC PERMEABILITY, BIIONIC POTENTIAL, MEMBRANE CONDUCTANCE

8722. SCHLEUTER, HERMAN and PATRICK COLLETTI. (Lab. Biochem., Swiss Fed. Inst. Technol., Zurich, Switz.). 171-176. **Characterization of neural and charged ionophores using vesicular artificial lipid membranes (liposomes).** / HYDROCARBON STRUCTURE, THERMODYNAMICS, MINERAL DIFFUSION

MEETING FORMAT

4718. SLOAN, GARTH and I. WILLINGSBY (Ed.). (Dep. Physiol. Pharmacol., Vet. Sch., Toulouse, Fr.) ANN Rech VET 10 (2/3): 157-502. 1985 [In Fr.] **Communications on Rumen Physiology and Forage Degradation at the 5th International Symposium on Ruminant Physiology.** Clermont-Ferrand, France. Sept. 3-7, 1985. [ABSTRACTS ONLY] ← Main Entry

4719. CROFT, L. (Rowett Res. Inst., Bucksburn, Aberdeen, Scott., U.K.) 157-159. [In Fr.] **Effects of cannulation on intestinal motility.** / SHEEP, OCCLUSION, SURGERY, RADIOLOGY ← Meeting Papers

4720. SCHWAB, R., E. SCHWAB, F. TANG and B. BEAUGEROUX. (Fac. Sci. Agron. Etat 5500 Grembloux, Belg.) 163-165. [In Fr.] **Rate of passage of digesta in sheep.** / HAY, FECES, RADIOTRACER

SOURCE: *The 1985 BIOSIS Information Catalog* (Philadelphia, PA: Biosciences Information Service, 1985).

Figure 10=3. - Promotional Material From Commercial Supplier of Full Texts of Scientific Publications



Pick an article-any article from this issue of CC! It can be in your hands *fast* when you order it from ISI'S document delivery service, *The Genuine Article*

The Genuine Article can supply you with original article tear sheets or quality photocopies of nearly all journal articles, editorials, letters, and other items you see in this issue of *Current Contents*. To order, simply fill out the coupon below and mail it to ISI, together with your check or money order

Price information: Any article of ten or fewer pages costs \$7.50 (when order includes ISI Accession Number). This amount includes first class mail delivery to the U. S. A., Canada, and Mexico. Air mail to all other locations costs \$8.50. For every additional ten pages or fraction of ten pages, there is an additional charge of \$2 per article.

For complete information on our document delivery service, write ISI Customer Services at the address shown.

FAST, RELIABLE DOCUMENT DELIVERY

To receive your order from *The Genuine Article*, fill in oval with ISI accession number adjacent to journal title in CC.



Journal _____
 Vol _____ No _____
 Date _____ 1st Page No _____
 Author _____

- Check enclosed.
- ISI Stamps attached
- Please send me information on volume discounts or deposit accounts

Name Title _____

Organization Dept _____

Address _____

City _____ State Province _____ County _____

ZIP Postal Code _____ Telephone _____

Prices are subject to change

Payment must accompany order

ISI Institute for Scientific Information®

The Genuine Article. 3501 Market Street, Philadelphia, PA 19104 U.S.A.
 Telephone. (21 5) 386-0100, ext. 1140, Cable SCINFO, Telex: 84-5305

European Office: 132 High Street, Uxbridge, Middlesex UB8 1 DP, UK.
 Telephone: 44-895-70016, Telex: 933693 UKISI.

1985 ISI
 CCS 7.16

SOURCE: Institute for Scientific Information, Inc., Philadelphia, PA.

calls sometimes refers information requests among member countries through its national correspondents.

A recent U.S.-led project of the OECD, generally referred to as "Switchboard," has also addressed the problems of obtaining information from other countries. Unpublished information may be requested through the Switchboard system for use in risk assessments or to otherwise protect health and the environment. A pilot system is to be run in which two requests per participating country

per month would be referred to some combination of government agencies, industry, academia, and research institutes that might have unpublished data relevant to the request. The mechanisms for referring requests on the national level and the enlistment of various organizations, either as requesters or responders, is the responsibility of Switchboard's national focal point. This project will begin on a small scale and will be monitored. If appropriate, it could be expanded (18).

COMPUTER SYSTEMS

Computers have two applications as an alternative to using animals in research, testing, and education. First, they can be used to model or simulate biological, chemical, and physical systems. In this way, a computer could be used as a direct replacement for some number of animals used in laboratories. This form of computer use is discussed in chapters 6, 8, and 9. Second, computers are used to disseminate information that has been generated from prior use of animals in research and testing, thus avoiding the needless repetition of a procedure by other scientists. It is this role of computers as information disseminators that is discussed in the rest of this chapter.

Advantages of Computers

Biological testing (see ch. 7) can be described as the repetitive use of a standard biological test situation, or protocol, employing different chemicals or different test parameters (e.g., species or biological end points). Because the protocols in testing are more stereotyped and less varied than those in research, biological testing is more amenable than research to the institution of a computerized data retrieval system. In fact, testing emerged in the 1970s as the first discipline in which such a system was developed.

