MANAGING ETHNOPHARMACOLOGICAL DATA: HERBARIA, RELATIONAL DATABASES, LITERATURE

J. R. Stepp
Department of Anthropology, University of Florida, USA
M. B. Thomas
Department of Botany, University of Hawai'i, USA

Keywords: database technology, relational databases, herabaria, standards, meta-data.

Contents

Summary

The proper management of ethnopharmacological data is crucial to further progress in the field. However, the interdisciplinary nature of the field can lead to difficulties in standardizing data. While scientific collections in the field date back many centuries, data management is rapidly changing as a result of what has been described as the quiet revolution of information technology and bioinformatics. Today, an increasing amount of data from ethnopharmacological research is being stored and curated electronically. The internet has opened up access to vast amounts of ethnopharmacological data within numerous databases.

Many herbaria are making images of voucher specimens accessible online. This is significant because properly collected voucher specimens and their associated data are of primary importance to ethnopharmacology. New database models are proposed that will incorporate phylogenetic, ecological, and ethnobotanical sampling approaches in data collection. Work is underway to create standardized metadata for ethnopharmacology. A model for a database system that provides a standardized environment for submission, storage, and retrieval of ethnomedicinal data has been developed. The model is based on object-oriented database technology, and is suitable for not only storing data, digital images, sound and video, but also for modeling domain knowledge associated with plant-based medicinal preparations utilized in systems of traditional medicine. The model incorporates both linguistic and semantic elements. These advances in management of ethnopharmacological data should lead to more comparative approaches and further development of theory.

1. Introduction

The management of ethnopharmacological data is a complex issue due to the interdisciplinary nature of the discipline. Research questions and methods vary widely depending on the training and background of the researcher(s). This can lead to difficulties in standardizing data and engaging in comparative approaches. The field of ethnopharmacology involves people's use of plants, fungi, animals, microorganisms and minerals, in the context of traditional medical systems. It is concerned with identifying the biological and pharmacological effects of these *materia medica*, and communicating that information based on the principles established through international conventions. Early humans confronted with illness and disease experimented and discovered a wealth of useful therapeutic agents in both the animal and plant kingdoms. The empirical knowledge of these medicinal substances and their toxic potential was passed on by oral tradition and sometimes recorded in herbals and other texts on *materia medica*.

Today, at least 121 plant derived pharmaceutical drugs including belladona alkaloids (e.g. atropine, hyoscyamine, and scopolamine), digoxin, cocaine, the opiates (codeine and morphine), tubocurarine, digoxin, reserpine, taxol, tubocurarine, quinine and reserpine have been discovered and commercialized through the study of traditional remedies. Natural product investigation based on ethnopharmacological leads continues, although the resurgence of interest in ethno-bioprospecting in the late 1980s and 1990s has declined lately, mainly due to the lack of new pharmaceuticals being created through such approaches. Of perhaps more current significance is the use of ethnopharmacological leads as a backbone in combinatorial chemistry to create new pharmaceuticals. Chemists continue to use plant-derived compounds (e.g. emetine, morphine, taxol, physostigmine, quinidine) as prototypes in their attempts to develop more effective and less toxic medicines.

In recent years, the preservation of local knowledge, the promotion of indigenous medical systems in primary health care, and the conservation of bio-cultural diversity have become vital issues to all scientists working at the interface of social and natural sciences but especially to ethnopharmacologists. Recognizing the sovereign rights of States over their natural resources, ethnopharmacologists are particularly concerned with local people's traditional rights to further utilization and development of their traditional knowledge and autochthonous resources. In view of that, today's ethnopharmacological research embraces multidisciplinary effort in the documentation of indigenous medical knowledge and scientific study of indigenous medicines in order to contribute in the long-term to improved health care in the communities of study as well as searching for pharmacologically unique principles from existing indigenous remedies.

Ethnopharmacological data management is similar to that of the broader field of ethnobiology, in that researchers rely heavily on voucher specimens of organisms and associated data collected with the voucher in order as the primary unit of data. However, ethnopharmacological data management has the added complexity of data related to medical practices and, sometimes, chemical analyses of the voucher specimens. Recent advances in database technology hold promise for the organization and analyses of complex data.

2. Historical Trends

Scientific collections and data in ethnopharmacology date back many centuries. As early as the thirteenth century humans started to systematically collect voucher specimens of plants and store them in collections which evolved into herbaria. The development of herbaria coincided with the development of the printing press and subsequently, libraries. In fact, herbaria had much in common with early libraries because plants were dried and pressed and then glued into blank pages of books. Some of the earliest herbaria are found at at the Naturkundemuseum in Kassel, Germany (1569) and at the universities of Bologna (1570), Basel (1588), and Oxford (1621).

