

ANNUAL REPORT

FISCAL YEAR 2003

DIVISION OF SPECIALIZED INFORMATION SERVICES

NATIONAL LIBRARY OF MEDICINE

SPECIALIZED INFORMATION SERVICES

Jack Snyder, M.D., J.D., Ph.D.
Associate Director

The Toxicology and Environmental Health Information Program (TEHIP), known originally as the Toxicology Information Program, was established 35 years ago within the NLM Division of Specialized Information Services (SIS). Over the years TEHIP has provided for the increasing need for toxicological and environmental health information by taking advantage of new computer and communication technologies to provide more rapid and effective access to a wider audience. We have moved beyond the bounds of the physical NLM, exploring ways to point and link users to relevant sources of toxicological and environmental health information wherever these sources may reside. Resources include chemical and environmental health databases and Web-based information resource collections. Development of HIV/AIDS information resources became a focus of the Division several years ago, and now includes several collaborative efforts in information resource development and deployment, including a focus on the information needs of other special populations.

The SIS Web server provides a central point of access for the varied programs, activities, and services of the Division. Through this server (<http://sis.nlm.nih.gov>), users can access interactive retrieval services in toxicology and environmental health, HIV/AIDS information, or special population health information; find program descriptions and documentation; or be connected to outside related sources. Continuous refinements and additions to our Web-based systems are made to allow easy access to the wide range of information collected by this Division. Our usage has continued to increase over the past year with access to all toxicology and HIV/AIDS data free over the Internet.

In FY2003 SIS continued to balance efforts to enhance and re-engineer existing information resources with efforts to provide new services in emerging areas. We further developed various prototypes that rely on geographical information systems, innovative access and interfaces for consumers, and graphical display of data from information sources. Highlights for 2003 include

- SSEUS (SIS SQL Entry Update System), a new data creation and maintenance system for TOXNET, our premier collection of databases on toxicology, hazardous chemicals, and related areas;
- Listserv, Automated Indexing, and other refinements of the interface and multi-database search capability for TOXNET;
- Household Products Database, which links the ingredients in over 4,000 consumer brands to

health effects described in Material Safety Data Sheets provided by manufacturers;

- Several new Toxicology & Environmental Health Special Topic Web resource pages, including Environmental Justice and Asian-American Health;
- TOXMAP, a prototype system that uses maps of the United States to help users visually view data about chemicals released into the environment and easily connect to related environmental health information;
- ToxTown, a graphical portal to chemicals encountered in everyday life, in everyday places;
- Continued support of PAHO/NLM Disaster Preparedness Information Centers in Honduras, Nicaragua, and El Salvador;
- Expanded Native American outreach initiatives; and
- Minority outreach activities with the Historic Black Colleges and Universities, United Negro College Fund Special Projects, and the National Medical Association.

Resource Building

The wide range of SIS resources related to toxicology and environmental health information, HIV/AIDS information, and special populations information includes many databases that are created or acquired as well as other services and projects.

Haz-Map database, released in 2002 at <http://hazmap.nlm.nih.gov>, is an occupational toxicology database designed to link jobs and hazardous job tasks to occupational diseases and their symptoms. It is a relational database of chemicals, jobs, and diseases that averaged nearly 20,000 queries per month in 2003. The *Haz-Map jobs table* is based on the 1997 Standard Occupational Classification (SOC) system. The *industries table* is based on the Standard Industrial Classification (SIC) system. The *diseases table* is based on the International Classification of Diseases (ICD-9). Information from textbooks, journal articles, and electronic databases was classified and summarized to create the database. A user may search this occupational database by chemical agent, occupational disease and by job type.