If a comprehensive, computerized registry of biological research or testing data were established, certain benefits might accrue. These benefits are predicated on the inclusion in the computerized registry of both control and experimental data, and of both positive and negative results. (Data

obtained from testing fall into two broad categories: those derived from untreated (control) subjects, and those from treated (experimental) subjects. Data obtained from treated subjects may either show an effect from the treatment ("positive results") or no effect ("negative results").) Furthermore, the advantages of such a registry depend on the acceptance by working scientists of the data contained in it—acceptance that seems possible only with the imprimatur of peer review of the data. The anticipated benefits of a computer-based registry of research or testing data include:

- **Decreased Use of Animals in Research or Testing.** In some instances, an investigator would locate the exact data desired, possibly from a previously unpublished source, thus avoiding unintentional duplication of animal research or testing. Baseline data could permit the selection of a dose, a route of administration, or a strain of animal without the need for new animal experiments to establish these factors. Efficiencies could also include the use of fewer doses on smaller numbers of animals. Conceivably, the number of animals required for control groups could be reduced, although many experimental protocols require the use of concomitant control subjects, rather than of data from a pool of control subjects, in order to achieve statistical significance.
- **A Check for Genetic Drift.** Certain experimental results can change over a span of many generations due to subtle, progressive changes in the underlying genetic constitution of the strain of animals ("genetic drift"). The regis-

try would provide baseline data within specified time frames of measurement, and make it easy to check for the possibility of genetic drift.

- **New Perspectives on Old Data.** By performing statistical comparisons across data sets and identifying relationships not already obvious, unforeseen relations could be established without animal experimentation.

The scientific community makes use of a number of computerized literature retrieval services to obtain bibliographic citations and abstracts to the published literature. Most abstracting and indexing services started as publications, but most are now available on-line as well. Others, such as *AGRICOLA*, are only available on-line.

Many handbooks and other numerical databases are also available on-line. Several numerical databases are sponsored by the Federal Government. The most comprehensive, the recently terminated Laboratory Animal Data Bank, is reviewed in detail in the next section. Two current systems, the Toxicology Data Bank and the Registry of Toxic Effects of Chemical Substances, are discussed in some detail here. Table 10-2 lists a number of databases available for searches of the research and testing literature. Table 10-3 lists some widely used databases of the NLM.

Toxicology Data Bank

The Toxicology Data Bank (TDB) was made public by NLM in 1978. It is designed to address some of the needs of the testing and regulatory communities for toxicity information. TDB is organized by individual chemicals or substances, now totaling more than 4,000. Its fixed format includes:

- data on the production and use of each chemical;
- a description of the physical properties of each chemical; and
- the results of pharmacological and biochemical experiments, and information on toxicological testing.

TDB is based on conventional published sources and does not include unpublished data. Thus, baseline data on control animals, which might be used in place of a control group, could not be included because so little has been published.

The most valuable feature of TDB is the fact that all the data it contains are peer-reviewed. As a consequence, its data summaries are acceptable to most users (5). (Another database containing only peer-reviewed data is the Environmental Protection Agency's Gene-Tox.)

Registry of Toxic Effects of Chemical Substances

The Registry of Toxic Effects of Chemical Substances (RTECS) has been published annually since 1971 by the National Institute for Occupational Safety and Health, under Section 20(a)(6) of the Occupational Safety and Health Act of 1970 (Public Law 91-596). RTECS is a compendium, extracted from the scientific literature, of known toxic and biological effects of chemical substances. RTECS does not evaluate the data it cites, leaving that responsibility to the reader. An example of the information contained in a typical substance entry in RTECS is given in figure 10-4.

By congressional mandate, those data that indicate a toxic effect of a chemical are to be included in RTECS; those that show no toxicity are to be excluded. Thus, RTECS does not include negative results. Moreover, a chemical might not be included in the registry for a variety of reasons, including the following:

- The test results could not be cited because the protocol of the study did not meet the RTECS selection criteria.
- The substance has not yet been tested or the results have not yet been published.
- The substance has been tested and the results published, but the information has not yet been entered into the RTECS file.

The exclusion of negative results from RTECS and its incompleteness for these other reasons may lead to the repetition of toxicity testing of essentially nontoxic substances.

The production of RTECS costs approximately \$500,000 per year. The current quarterly update includes a total of 68,000 compounds, and it continues to grow steadily toward the estimated 100,000 unique substances for which toxicity data may be available. If RTECS were expanded to include all results of whole-animal toxicity testing, including

Table 10.2.—Examples of Databases Available for Searches of Literature Involving **Animal Research and Testing**

Database	Description	First year covered
AGRICOLA	Worldwide journal and monograph literature on agriculture and related subjects; from the National Agricultural Library	1970
AQUACULTURE	Growth requirements, engineering, and economics of marine, brackish, and freshwater organisms; from National oceanic and Atmospheric Administration	1970
AQUALINE	Abstracts from world literature on water, waste water , and aquatic environments; from Water Research Centre, Stevenage, U.K.	1974
ASFA (Aquatic Sciences and Fisheries Abstracts)	Life sciences of seas and inland waterways plus legal, political, and social implications of aquatic life; from UNESCO	1978
BIOSIS Previews	International coverage of life science research; from <i>Biological Abstracts</i>	1969
CA Search	International coverage of chemical sciences; from Chemical Abstract Service	1967
Comprehensive Dissertation Abstracts	Author, title, and subject guide to nearly all American dissertations since 1861 and many from foreign countries; abstracts added beginning in July 1981; from Xerox University Microfilms	1861
Conference Papers Index	Records of scientific and technical papers presented at major regional, national, and international meetings each year; from Data Courier, Inc.	1973
CRIS (Current Research Information System)	Research in agricultural sciences; from U.S. Department of Agriculture's State Research Service	1974
Enviroline	International coverage of biology, chemistry, economics, geology, law, management, planning, political science, and technology of environmental issues; from Environment Information Center, Inc.	1971
Environmental Bibliography	Atmospheric studies, energy, general human ecology, land resources, nutrition and health, and water resources; from Environmental Studies Institute	1973
Excerpta Medica	Worldwide citations and abstracts from 3,500 biomedical journals; from Excerpta Medica	1974
INSPEC	Coverage of literature in computers, electrotechnology, and physics; from the American Institute of Electrical Engineers	1969
I PA (International Pharmaceutical Abstracts)	Literature on drug development and use of drugs; from the American Society of Hospital Pharmacy	1970
IRL Life Sciences Collection	Worldwide coverage of life sciences including conferences; from Information Retrieval, Ltd.	1978
ISI/BIOMED	Index of 1,400 biomedical journals; from the Institute of Scientific Information	1979
ISI/COMPUMATH	Covers literature in computer science, mathematics, statistics, operations research, and related areas; from the Institute for Scientific Information	1976
ISI/ISTP&B	Computerized version of Scientific and Technical Proceedings and Books. Covers 3,000 proceedings and 1,500 books annually; from the Institute for Scientific Information	1978
LISA (Library Science Abstracts)	international coverage of library and information science literature; from Learned information, Ltd.	1969
Microcomputer Index	Subject and abstract guide to 21 microcomputer journals; from Microcomputer information Services	1981
NIMH	Mental health literature from 950 journals, symposia, government reports, and other sources; from the National Institute of Mental Health	1969
Oceanic Abstracts	International literature on geology, governmental and legal aspects of marine resources, marine biology, marine pollution, meteorology, and oceanography; from Data Courier, Inc.	1964
Pollution Abstracts	Literature on the sources and control of environmental pollution; from Data Courier, Inc.	1970
Population Bibliography	international coverage of population research: abortion, demography, family planning, fertility studies, and migration; from Carolina Population Center, University of North Carolina	1966
Psychological Abstracts	Worldwide coverage of literature in psychology and related social-behavioral literature; from the American Psychological Association	1967
SCISEARCH	International literature of sciences and technology; from the institute for Scientific Information	1974

