SOURCES USED IN HEALTH POLICY RESEARCH AND IMPLICATIONS FOR INFORMATION RETRIEVAL SYSTEMS

BRADFORD H. GRAY, PHD

This conference, *Accessing Useful Information: Challenges in Health Policy and Public Health*, grew from the sense of the organizers that the bibliographic resources available to researchers and practitioners in the fields of public health and health policy may not meet the needs of users as well as they might and do not match the resources available to biomedical researchers and practitioners. The goals were to try to understand the dimensions of the problem and to begin to consider ways that it might be addressed. My contribution to this discussion is as a researcher and user of the literature of health services research and policy. I have had occasion to use this literature as background for my own research and as support for the kinds of meta-usage that occurs at the Institute of Medicine.

I should emphasize at the outset that the progress of recent years in making information available through electronic means is astonishing. I am grateful to the people who have brought us to the present state of affairs because of what they have already accomplished and their manifest commitment to continued improvement of the resources available to researchers and practitioners in health policy and public health. Although this paper identifies some limitations of and problems in the present state of affairs, it should in no way detract from my appreciation for the resources that are available from the nation's libraries and on-line services.

Dr. Gray is Director, Division of Health and Science Policy, New York Academy of Medicine, 1216 Fifth Avenue, New York, NY 10029-5293. (E-mail: bgray@nyam.org)

THE NATURE OF THE PROBLEM

How does one put one's hands (or eyes) on the material he or she needs as a researcher, and how does one evaluate its quality? Many researchers try to obtain their own copies of everything that pertains to their current work and topics of potential future interest. One does the latter because so much material that pertains to health services research and health policy is ephemeral. Thus, in addition to journal subscriptions, one gets on mailing lists for various government agencies and think tanks; reads many newspapers, trade publications, newsletters, and journals; and orders all publications and sources that appear to be "within scope." In an earlier era, this strategy might have worked, but today, the volume and diversity negate the strategy. It is difficult to gain awareness of everything that is available. Moreover, the volume of material can quickly overwhelm one's storage space, and it is difficult to develop a way of organizing materials in a way that allows timely retrieval. I find it increasingly difficult to keep effective track of publications other than those for which I have an immediate need. In short, I need a librarian. Fortunately, my employer, the New York Academy of Medicine, has an excellent library.

Librarians and researchers face the challenge of knowing where pertinent material may be located. We need efficient ways of keeping up to date on the background and current status of information and research on problems of concern. We would like an electronic system to help us do this, but the challenges posed by the development of such a system and by making it usable by nonexperts are formidable:

- One has to know how the problem of interest might be labeled. For example, one may notice that hospitals and companies have been purchasing doctors' practices and wish to know the significance of that development. What might the topic be termed in an indexing system? To cite another example, I am now studying the capitated health plans that have been created by health centers and hospitals that traditionally have served poor populations. Such plans do not have standard names; in New York, some of them are called Prepaid Health Services Plans (PHSPs); they are called by other names in other states. It is difficult even for specialists to keep up, and it is a challenge to any vocabulary-based retrieval system.
- The health policy world spawns new terms and language with astounding rapidity. In many cases, the new words refer to a new phenomenon—ERISA (Employee Retirement Income Security Act of 1974) and HMO (health maintenance organization) and PSRO (Professional Standards Review Organization) in the 1970s; TEFRA (Tax Equity and Financial Responsibility Act)

limits and prospective payment system in the early 1980s; the resource-based relative-value scale of the late 1980s; and the provider-sponsored organizations, integrated delivery systems, physician practice management companies, 1115 waivers, and medical savings accounts of the 1990s. Most recently, the False Claims Act and other fraud and abuse legislation have become prominent.

- Health policy lacks clean boundaries. The major determinants of the health of populations mostly fall outside the traditional definitions of health care—social inequality, racism, bad housing, poor education, unemployment, poverty. Further, much work of great importance for public health is done by people who have little or no interest in health *per se* and who do not publish their work in health-related outlets. My work illustrates the boundary problem. Much of my research for the past 15 years has been aimed at understanding the significance of the growth of corporate ownership of health care organizations. Work directly pertinent to that topic is published in sociology, economics, legal, business, ethics, and history journals; in publications by government agencies, trade associations, advocacy groups, and all manner of think tanks; in corporate annual reports and in filings with the Securities and Exchange Commission; in regulations and court decisions; and in the trade, investment, and general press. For afficionados of the topic, following the saga of Columbia/HCA Healthcare Corporation in 1997 in the *New York Times*, the *Wall Street Journal*, and *Modern Healthcare* was riveting.
- Health policy is very topical; policy decisions by Congress, regulatory agencies, and state governments have enormous consequences on diverse parties. As a result, dozens (perhaps hundreds) of newsletters track different issues, policy-relevant material emanates from countless sources in the public and private sectors, and much available information comes from sources that are not at all disinterested.