Later herbaria starting around the eighteenth century began using specimens mounted on loose sheets of paper. This allowed for the shuffling of collections as classifications schemes matured and changed. The publication of *Species Plantarum* by Carl Linnaeus in 1753 revolutionized classification by providing a global system for the classification of organisms. Through the development of a binomial nomenclature, Linnaeus standardized the way herbaria cataloged collections.

Because herbarium collections are spread across many institutions and geographic regions, they have been historically difficult to utilize and compare. Retrieving ethnopharmacological research data has always been challenging because much of this information is published in disciplinary journals from diverse fields (e.g. botany, biology, anthropology, conservation biology) or in "gray literature" (e.g. unpublished works, government documents and technical reports) that are not widely accessible. The diffuse distribution and variable quality of this data limit the ability of scientists to easily obtain access to either legacy data or current published ethnobotanical research. Large poorly maintained databases, some of which are proprietary, also exist in institutions but are not accessible to the wider scientific community.

A decade or two ago, a researcher interested in learning about the medical use of a particular plant species may have had to scan index cards, field notes and personal journals in order to find what had been collected, where and for what use. To learn what was contained in herbaria and museum collections a researcher had to physically visit them. This situation, however, is rapidly changing as a result of what has been described as the quiet revolution of information technology and bioinformatics. Today, an increasing amount of data from ethnopharmacological research is being stored and curated electronically.

3. Present Trends

3.1. Ethnopharmacological Databases

Probably, the largest collection of ethnopharmacological data is held by the Program for Collaborative Research in the Pharmaceutical Sciences College of Pharmacy, University of Illinois at Chicago NAPRALERT database. An acronym for NAtural PRoducts ALERT, NAPRALERT is the largest relational database of world literature describing the ethnomedical or traditional uses, chemistry, and pharmacology of plant, microbial and animal (primarily marine) extracts. While at its core the database is concerned with

coverage of natural products, whether used by humans or not, a substantial portion of the database is within the realm of ethnopharmacology. NAPRALERT is a relational database that is fee-based, except for researchers from Third World Countries whose access is free of charge. It contains bibliographic and factual data on natural products, including information on the pharmacology, biological activity, taxonomic distribution, ethno-medicine and chemistry of plant, microbial, and animal (including marine) extracts (Figure 1). In addition, the file contains data on the chemistry and pharmacology of secondary metabolites that are derived from natural sources and that have known structure. It is a "source" type of database as opposed to the more common type known as a "bibliographic resource", which only contains citation information.

NAPRALERT currently contains the extracted information from over 170,000 scientific research articles. The NAPRALERT File contains records from 1650 to the present. Approximately 80% of the file is from systematic survey of the literature from 1975 to the present. The remaining records were obtained by selective retrospective indexing dating back to 1650. Over 151 000 plant, animal, marine and microbial organisms are covered with over 1.5 million entries on their biological activities. The database contains four main record sets for each entry: demographic (similar to that of a standard bibliographic file but with select additional information added); organism (full taxonomic description, part of organism studied); compound (when available, information is provided on the chemical composition and percentages of the organism); and pharmacology (biological activities and effects). The database is kept current by a team of scientists who systematically review relevant literature and extract pertinent information into the database. Over 600 articles are added each month from over 700 journals. Comprehensive abstracting services are also utilized and scanned for relevant articles. The secondary literature indices that are incorporated into NAPRALERT include Index Catalog of the Surgeon General (1880-1961), Index Medicus (1897-1927; 1960 to present), Chemical Abstracts (1907 to present), Quaterly Cumulative Index Medicus (1916-1956), Biological Abstracts (1926 to present), Current List of Medical Literature (1941-1959), United States Armed Forces Medical Journal (1950-1960), National Library of Medicine Current Catalog (1966 to present), and Current Contents—Life Sciences (1967 to present). Over particular utility in the database are the numerous search functions available. Some search fields available that would be of particular interest in ethnopharmacological research are:

- verify, where a search is conducted to determine if a Latin binomial is valid within the database or if a synonym is required;
- common—this feature provides all common (ethnobiological) names for a particular Latin binomial of an organism;
- ethno—which provides information on traditional uses of a particular genus or species;
- exper or biol—information on all experimental biological testing on a particular extract or compound from an organism;
- cmpd—a list of all compounds in a genus or species
- occ—a list of all organisms from which a compound has been identified.