Table 10-2.—Examples of Databases Available^a for Searches of Literature Involving Animal Research and Testing (Continued)

Database	Description	First year covered
SOCIAL SCISEARCH	Worldwide coverage of social and behavioral sciences literature; from the Institute for Scientific Information	1972
TELEGEN	Covers literature on biotechnology and genetic engineering in 7,000 sources including conference and symposia papers, government studies, periodicals, and the popular press; from Environment Information Center, Inc.	1973
Zoological Record	Covers zoological literature from 6,000 journals; from Biosciences Information Service and the Zoological Society of London	1978

^aThese databases are available by telephone connection to one or more of the following: Lockheed Information System DIALOG, System Development Corp 'S ORBIT, and Bibliographic Retrieval Service, Inc.

SOURCE: Adapted from R.V. Smith, *Graduate Research* (Philadelphia, PA: ISI Press, 19S4).

Table 10-3.—Examples of On-Line Databases of the National Library of Medicine

Name	Description	Number of records (average length)	Type of record
MEDLINE	1966-present. Bibliographic citations and abstracts from primary biomedical literature	3,300,000 (1,250 char.)	Bibliographic
TOXLINE	1965-present. Abstracts from primary toxicological literature	1,400,000 (1,050 char.)	Bibliographic
CHEMLINE	Dictionary to chemicals contained in TOXLINE and other MED-LARS data bases	500,000 (275 char.)	Chemical compound
RTECS	Brief summaries of toxicity results from primary literature	68,000 (1,000 char.)	Chemical compound
TDB	Detailed chemical, pharmacological, and toxicological data and extracts from monographs and handbooks	4,000 (17,000 char.)	Chemical compound

SOURCE: Office of Technology Assessment

negative results, its size would be increased by an estimated 10 to 15 percent (11). RTECS is available in hard copy (19), on microfiche, on magnetic tape, and on-line from both the MEDLINE service of NLM and the Chemical Information System, a joint resource of several Federal agencies that is managed by EPA.

On-Line Literature

The research community makes use of a number of computerized literature retrieval services to obtain bibliographic citations and abstracts from primary literature. Among these, for example, is NLM's MEDLINE database, a bibliographic file now exceeding 3,300,000 entries. In the private sector, Biosciences Information Services prepares hundreds of thousands of abstracts each year, providing access to essentially the entire published biological research literature. However, the re-

search community is not presently served by a computerized database that includes comprehensive descriptions both of experimental protocols and of the resulting data.

Movement toward on-line delivery of the full text of scientific publications has begun in the private sector. For example, Mead Data Central (Dayton, OH) offers MEDIS, a medical literature database. In 1985, the MEDIS service included about 70 publications, with some stored journal articles going back to 1980. MEDIS includes the full text of the *Journal of the American Medical Association* (since 1982), *Archives of Internal Medicine*, and some textbooks and newsletters. In 1984, Bibliographic Retrieval Services (Latham, NY) joined with publisher W.B. Saunders Company to offer the full text of the *New England Journal of Medicine* and several other journals on-line. A serious limitation to any current full-text literature retrieval system is the inability to retrieve graphs, photographs, and other images (7).

Figure 10-4.—A Typical Substance Entry in the Registry of Toxic Effects of Chemical Substances (RTECS)