ChemIDplus (Chemical Identification File) is an NLM online chemical dictionary that contains nearly 370,000 records, primarily describing chemicals of biomedical and regulatory importance, and available to users on the Internet at <http://chem.sis.nlm.nih.gov/chemidplus>. ChemIDplus features include chemical structure search and display for over 177,000 chemicals, and hyperlinked locator fields that retrieve data for a given chemical from other resources such as TOXLINE, MEDLINE or HSDB

as well as EPA and ATSDR. Over 15,000 records of regulatory interest collectively known as SUPERLIST are also available and hyperlinked in ChemIDplus. During FY2003 over 75,000 queries per month were made of this database. To assist with spelling errors, a chemical spell checker was released to help users retrieve substances more efficiently by chemical name. The checker, which can be instantly revised using the SIS DBMaint2 online update system, contains new spelling indices for more than 1.3 million chemical names and synonyms. The database was enhanced by the addition of various new locators pointing to international resources, including coverage of agents found in NIAID ChemDB, ATSDR Medical Management Guidelines, and ClinicalTrials.gov. In FY2003, new test versions for the new ChemIDplus "Light" and "Heavy" systems were ported to a UNIX-based chemserver for testing and further development. The new system capabilities include a simpler Web front end that does not require plug-ins for structure display, and an advanced version that allows numeric searching by acute toxicity data and effect, and chemical/physical properties.

The Hazardous Substances Data Bank (HSDB) continues to be a highly used resource, averaging 50,000–60,000 searches each month (a 5% increase over FY2002). Increased emphasis continues to be placed on providing more data on human toxicology and clinical medicine within HSDB, in keeping with past recommendations of the Board of Regents' Subcommittee on TEHIP. In 2003, there has also been a continued emphasis on adding to HSDB new chemicals with the potential for high toxicity and high human exposure. Over 135 new chemicals were added in 2003, including new pesticides, drugs, and environmental pollutants. The emphasis on the addition of new chemicals will continue in the coming year. Newer sources of relevant data are being examined for incorporation into new and existing data fields within the current 4,757 HSDB records. Because of increased staff efforts, more records are being processed through special enhancements, including source updates from various peer-reviewed files. Special summary information is being prepared to allow easier presentation of information at a health consumer level. The process of developing a new Web-based system for HSDB creation, review, and maintenance is continuing. As part of this effort, SSEUS (a relational HSDB database using the MySQL database application) was created, and a new client-server interface was programmed to allow easier updates. The new maintenance system is now poised for integration with other new features, including numeric searching and automatic indexing.

The Toxicology Data Network (TOXNET), NLM's information system providing database management for many of its toxicology files, has moved from a networked microprocessor environment to a UNIX-based platform (Solaris Version 2.6) on a SUN Enterprise 3000

computer. SIS continues to integrate this configuration with other database creation systems and Web access to them. Further refinements of the SIS search interface enhance the ability of users to simultaneously search HSDB, TOXLINE, CCRIS, Gene-Tox, DART/ETIC, IRIS, TRI and ChemIDplus from one input screen. Based on recommendations from the Institute of Medicine, users are presented with a basic search screen with just a single input box for searching, with customized screens for more sophisticated users. These advanced features include Boolean searching and the ability to limit search terms to specific fields. Feedback from TOXNET user online surveys has provided a basis for current and future planning, and as result, SIS will implement a chemical spellchecker, automated indexing, and a virtual meta-search tool during the coming years.

Alternatives to Animal Testing (ALTBIB)—SIS continues to compile and publish references from the MEDLARS files that were identified as relevant to methods or procedures that could be used to reduce, refine, or replace animals in biomedical research and toxicological testing. Staff members search, edit, and categorize citations to create a true value-added resource in this field. The 22 bibliographies issued during the past 10 years are available on the Internet through the SIS Web server, and the primary distribution mechanism for this project is now the Internet, through a new online resource named ALTBIB, which allows search access to all of the 7,595 citations organized from previous bibliographies. ALTBIB uses the TOXNET search engine, and is available at <http://toxnet.nlm.nih.gov/altbib.html>. A user may search by keyword, author, or one of the 16 subdivisions such as "Quantitative Structure Activity Studies."

TOXLINE (Toxicology Information Online) is a large NLM bibliographic database traditionally produced by merging "toxicology" subsets from secondary sources. By the end of FY2003, the database included over 3 million citations to toxicology literature dating back to 1965. In 2003, users accessed standard journal literature in toxicology and environmental health as part of the enlarging MEDLINE database, while NLM continued to add journals in the area of toxicology and environmental health to MEDLINE to cover some of the literature formerly provided by outside sources. For the non-standard journal literature in this area, SIS further enhanced a Web-based system on TOXNET that allows efficient acquisition and updating of these components. Easy access to this TOXLINE Special database and to TOXLINE Core, the standard journal literature on PubMed, is available from the improved TOXNET user interface.