DISPLAY ALL

```
AN   92:17094  NAPRALERT
DN   H06008
TI   (+)-ALPHA-VINIFERIN, AN ANTI-INFLAMMATORY COMPOUND FROM CARAGANA
     CHAMLAGU ROOT
AU   KITANAKA S; IKEZAWA T; YASUKAWA K; YAMANOUCHI S; TAKIDO M; SUNG H
K;
     KIM I H
CS   COLL PHARM, NIHON UNIV, CHIBA 274 JAPAN
SO   CHEM PHARM BULL (1990) 38 (2) p. 432-435.
DT   (Research paper)
LA   ENGLISH
CHC  1424
ORGN Class: DICOT  Family: LEGUMINOSAE  Genus: CARAGANA  Species:
      CHAMLAGU [LAMK.]
      Organism part: DRIED ROOT
      Geographic area (GT): SOUTH KOREA; EAS
      TYPE OF STUDY (STY): IN VIVO.  Classification (CC):
          ANTIINFLAMMATORY ACTIVITY
          Extract type: ETHER EXT
          Dosage Information: INTRAGASTRIC; MOUSE; DOSE: 300.0 MG per KG
          Qualitative results: ACTIVE
          Comment(s): VS.CARRAGEENIN-INDUCED PEDAL EDEMA..
      TYPE OF STUDY (STY): FOLKLORE.  Classification (CC): ANALGESIC
          ACTIVITY
          Extract type: HOT H2O EXT
          Dosage Information: ORAL; HUMAN ADULT
          Comment(s): USED AS AN ANTI-NEURALGIC..
      TYPE OF STUDY (STY): FOLKLORE.  Classification (CC):
          ANTIINFLAMMATORY ACTIVITY
          Extract type: HOT H2O EXT
          Dosage Information: ORAL; HUMAN ADULT
          Comment(s): USED AS AN ANTI-RHEUMATIC AND AN ANTI-ARTHRITIC..
      TYPE OF STUDY (STY): ISOLATION.
          COMPOUND. Chemical name (CN): VINIFERIN,ALPHA: (+)
               CAS Registry Number (RN): 62218-13-7
               Class identifier (CI): OXYGEN HETEROCYCLE
               Yield: 00.00039%
      TYPE OF STUDY (STY): IN VIVO.  Classification (CC):
          ANTIINFLAMMATORY ACTIVITY
          Dosage Information: INTRAGASTRIC; MOUSE; DOSE: 10.0 MG per KG
          Qualitative results: ACTIVE
          Comment(s): VS.CARRAGEENIN-INDUCED PEDAL EDEMA..
          COMPOUND. Chemical name (CN): VINIFERIN,ALPHA: (+)
               CAS Registry Number (RN): 62218-13-7
               Class identifier (CI): OXYGEN HETEROCYCLE
      TYPE OF STUDY (STY): ISOLATION.
          COMPOUND. Chemical name (CN): GLYCEROL-ALPHA-LIGNOCERATE
               Class identifier (CI): LIPID
          COMPOUND. Chemical name (CN): GLYCEROL-ALPHA-CEROTATE
               Class identifier (CI): LIPID
          COMPOUND. Chemical name (CN): GLYCEROL-ALPHA-MONTANATE
               Class identifier (CI): LIPID
```

Figure 1. NAPRALERT Search Example

The Internet provides a range of other databases based on either primary research or secondary research, such as literature searches. Most of these tend to not be global in extent like NAPRALERT but tend to focus on particular socio-linguistic groups or regions. Of particular note is the database developed by Daniel Moerman at the University of Michigan-Dearborn on Native American Ethnobotany (http://herb.umd.umich.edu). The database contains information on plant derived foods, drugs, dyes and fibers of Native American Peoples. As a medical anthropologist, Moerman has devoted a significant portion of the database to ethnopharmacological data. Moerman has spent over 25 years in developing the database through the systematic exploration of scientific literature, ethnographic accounts and historical documents. The result is virtually a census of plant use by Native Americans. The database contains 44,691 items which represents uses by 291 Native American groups of 4029 species from 243 different plant families of which 24,945 are ethnopharmacological entries representing 2582 species. In addition, the database is linked to the United States Department of Agriculture PLANTS database (http://plants.usda.gov/). This allows for the cross referencing of botanical information for each plant along with pictures, range maps and endangerment status.

-
-
-

Bibliography

Bye R. (1986). Voucher specimens in ethnobiological studies and publications. *Journal of Ethnobiology* **6**(1), 1-8. [This article notes the importance of properly collected herbarium voucher specimens in ethnopharmacology and related disciplines]

Cook F. (1995). Economic Botany Data Collection Standard. Prepared for the International Working Group on Taxonomic Databases for Plant Sciences (TDWG). Kew, Royal Botanic Gardens, UK. [This article has raised awareness and created much needed discussion on ethnopharmacological and ethnobiological data standards.]