A	1	3	4	7	
AA1111111	RTECSIUM				
②	UPDT: 8209	CAS: 9999-99-9	MW: 357.88	MOLFM: C. Re-S-Te	
⑥	SYN: NIOSHRTECSIUM • RTECS (DOT) • RTECSIAM (French) " TSL (06 S.)				
⑦	IRRITATION DATA AND REFERENCES:				
	skn-hmn 5%/1M nse MOD				⑫ JAMAAP 135,1920,68
	skn-rbt 10 mg/24H open MLD				A M I H B C 10,61,64
	eye-rbt 20 mg				AMIHBC 10,61,64
⑧	MUTATION DATA AND REFERENCES:				
	mma-sat 50 ug/plate				PNASA6 72,513.75
	cyt-hmn: leu 100 umol/L/8H				MUREAV4,53,67
	dlt-mus-ori 35 mg/kg				TXAPA9 23,288,72
	hma-mus/sat 400 ppm				GISAAA 42(1),32,77
⑨	REPRODUCTIVE EFFECTS DATA AND REFERENCES:				
	1T03	ori-rat TDLo:270 mg/kg (5D male)		TJADAB 19.41A,79	
	3T25T46T52	ihl-rat TCLo:100 ppm (3D pre/4-7D preg)		NTIS** AD-900-000	
	2T35T75	mul-mus TDLo:900 mg/kg (1D male/6-15D preg/5D post)		TOXID9 1,125,81	
⑩	TUMORIGENIC DATA AND REFERENCES:				
	2V01M61	ori-rat TDLo:90 gm/kg/78W-l		29ZUA8 -,183,80	
	1V03	ihl-mus TCLo:400 ppm/8W-l		JTEHD6 - [SUPPL.2:,69,77	
	2V02K60	ori-rat TD :100 gm/kg/104W-l		BJCAAI 16,275,62	
⑪	TOXICITY DATA AND REFERENCES:				
		ihl-hmn TCLo:500 mg/m3/7M		AIHAAP 23,95.62	
	2J13K18	ori-rat LD50:1500 mg/kg		MarJV# 26 Apr 76	
		scu-mus LDLo:1200 mg/kg		FCTXAV 17(3),357.79	
⑬	AQUATIC TOXICITY RATING: TLm96:100-10 ppm				WQCHM" 2,-.74
⑭	REVIEW: CARCINOGENIC DETERMINATION: ANIMAL POSITIVE				IARC 20,151,80
	REVIEW: TOXICOLOGY REVIEW				PLMJAP 6(1),160,75
	REVIEW: TLV-TWA 10 ppm; STEL 25 ppm				DTLVS* 4,368,80
	REVIEW: TLV-SUSPECTED CARCINOGEN				DTLVS" 4,358.80
⑮	STANDARDS & REGULATIONS:				
	OSHA STANDARD-air:TWA 50 ppm				FEREAC 39,23540,74
	DOT-ORM-A, LABEL: NONE				FEREAC 41,57018,76
⑯	CRITERIA DOCUMENT: OCCUPATIONAL EXPOSURE TO RTECSIUM recm				NTIS**
	std-air:CL 10 ppm/60M				
⑰	STATUS. SELECTED BY NTP FOR CARCINOGENESIS BIOASSAY AS OF SEPT 1982				
	STATUS: NTP SECOND ANNUAL REPORT ON CARCINOGENS, 1981				
	STATUS: NIOSH MANUAL OF ANALYTICAL METHOOS, VOL 3 S255				
	STATUS: NIOSH CURRENT INTELLIGENCE BULLETIN 41, 1980				
	STATUS: REPORTED IN EPA TSCA INVENTORY, 1962				
	STATUS: EPA TSCA 8(a) PRELIMINARY ASSESSMENT INFORMATION				FEREAC 47,26992,82
	FINAL RULE				

Key to Figure 10-4

- A. RTECS accession number, a sequence number assigned to each substance in the Registry.
1. Substance name.
 2. Date when substance entry was last revised.
 3. American Chemical Society's Chemical Abstracts Service unique identification number for the substance.
 4. Molecular weight of the substance,
 5. Molecular or elemental formula of the substance.
 6. Synonyms, common names, trade names, and other chemical names for the substance. Skin and eye irritation data.
 8. Mutation data.
 9. Reproductive effects data.
 10. Tumor-causing data.
 11. Toxicity data.
 12. Acronyms for the references from which the data and other citations were abstracted.
 13. Aquatic toxicity rating.
 14. Reviews of the substance.
 15. Standards and regulations for the substance promulgated by a Federal agency.
 16. A Criteria Document supporting a recommended standard has been published by NIOSH.
 17. Status information about the substance from NIOSH, EPA, and the National Toxicology Program.

SOURCE: U.S. Department of Health and Human Services, National Institute for Occupational Safety and Health, Registry of Toxic Effects of Chemical Substances, R.L. Tatken and R.J. Lewis, Sr. (eds.) (Cincinnati, OH: DHHS (NIOSH) Pub. No. 83-107, 1983).

LABORATORY ANIMAL DATA BANK

The Laboratory Animal Data Bank (LADB) is a computerized set of records of baseline data of physiological, histological, and other biological properties of mammalian species (largely rodents) used in research and testing. The data contained in LADB were derived from both research and testing, and are relevant to both areas of animal use. Although LADB exists today only as an archival reference, and is no longer publicly available online, it is of great historical interest in a consideration of computer-based information resources.

In 1970-73, as the carcinogenesis bioassay program of the National Cancer Institute (NCI) was developed, NCI's Division of Cancer Cause and Prevention anticipated needing better access to baseline data for experimental animals. In 1973-74, NLM helped formulate the concepts leading to LADB. The major contributor of funding for LADB was NCI.

Data for LADB were derived from published and unpublished reports. Only control, or baseline, data from groups of animals were included. The data were collected and entered into LADB via a standard, eight-page form (reproduced in ref. 2) that surveyed 306 variables, including:

- name and manufacturer of the animals' feed,
- vaccinations given to the animals,
- organs or tissues routinely examined at autopsy,
- blood variables that were analyzed,
- detergent used in washing cages, and
- source of the animals.

The first page of that form is reproduced in figure 10-5.

Building Phase, 1975-so

Battelle Laboratories (Columbus, OH) was awarded an NLM contract in 1975, after a competitive procurement, and began detailed design activities in 1975-76. Methods for obtaining data were developed, and the data file was designed to permit interactive access, or time-sharing, by users. Sufficient data were entered to permit initial study by NLM staff in 1976, and in the following year 13 outside users were allowed to test the system.