DIRLINE (Directory of Information Resources Online) is NLM online directory of resources including organizations, databases, bulletin boards, as well as

projects and programs with special biomedical subject focus. These resources provide information to users which may not be available from one of the other NLM bibliographic or factual databases. DIRLINE continues to receive a high level of use (nearly 7000 searches per month). The interface supports direct links to the Web sites of the organizations listed in the database, as well as direct e-mail connections. The quality and utility of the database continues to improve as duplicates have been eliminated through changes in policy and streamlining of maintenance. *Health Hotlines*, the always popular publication of health-related toll-free telephone numbers, has a recently updated Web version which also indicates the availability of Spanish speaking customer service representatives and Spanish language publications from the resources listed.

The **Toxics Release Inventory (TRI)** series of files now includes online files TRI86 through TRI2001. These files remain an important resource for environmental release data and are a useful complement to other SIS databases. Mandated by the Emergency Planning and Community Right-to-Know Act (Title III of the Superfund Amendments and Reauthorization Act of 1986), these EPA databases contain environmental release data for air, water, and soil for over 600 EPA-specified chemicals. These files are used in the new SIS R&D project using a geographical information system, TOXMAP.

The **Chemical Carcinogenesis Research Information System (CCRIS)** continues to be built, maintained, and made publicly accessible at NLM. This data bank is supported by the National Cancer Institute (NCI) and has grown to over 8,000 records. The chemical-specific data covers the areas of carcinogenesis, mutagenesis, tumor promotion, and tumor inhibition.

The **Integrated Risk Information System (IRIS)**, EPA's official health risk assessment file, continues to experience high usage and be very popular with the user community. EPA has had a version of IRIS on the agency's Web page since 1996, and we will continue to consider how best to integrate our Web service with what EPA provides. IRIS now contains 540 chemicals.

The **GENE-TOX** file is built directly on TOXNET by EPA scientific staff. This file contains peer-reviewed genetic toxicology (mutagenicity) studies for about 3,200 chemicals. GENE-TOX receives a high level of interest among users in other countries.

The **Registry of Toxic Effects of Chemical Substances (RTECS)** is a data bank based upon a National Institute for Occupational Safety and Health (NIOSH) file by the same name which NLM restructured and made available for online searching. With our move to free Internet access to all databases, NIOSH requested that we no longer include RTECS on our system. SIS continues to use RTECS in the creation of the Hazardous Substance

Data Bank.

The **Developmental and Reproductive Toxicology (DART)** database now contains over 240,000 citations from literature published since 1989 on agents that may cause birth defects. DART is a continuation of the Environmental Teratology Information Center backfile (ETICBACK) database. In FY2003, next generation DART consisted of two subsets: DART Core on PubMed, containing over 170,000 citations to the journal literature, and DART Special, containing nearly 70,000 citations to specialized resources (including meeting abstracts, books, technical reports) in this subject area. In FY2003, a new contract was awarded, thousands of new records were added, and easy access to DART Special and to DART Core was maintained at the new TOXNET interface. DART is funded by NLM, the Environmental Protection Agency, the National Institute of Environmental Health Sciences, and the FDA's National Center for Toxicological Research, and is managed by NLM.

The **Environmental Mutagen Information Center (EMIC)** database contains over 24,000 citations to literature on agents that have been tested for genotoxic activity. A backfile for EMIC (EMICBACK) contains over 75,000 citations to the literature published from 1950-1991. The Environmental Protection Agency, the National Institute of Environmental Health Sciences and NLM, collaborating partners in this effort, stopped compiling this special collection as of December 1999, but SIS will keep the collections as part of the TOXLINE Special database on TOXNET.

On March 21, 2002, SIS sponsored a Children's Environmental Health Information Resources Satellite Broadcast via the CDC Public Health Training Network. The program demonstrated selected online resources in the context of important children's environmental health issues. Topics included exposure of children to pesticides, environmental triggers of childhood asthma, methylmercury and fish contamination, the use of Geographic Information Systems for environmental health data, Healthy People 2010 resources, and lead poisoning prevention funding resources. The program was designed for physicians, nurses, physician assistants, nurse practitioners, epidemiologists, public health educators, librarians, counselors, administrators, or anyone else providing environmental health-related services. A Webcast is available at on the CDC Web site under public health training network. In addition, a children's environmental health resource sampler was developed and put on the National Network of Libraries of Medicine Web site.