Lee W.L. and Bell B.M. (1982). Guidelines for acquisition and management of biological specimens. Association of Systematics Collections, Lawrence, Kansas. [This presents important guidelines and standards relevant to the collection of ethnopharmacological data.]

Loub W., Farnsworth N.R., Soejarto D.D. and Quinn M.L. (1985). NAPRALERT: Computer Handling of Natural Products Research Data. *Journal of Chemical Information and Computer Science* 25:99-103. [This article introduces the NAPRALERT database and demonstrates the utility of such an approach.]

Walter S. Judd, Christopher S. Campbell, , Elizabeth A. Kellogg, Peter F. Stevens, and Michael J. Donoghue (2002). Plant Systematics: A Phylogenetic Approach. Sinaur Press. [This is the most up to date text on a rapidly changing field crucial to ethnopharmacology.]

Martin G. (1995). Ethnobotany: A methods manual. London, Chapman and Hall. [This is a seminal work on methods for collecting ethnobotanical data.]

Moerman D. (1998). Native American Ethnobotany. Portland, Timber Press. [This work represents the most complete inventory of ethnopharmacological data for Native North Americans and sets new standards for data management.]

Salick, J., J. Alcorn, E. Anderson, C. Asa, W. Balee, M. Balick, S. Beckerman, B. Bennett, J. Caballero, G. Camilo, A.B. Cunningham, E. Elisabetsky, L. Emperaire, G. Estabrook, G. Fritz, L. Gross, E. Hunn, T. Johns, E. Luoga, G. Martin, W. McClatchey, J. Miller, P. Minnis, D. Moerman, M. Paletti, D. Pearsall, C. Ramirez-Sosa, J. Rashford, B. Schaal, D. Spooner, J. Stepp, M. Thomas, T. Ticktin, N. Turner, J. Xu (2003). Intellectual Imperatives in Ethnobiology. NSF Biocomplexity Workshop Report. Missouri Botanical Garden. St. Louis, MO. [This is a foundational statement by leading ethnobiologists on emerging research, theoretical and ethical issues.]

Thomas M.B., Lin N. and Beck H.W. (2001). A Database Model for Integrating and Facilitating Collaborative Ethnomedicinal Research. Pharmaceutical Biology. Vol 39, Supplement, pp. 41-52. [This introduces a new object oriented database model with significant potential for ethnopharmacological data.]

Thomas M.B. (2003). Emerging Synergies Between Information Technology and Applied Ethnobotanical Research. Ethnobotany Research and Applications. [This work covers recent advances in information technology and how they might be applied to ethnopharmacology.]

Biographical Sketches

John Richard Stepp is an associate professor of Anthropology at the University of Florida. Since the mid 1990s he has conducted ethnobotanical research with the Tzeltal Maya in Highland Chiapas, Mexico. More recently, he has begun work with the Q'eqchi' Maya, Mopan Maya and Garifuna along the border between Guatemala and Southern Belize. His research explores persistence, change and variation of traditional ecological knowledge. He is a U.S. Environmental Protection Agency-STAR and David L. Boren-NSEP fellow. He is the co-editor of *Ethnobiology and Biocultural Diversity* (2002, Univ. of Georgia Press) and the founding editor of the *Journal of Ecological Anthropology*. He was program chair for the 7th International Congress of Ethnobiology and has served on the American Anthropological Association executive board. He is a recipient of the Society for Ethnobiology Barbara Lawrence Award. His research has been profiled in *Trends in Plant Science*, *Lancet*, and the *New Scientist*.

Michael B. Thomas is Director, Centre for International Ethnomedicinal Research and Education, and Affiliate Researcher at the University of Hawaii at Manoa. Dr. Thomas' research is focused in several areas. One is institutional capacity building through the integration of information technology, specifically database technology with ethnobiological data and the development of virtual herbaria. Of special interest is the methodology in which ethnobotanical research data is recorded and digitally stored, not only in the Pacific, but around the world. A second area of interest is international development concerning trade and the sustainable harvest of Non-Timber Forest Products/agroforestry-based products and industries. He has worked as a natural resource and agribusiness consultant implementing community-based natural resource management, research and field-level development projects with the Peace Corps, United States Agency for International Development and Winrock International. He specializes in improving capacity of private sector agribusiness enterprises to access and supply domestic and foreign markets, increase integration of smallholder agriculture in the supply chain for value-added agro-processing, medicinal and aromatic plants (MAPs) and non-timber forest products (NTFPs). He continues working as a consultant on agribusiness, and rural agricultural development.