In June 1976, NLM requested the Institute of Laboratory Animal Resources (ILAR) of the National Academy of Sciences to provide advice on scientific and technical aspects of LADB. A Committee on Laboratory Animal Data was formed by ILAR

Figure 10-5.—A Representative Page of the Eight-Page Data Input Form for the Laboratory Animal Data Bank

Animal Group Environment and Husbandry Conditions

FOR INTERNAL LADB USE ONLY

LADB Animal Group Number
Related Animal Group Number[†]

NS = Not Specified NA = Not Applicable

01 Performing Laboratory (where animals were housed)		02 City		03 State	04 Altitude feet
05 Country United States <input type="checkbox"/> Other <input type="checkbox"/>		06 Latitude deg North	07 Latitude deg South	08 Data Donor Organization (if different from performing lab)	
09 City		10 State		11 Country United States <input type="checkbox"/> Other <input type="checkbox"/>	
12 Animal Supplier or Breeder		13 City		14 State	15 Country United States <input type="checkbox"/> Other <input type="checkbox"/>
16 Testing Organization (if different from performing lab or data donor organization)		17 City		18 State	19 Country United States <input type="checkbox"/> Other <input type="checkbox"/>
20 Tests Performed by Testing Organization CBC <input type="checkbox"/> Serum Chemistries <input type="checkbox"/> Bone Marrows <input type="checkbox"/> Histopathology <input type="checkbox"/> Other <input type="checkbox"/>		21 Common Animal Name <input type="checkbox"/> Monkey <input type="checkbox"/> Ape <input type="checkbox"/> Cat <input type="checkbox"/> <input type="checkbox"/> Dog <input type="checkbox"/> Rabbit <input type="checkbox"/> Rat <input type="checkbox"/> <input type="checkbox"/> Mouse <input type="checkbox"/> Minkipg <input type="checkbox"/> Hamster <input type="checkbox"/> Other <input type="checkbox"/>			
22 Species Name		23 Strain, Breed, or Stock Name		24 Substrain Name (if not included in # 23)	

Rodents/Lagomorphs Now go to: Question 25 if Rodent or Lagomorph Question 30 if Carnivore

25 Type of Breeding Outbred <input type="checkbox"/> now go to question (32) Inbred <input type="checkbox"/> now go to question (26) Hybrid <input type="checkbox"/> now go to question (27) Congenic <input type="checkbox"/> now go to question (28) Mutant <input type="checkbox"/> now go to question (29)		26 Type of Inbreeding BxS <input type="checkbox"/> FxD <input type="checkbox"/> MxS <input type="checkbox"/> Gen (F #) _____ NS <input type="checkbox"/> Now go to question 32		27 Hybrid Parentage Strain of Dam _____ Specify Strain of Sire _____ Specify Now go to question 32	
28 Congenic Parentage Background Strain (inbred partner) _____ Specify Donor Strain _____ Specify Now go to question 32		29 Mutant Parentage Strain on which carried _____ Specify BxS <input type="checkbox"/> Backcross <input type="checkbox"/> Gen (F #) _____ Specify (N #) _____ Specify NS <input type="checkbox"/> Now Go To Question 32			

Carnivores

30 Mating System Outbred <input type="checkbox"/> Inbred <input type="checkbox"/> Linebred <input type="checkbox"/>		31 Type of Source Random Source, Unconditioned <input type="checkbox"/> Random Source, Conditioned <input type="checkbox"/> Kennel (Colony) Produced <input type="checkbox"/> Now go to question 32		32 Microbial Status (check only one) Conventional <input type="checkbox"/> now go to question 35 Axenic (germ free) <input type="checkbox"/> Gnotobiotic <input type="checkbox"/> now go to question 33 Defined Microbially Associated <input type="checkbox"/> Barrier Maintained (SPF) <input type="checkbox"/> now go to question 34 Monitored <input type="checkbox"/>	
33 Designate Organisms			34 Tests by Which Monitored		
35 Type of Control Animal Negative (untreated) <input type="checkbox"/> now go to question 42 Sham <input type="checkbox"/> now go to question 37 Vehicle <input type="checkbox"/> thru 41 Pre-experiment * <input type="checkbox"/> now go to question 36 Quarantine <input type="checkbox"/>		36 Duration (days) _____ Now go to question 42		37 Designate Sham or Vehicle Distilled Water <input type="checkbox"/> Tragacanth <input type="checkbox"/> Saline Solution <input type="checkbox"/> Agar <input type="checkbox"/> Gelatin Capsule <input type="checkbox"/> Corn Oil <input type="checkbox"/> Other _____	
38 Route and Site of Administration of Vehicle Route _____ Site _____ Oral (feed) <input type="checkbox"/> N.A. Oral (gavage) <input type="checkbox"/> N.A. Inhalation <input type="checkbox"/> N.A. Topical Application <input type="checkbox"/> _____ Parenteral Injection <input type="checkbox"/> Specify I.M. <input type="checkbox"/> IV <input type="checkbox"/> Specify Subcut <input type="checkbox"/> Specify I.P. <input type="checkbox"/> Specify Other _____ Specify		39 Schedule of Vehicle Administration/Sham Procedure Continuous <input type="checkbox"/> Other _____ Periodic _____ Times/Day for _____ Days/Wks for _____ Wks 40 Concentration of Vehicle (if applicable) (e.g. 0.85% Saline Solution, etc.) _____ 41 Amount of Vehicle Administered Constant Amount _____ ml (or) _____ mg (or) _____ Other Variable Amount (amount per unit body weight or body surface area) _____ mg/kg _____ mg/m ² _____ ml/kg _____ ml/m ² Other _____			

[†]Includes postquarantine and under environmental conditions prior to treatment regimes

SOURCE: P.L. Altman and K.D. Fisher, *Review of Standards Related to the Laboratory Animal Data Bank—Interim Report* (Bethesda, MD: Federation of American Societies for Experimental Biology, 1980),

to advise NLM. It met in 1977 and 1978 to consider reports from the NLM and Battelle staff and to respond both to specific requests for guidance concerning LADB developmental aspects and to feedback from the 13 outside users. The NLM staff further requested ILAR to review the basic concept, purpose, scope, validation of data, and utility of LADB.