AIDS Information Services

NLM has continued its successful AIDS Community Information Outreach Program with 15 new awards in FY2003, bringing the total number of awards made to 172. In addition to these awards, NLM continues to work with other organizations to raise awareness of HIV/AIDS information resources among small community organizations at a grassroots level. A new "HIV/AIDS and Older Adults" webpage is at <http://sis.nlm.nih.gov/HIV/HIVOlderAdults.html>.

NLM remains as the project manager for the multi-agency AIDS Clinical Trials Information Service (ACTIS) and the HIV/AIDS Treatment Information Service (ATIS), which now have been merged into a new service entitled "AIDSinfo." This new service will continue to provide access to AIDS-related clinical trials information (through Clinicaltrials.gov) and federally approved treatment guidelines. The contract for this service also provides support services for Clinicaltrials.gov.

Outreach / User Support

Special Population Web Sites: The Arctic Health web site (<http://arctichealth.nlm.nih.gov>), initially developed by SIS staff, is now updated by the University of Alaska at Anchorage; the newly released Asian-American Health web site was announced by HHS Secretary Thompson during the celebration of Asian and Pacific Islanders Health Month, and the forthcoming American Indian Health Web site is now undergoing review. These Web sites include relevant policy, legislative, and organizational information as well as organized links to health and environmental issues of concern to the designated population.

NLM-Tox-Enviro-Health-L listserv was created in June 2003 to send announcements about SIS toxicology and environmental health resources. Messages sent to the nearly 500 subscribers include lists of new chemicals added to Hazardous Substances Databank, announcements about the new Household Products Database, and new environmental health topics for consumers added to Tox Town or MedlinePlus.

In FY2003, the Toxicology Information Outreach Panel (TIOP) evolved a new strategic plan and was renamed the Environmental Health Information Outreach Panel (EnHIOP). Dr. Henry Lewis, Dean of the School of Pharmacy at Florida A&M University, became Chair of the new group. The new EnHIOP includes representation from additional Historically Black Colleges and Universities (HBCU's) as well as from Tribal Colleges and Hispanic Serving Educational Institutions. The panel will address a broader spectrum of environmental health issues in the coming years.

SIS continued its health information training programs at national and regional meetings of the National Medical Association. These programs cover all NLM online resources, including TOXNET, PubMed, ClinicalTrials.gov, and Medlineplus.

A more recent addition to NLM's outreach programs is one to improve access to health-related disaster information in three disaster-prone Central American countries: Nicaragua, Honduras, and El Salvador. In FY2003, SIS continued its support of the Regional Disaster Information Center for Latin America and the Caribbean to strengthen the capacity of these countries to collect, index, manage, store, and disseminate public health and medical information related to disasters. The objective of this project is to contribute to disaster reduction by capacity building activities in the area of disaster-related information management. Selected libraries and information centers have been provided with the knowledge, training and technology resources in order to act as reliable information providers to health professionals and others in their countries. Through this initiative, the participating libraries and information centers have been strengthened in several areas:

- Technological infrastructure (Internet connectivity and computer equipment)
- Information Management (health science librarian training)
- Information Product Development (digital library, Web sites)

This project is also assisting SIS in developing models for collecting and exchanging health information in geographically isolated and disaster-prone environments and for handling non-traditional or unpublished literature, in this case on the health aspects of disasters.

SIS exhibited at over 40 conferences in this fiscal year. Several of these provided opportunities for presentations or workshops about NLM information resources. In addition, NLM-SIS hosted the National Congress of American Indians President's Task Force on Health Information, and a national conference on Refugee Health Information. NLM also sponsored the e-health track at the Technology Partnerships Conference (formerly the HBCU/MI Technology Expo) held at the Georgia Centers for Advanced Telecommunications and Technology, Atlanta Georgia.