In 1978, ILAR prepared such a report and recommended that peer review of data for inclusion in LADB be performed, together with peer examination of the criteria for data acceptability (13). NLM contracted with the Life Sciences Research Office (LSRO) of the Federation of American Societies for Experimental Biology (FASEB) to organize an ad hoc LADB User Assessment Panel to review data descriptors and coverage of various disciplines by LADB.

In 1979, each member of the LSRO ad hoc panel had a computer terminal with unlimited access to the LADB database. The resulting hands-on experience provided the basis for an objective assessment of the data descriptors and the scope of coverage of LADB. The ad hoc panel's report, published in 1980 (2), made some 20 suggestions for improving LADB. The recommendations focused on increasing the data coverage, ensuring the quality of the data included, and facilitating statistical comparisons of data within LADB. The panel's principal recommendations were:

- put the individual animal data files on-line with the grouped animal data files;
- standardize diagnostic terms for pathology data, by using a system such as Systematized Nomenclature of Medicine;
- add new data elements to LADB for growth, development, reproduction, and teratology;
- provide capability for on-line statistical analysis for determining relationships between different data sets; and
- adopt new acceptance criteria for data submitted to LADB.

Public Accessibility, 1980-81

LADB first became available on-line to the public in April 1980, via the Battelle computer in Columbus, OH. Some 100 subscribing organizations logged 96 billable hours over the first 6 months

of public availability. This usage was far lower than that of other databases operated by NLM, even in their beginning stages. The paucity of user hours, coupled with other financial considerations (see following section) led NLM to stop LADB file-building in January 1981. Slightly more than 1 million animal measurements were contained in LADB at this point, mainly obtained from 30,000 rats and mice. A small amount of data came from cats, dogs, hamsters, minipigs, monkeys, and tree shrews (2). Approximately 80 to 85 percent of the total was obtained from investigators holding contracts awarded by the National Cancer Institute, under its cancer treatment and bioassay programs. When the collection of data was halted, 44 organizations and 15 Federal contractors had contributed data, and 9 other sources had agreed to do so. The Federal commitment to LADB from 1975 through 1980 totaled slightly over \$3 million.

Battelle continued to make the file available to the public, but usage did not increase sufficiently to make the project self-sustaining. In early 1982, just 2 years after becoming available on-line, LADB was taken off -line. The file was turned over to the National Technical Information Service for public distribution via licensing. One copy of LADB has been licensed by NTIS to date, to Pergamon International Information Corporation. In 1985, Pergamon, in a joint venture with FASEB, published hard-copy data books created from LADB records (1). There are no plans to add data to the existing file, or to make it publicly available on-line.

Reasons for the Failure of LADB

Financial Considerations

As mentioned, only 96 on-line hours were logged by about 100 LADB users over the first 6 months of public availability. This total was far too small to provide any useful base for self-sufficiency—one of the initial goals for the system. When the Federal Government was vigorously seeking ways to reduce its long-term financial commitments in late 1980, NCI dropped its major financial support in December, and other agencies of the Department of Health and Human Services declined to pick up the slack, under pressure from the Office of Management and Budget to reduce expenses. The LADB contract with Battelle was terminated in early 1981 for lack of funding.

User Friendliness and User's Needs

The interactive software used in LADB was designed in 1975. As such, it predated many major software developments that have emphasized "user friendliness." The users the system was aimed at—biologists—found it hard to retain procedural familiarity with infrequent use.

Another problem with the use of the LADB data, according to the FASEB ad hoc panel (2), was the inability to perform on-line statistical comparisons between different data sets. This limitation, which makes some desirable statistical comparisons difficult to perform, arose from inadequate design and would probably not be a problem with today's software.

User Community

LADB was publicly available for too short a time to permit many conclusions to be drawn about the users. By definition of the content, its users would be expected to be pharmacologists and toxicologists concerned with toxicity testing, particularly chronic toxicity testing. This community, numbering about 3,000 to 5,000 scientists, is far smaller than the community of basic biological scientists (about 200,000). The pool of prospective users of LADB, therefore, seems too small to sustain it.

Peer Review of LADB Design and Data

Although the Institute of Laboratory Animal Resources of the National Academy of Sciences evaluated LADB in 1978, it had not been involved in the original design considerations. Similarly, FASEB entered the review process in 1979—too late to have substantial impact on the design and most of the file-building process. The March 1980 FASEB review (2) pointed out several major design problems, including lack of on-line availability of the individual animal data files. The LADB records that are searchable on-line are composites from groups of animals. Failure to include data on individual animals prevents users from performing statistical comparisons between different data sets.

Lessons Learned From LADB

The acceptance of a biology data bank by the user community and its success in supplying useful research and testing data are actually determined well in advance of the collection or dissemination of data. The first step in assembling a computerized data registry should be the clear definition of its potential users and their specific needs. No adequate study of this nature was performed prior to the original design of LADB. The results of a preliminary feasibility study should identify the various users, their needs, and their desire (or lack thereof) to use and support the proposed database (3).

A 1981 FASEB report, "Guidelines for Development of Biology Data Banks" (4), emphasized three important steps in planning and developing a data bank of biological information. First, the stimulus for establishing a research and testing data bank may be the realization by a scientist, a government agency, or a private organization that the required information is not readily accessible from published, unpublished, or on-line resources. Nevertheless, the need for such an information resource must be determined independently. Most appropriately, this is done by an organization unrelated to the proposing institution. Determination of need involves answering the following questions:

- How many institutional, organizational, or individual users would find the database useful?
- How many would be willing to subscribe, and to what extent would cost be a factor in subscribing?
- How many institutions, organizations, or individual scientists could supply data? How many would?
- How are potential users distributed among disciplines?
- How much unpublished and presently inaccessible data could be made available to investigators by developing a data bank?

If the responses to these questions indicate a solid foundation of perceived need, then the establishment of the data bank is probably justified.

Second, the collection of descriptive data on those scientists interested in the proposed database and on their disciplinary specialties provides a basis for matching the scope of the database to the breadth of disciplinary interest. Specifically, the scope and design of the database depend on the range of purposes for collecting the research and testing data, the size of the prospective audience, and the needs of the users. It is essential to recognize that the needs of any user audience are dynamic and subject to change. A feasibility study should include an analysis of current trends in user application as a basis for inclusion of sufficient flexibility to permit later modification.

In identifying the potential user community, the following considerations are key:

- Can a model be developed to estimate with a high level of accuracy the number of potential users?
- Can a projection of the number of potential donors be made from a similar model?
- To what extent will the user community sup-

port assessment of operational charges to defray costs?

- What will it cost to collect, systematize, store, and retrieve the data for a computerized, on-line system?

Third, critical to the acceptance and success of a registry of research and testing data is peer review by experts, at all levels of database development. These levels include:

- system design;
- definition of data elements;
- establishment of standards for data acceptance;
- compilation and building of data files; and
- post-hoc evaluation of the system (i.e., feedback resulting from experience gained by actual use of the system).

The peer-review process assures that experienced researchers have judged the design, standards, and data to be used. The process enhances quality control, although it imposes the penalty of high costs and slow input of data.

EXPANDING THE LADB CONCEPT: A COMPUTERIZED REGISTRY OF RESEARCH AND TESTING DATA

The concept behind the LADB could be expanded in at least two important dimensions. First, the scope could be broadened beyond baseline results to include experimental results from research and testing. How great an increase in size would this be? For every measurement obtained from a group of control animals, measurements are obtained from an estimated one to nine groups of experimental animals. This makes a registry of control and experimental data from **2 to 10 times** the size of a registry of baseline data alone.

Second, the coverage could be enlarged beyond principally rodents to all vertebrate species. How great an increase would this entail? Several hundred vertebrate species could be involved. The number of species would increase by a factor of more than 100. Yet the bulk of the results would still be derived from rats and mice, since rats and mice account for 12 million to 15 million of the 17 million to 22 million animals used annually in

the United States (see ch. 3). Increasing the scope from rats and mice to all vertebrates would therefore likely enlarge the size of the data registry by **a factor of 1.5** (17 million to 22 million animals divided by 12 million to 15 million rats and mice).

The creation and maintenance of a computerized registry of baseline and experimental results from all species of vertebrate animals would represent an enterprise 3 to 15 times more complex than the unsuccessful Laboratory Animal Data Bank.

The factors that led in the 1970s to the assignment of the LADB project to the NLM remain valid today should a similar project be undertaken. NLM has related experience in handling substance-oriented databases (as detailed in table 10-3), such as the TDB and RTECS. NLM also operates much larger databases, such as TOXLINE and MEDLINE, that are bibliographic rather than substance-oriented.

Other entities that could be considered for operating a centralized registry of research and testing data include:

- National Toxicology Program. NTP never had as its mission the development of a data bank, and it is not presently equipped to do so. The scope of NTP's mission would have to be redefined if it were to undertake this responsibility.
- National Bureau of standards. Although NBS specializes in physical, chemical, and engineering databases, it has never been involved in a biological database operation. NBS does not appear to be a viable candidate.
- National Agricultural Library (NAL). Unlike NLM, NAL has not developed any specialized computerized biological data registry systems. It does not appear to be a viable candidate for operating a centralized registry of research and testing data. The 1985 amendments to the Animal Welfare Act (see ch. 13) directed NAL—in cooperation with NLM—to provide information that could prevent unintended duplication of animal experimentation, and information on improved methods of animal experimentation.
- Chemical Industry Institute of Toxicology. Unless the chemical industry chose to increase funding to CIIT for this express purpose, it could not support this activity. Also, CIIT lacks personnel experienced in large-scale database development and operation,
- Pharmaceutical Manufacturers' Associa-

tion(PMA). PMA is not independent of direction by its members (as is CIIT, for example). Further, PMA is not engaged in large database efforts, making it an unlikely candidate.

- Federation of American Societies for Experimental Biology. FASEB has published handbooks of biological data and is currently embarked on a venture to extract some data from LADB files. However, because of limited resources for data-base development and operation, FASEB's most appropriate role might be as the coordinator of peer-review groups.
- Chemical Abstracts Service of the **American Chemical Society**, and Biosciences Information Services Each of these services annually prepares hundreds of thousands of abstracts that report biological research and testing results. These files are document-oriented and indexed systematically. However, the detail of the abstracts published does not begin to approximate the depth of information found even in LADB. Both services could conceivably undertake the development and operation of a computerized data registry, particularly with NLM supervision.

In summary, it appears that virtually no existing private or public entity, save the NLM, has the resources and expertise to design, develop, and maintain a computerized registry of research and testing data. If NLM were to undertake such a task, it would probably rely on contractors from the private and nonprofit sectors.

SUMMARY AND CONCLUSIONS

The sharing of information on research and testing is vital to scientific progress. There are a variety of ways in which such information can be shared.

Published materials, especially articles appearing in scientific journals, are an indispensable source of information on the results of completed research and testing. Unfortunately, a substantial body of information is not published, although some of it is publicly available.

Publication is a means of establishing a reputation in the scientific community. This is especially important to academics. For scientists in industry, however, the efforts required for publication

compete with other demands on resources, as well as with the need to keep information with some proprietary value confidential.

Much data generated by the Government are published. Yet, when an agency's mission is regulatory, less attention is given to publication than to other concerns.

Because of the importance of journals, their publication policies have a great impact on the kinds of information available. The most troublesome policy is the tendency to publish only results that show an effect. Thus, protocols that yield negative results may be unintentionally duplicated in subsequent experiments.