Research and Development Initiatives

To meet the mission of providing information on toxicology, environmental health, and targeted biomedical topics to the world, SIS has been developing new ways of presenting the world of hazardous chemicals in our environment to a wider audience. For example:

The **Household Products Ingredients Database** (<http://householdproducts.nlm.nih.gov>) provides a Web resource for consumers that links brand name household products (more than 4,000) with their ingredient chemicals (more than 2,000) and potential adverse health effects. Information derived from manufacturer's Material Safety Data Sheets and from SIS databases can provide answers to various questions, including: a) what

chemicals are contained in specific brands and in what percentage; b) which products contain specified chemicals; c) who manufactures a specific brand and how can that manufacturer be contacted; d) what are the potential acute and chronic health effects of the chemical ingredients found in a specific brand; and e) what other information is available about such chemicals in the toxicology-related databases of the NLM.

The **ToxTown** (<http://toxtown.nlm.nih.gov>) project explores how best to provide environmental health information to a general audience. **ToxTown** is an interactive guide to commonly encountered toxic substances, your health, and the environment. It uses color, graphics, sounds and animation to convey connections between chemicals, the environment, and the public's health. Tox Town is designed to provide:

- Facts on everyday locations where toxic chemicals might be found;
- Information about how the environment can affect human health;
- Non-technical descriptions of chemicals;
- Links to authoritative chemical information on the Internet; and
- Internet resources on environmental health topics.

Tox Town helps users explore an ordinary town or city or farm to identify its common environmental hazards. The city, town, or farm can be toured by selecting Location or Chemical links. Locations, like the school, home or office building, can be opened for cutaway views and for detailed information about potentially hazardous chemicals that might be found there, as well as for links to environmental health resources. Tox Town also offers some resources in Spanish.

In FY2003, SIS began beta-testing of **TOXMAP**, a prototype system that uses maps of the US to help users visualize data about chemicals released into the environment. TOXMAP integrates data from the EPA's Toxic Release Inventory with information about health effects, research citations, etc. found in TOXNET databases. Users can create nationwide or local area maps that show where chemicals are released into the air, water, and ground. TOXMAP also integrates data from other sources, such as demographic data from Census Bureau. TOXMAP provides region-specific links to chemical and bibliographic information.

WISER (Wireless Information System for Emergency Responders) is designed to provide critical chemical information quickly and conveniently on a PDA for use by emergency responders (first 24 hours in hot-zone). The application is being developed in partnership with the Agency for Toxic Substances and Disease Registry, using ATSDR Medical Management Guidelines for Acute Chemical Exposures, which were developed to aid emergency department physicians and other

emergency health care professionals who manage acute exposure following chemical incidents. The WISER prototype has focused on approximately 400 agents found in the Hazardous Substances Data Bank, and current deployment plans include a user's guide, a tutorial, evaluation methodology, and "in-field" testing.

ToxSeek provides a virtual meta-search tool for simultaneous searching of target information systems, displaying search results from targeted systems, and harvesting related concepts. This tool can be configured to define a set of target information/search tools, which for SIS are T&EH databases and searchable resources on the web. Testing of the prototype is under way and a beta version will be ready for public release in FY2004.

Other Interagency Initiatives

SIS personnel continued their leadership of the Interagency Tox-to-Consumer Workshop, which convened in June 2003 to finalize plans for an Inventory of Federal Government Consumer Environmental Health Resources.

Evaluation Activities

In FY2003, SIS completed ten Reviews of PDA (Personal Digital Assistant) Applications in Toxicology and Environmental Health. Each review typically covers the following topics: General Information, Intended Users, Data Source/Authorship, Contents, Navigation, System Requirements, License Type/Price, Availability, and Useful Web Links.

With support from an NIH Evaluation Express Award, SIS established online professional and consumer focus groups to evaluate various information products, including HAZMAP, Household Products, and ToxTown.

The NIH Office of Evaluation also provided set-aside funds to NLM for the evaluation of TOXNET, AIDSinfo, MedlinePlus, and the NLM Web site using the American Customer Satisfaction Index.

In these and other new initiatives, SIS continues to search for new ways to be responsive to user needs in acquiring and using toxicology and environmental health, HIV/AIDS, and other specialized information resources.