Federal Government agencies have access to some of the unpublished information held by industry, through reporting rules promulgated under the Toxic Substances Control Act, for example, and through registration requirements of the Federal Insecticide, Fungicide, and Rodenticide Act. This information is used for a variety of regulatory activities and is frequently available to the public under the Freedom of Information Act. It may also be added to databases.

There are several barriers to using available information. One is that users who wish to base important decisions on data need to know how reliable the data are. In assessing reliability, scientists will consider not only the protocol used but also the professional reputation of the scientist, the journal in which the article is published, and where the research or testing was done. If the format of the data (e.g., a numerical database) does not allow the quality to be assessed, the data may have little value. The imprimatur of peer review is an additional factor when assessing data quality.

International barriers to sharing information include language, the delays and expense of communication, the lack of personal acquaintances who could facilitate net working, and political and institutional differences.

Hundreds of thousands of research and testing articles are published each year. Most of these articles and other resources are available through abstracting and indexing services and through bibliographic services. No service, unfortunately, is so comprehensive that it can be relied on as a sole source. However, when multiple sources are used, there can be a great deal of overlap. Another problem is that the summary information may be in-

adequate to judge whether the complete article should be obtained. Citation services also exist for unpublished data and ongoing experiments, some on an international level.

Computers are quite valuable in obtaining access to information. Many of the abstracting and indexing services and bibliographic services are available on computer. Recently, the full text of some scientific journals—except for graphs and images—has become available on-line. In principle, a computer-based registry of research and testing data could reduce the use of animals in research and testing. In practice, the best design of such a computerized database remains uncertain.

One attempt toward a modest, well-defined data registry, the Laboratory Animal Data Bank, failed. Any new effort to establish a comprehensive database that includes descriptions of experimental protocols, control and experimental results, and peer review will benefit from the lessons learned from LADB. The creation and maintenance of a computerized registry of baseline and experimental results from all species of vertebrate animals would be 3 to 15 times more complex than the defunct LADB.

The initial step towards assembling a computerized data registry is the clear definition of both its potential users and their specific needs. The acceptance of a new biological data bank by the user community and the registry's success in supplying useful research and testing data are closely linked to how well the databank meets user needs. Thus, the probable success or failure of a new data bank can be predicted in advance of the collection or dissemination of the data,

CHAPTER 10 REFERENCES

1. Altman, P.L. (ed.), *Pathology of Laboratory Mice and Rats* (New York: Pergamon Press, 1985).
2. Altman, P.L., and Fisher, K.D., *Review of Data Standards Related to the Laboratory Animal Data Bank — Interim Report* (Bethesda, MD: Federation of American Societies for Experimental Biology, 1980).
3. Altman, P. L., and Fisher, K. D., *Assessment of the Laboratory Animal Data Bank in Meeting Needs of*
4. Altman, P. L., and Fisher, K. D., *Guidelines for Development of Biology Data Banks* (Bethesda, MD: Federation of American Societies for Experimental Biology, 1981).
5. Altman, P. L., and Fisher, K. D., *A User Assessment of the Toxicology Data Bank* (Bethesda, MD: Feder-

Users (Bethesda, MD: Federation of American Societies for Experimental Biology, 1980).

- ation of American Societies for Experimental Biology, 1982).
6. BIOSIS (Biosciences Information Service), "The 1985 BIOSIS Catalog," Philadelphia, PA, 1985.
 7. Cohen, M. F., and Flagle, C. D., "Full-Text Medical Literature Retrieval by Computer" *JAMA* 254:2768-2774, 1985.
 8. Cooper, M., "Secondary Information Services in Science and Technology: A Wide-Angle View," *J. Am. Soc. Info. Sci.* 33:152, 1982.
 9. Cummings, M. M., Consultant, Council on Library Resources, Inc., Washington, DC, personal communication, April 1985.
 10. Davidson, J., "The Scope of Biological Testing Today," paper presented at the Annual Meeting of the Society for Information Science, Annual Meeting, Philadelphia, PA, October 1984.
 11. Fang, J., Tracer Jitco, Inc., Rockville, MD, personal communication, June 1984.
 12. Fox, J. L., "News and Comment" *Science* 226:816, 1984.
 13. Institute of Laboratory Animal Resources, Committee on Laboratory Animal Data, *Review and Evaluation of the National Library of Medicine Laboratory Animal Data Bank* (Washington, DC: National Academy Press, 1979).
 14. Loosjes, Th.P., *On Documentation of Scientific Literature* (London: Archon Books, 1967).
 15. Meyer, D. E., Mehlman, D. W., Reeves, E. S., et al., "Comparison Study of Overlap Among 21 Scientific Databases in Searching Pesticide Information," *Online Review*, 7(1):33, 1983.
 16. National Bureau of Standards, Workshop on Data Quality Indicators, Gaithersburg, MD, Feb. 10-12, 1982.
 17. Organization for Economic Cooperation and Development, "Decision on the Mutual Acceptance of Data in the Assessment of Chemicals," [C(81)30 (Final)], addenda 1 to 4, adopted May 12, 1981.
 18. Organization for Economic Cooperation and Development, "Improved Access to Unpublished Information," Report of the Expert Group (Paris: 1984).
 19. U.S. Department of Health and Human Services, National Institute for Occupational Safety and Health, *Registry of Toxic Effects of Chemical Substances*, R.L. Tatken and R.J. Lewis, Sr. (eds.), DHHS (NIOSH) Pub. No. 83-107 (Cincinnati, OH: NIOSH, June 1983).
 20. U.S. Executive Office of the President, Council on Environmental Quality, Toxic Substances Strategy Committee, *Toxic Chemicals and Public Protection* (Washington, DC: U.S. Government Printing Office, May 1980).