All of these aspects of health policy and public health are pertinent to the challenge of creating electronic databases that will meet the needs of users.

A TEST OF ONE TOOL

When I was asked to contribute an article on the use of information-retrieval systems from the point of view of a researcher and user, I conducted some research. My experience with MEDLINE and HealthSTAR has made me appreciate what is available, but I also sensed that there are problems in locating much important information. I decided to test my sense by collecting data.

The premise of my research for this paper was that the citations in core health

services research and policy journals show what sources working researchers are actually using, as does the literature cited in reports by two "consumers" of research—the Institute of Medicine (IOM) and the US government's General Accounting Office (GAO). With the assistance of Ms. Cathy Rowe from the Academy staff, I compiled a list of citations from the lead articles in recent issues of six core health services research/policy journals and two recent reports from the IOM and GAO. Table I shows the articles and reports that we reviewed. They are not representative of anything; rather, they are diverse, at hand, and from mainstream sources. Ms. Rowe and I developed a set of categories into

TABLE I Publications Included in the Project

Publication No.	Authors	Title	Publication
1	Mick S, Comfort M	The quality of care of IMGs: How does it compare to that of the US medical graduates?	*Med Care Res Rev.* 1997;54(4): 379–413
2	Feder J, Lambrew J, Hukaby M	Medicaid and long term care for the elderly: Implications of restructuring	*Milbank Q.* 1997;75(4):425–459
3	Marstellar J, Bovbjerg R, Nichols R, Verrilli D	The resurgence of selective contracting provisions	*J Health Politics, Policy, Law.* 1997;22(5):1133–1413.
4	Baxter RJ, Mechanic RE	The status of the local health care safety net	*Health Aff.* 1997;J/A 16(4): 7–23
5	Young G	Insider representation on the governing boards of nonprofit hospitals: Trends and implications for charitable care	*Inquiry.* 1997; 33(winter): 352–362
6*	Brown R, Phillips B, Bishop B, et al.	The effects of predetermined payment rates for Medicare home healthcare	*Health Serv. Res.* 1997;32(4): 397–413
7	GAO	Health insurance for children: State and private programs create new strategies to insure children	Washington, DC: US General Accounting Office; January 1996. GAO/HHS 96-35
8	GAO	Managed care, explicit gag clauses not found in HMO contracts by physician concerns remain	Washington, DC: US General Accounting Office; August 1997. GAO/HHS 97-175
9	Institute of Medicine. (Jones S, Lewin M, eds.)	*Improving the Medicare Market: Adding Choice and Protection*	Washington, DC: National Academy Press; 1996
10	Institute of Medicine (Edmunds M, Frank R, Hogan M, McCarty D, Robinson-Beale R, Wisner C, eds.)	*Managing Managed Care: Quality Improvement in Behavioral Health*	Washington, DC: National Academy Press; 1997

*Not included in the NLM review of database coverage.

which the citations might be put, and, after some refining of the categories and some discussion of the proper fit for certain publications, we assigned each citation to a category. The results are shown in Table II.

Among the striking aspects of Table II is that only a third of the citations were to what we took to be peer-reviewed journals. Another 8% of the citations were to books. The remaining 60% were mostly to what has come to be known as the grey literature—trade publications newsletters, reports, and other publications from advocacy organizations, trade and professional associations, think tanks and research institutes, government agencies, legal references, and a miscellaneous mix of other sources. In this regard, the literature of health services research and health policy is quite different from the biomedical literature, for which the peer-reviewed literature is the dominant source of citations.

What were the implications for the National Library of Medicine's (NLM) databases, which cover health policy and public health? More than any other

TABLE II Sources Cited in Recent Publications in Health Services Research and Policy

Publication No.[a]	PJ[b]	PTN[c]	GP[d]	AA[e]	TF[f]	GA[g]	BK[h]	PRO[i]	LGL[j]	NP[k]	Total
1	64	3	0	4	2	13	4	3	0	0	93
2	4	1	1	3	7	7	1	0	0	2	26
3	8	23	3	13	4	6	0	6	1	5	69
4	5	0	3	0	2	1	1	1	0	2	15
5	23	8	0	1	2	3	6	0	0	0	43
6	3	0	0	0	10	0	0	0	0	0	13
7	5	0	0	0	4	5	0	0	1	0	15
8	2	0	0	3	1	0	0	0	3	3	12
9	32	3	6	4	26	36	2	7	0	3	119
10	144	32	3	60	75	78	53	18	1	20	494
Total	290	70	16	88	133	149	67	35	6	35	889
Percentage of all references	32.62	7.87	1.8	9.9	14.96	16.76	7.54	3.94	0.67	3.94	100.0

[a]As indicated in Table I.
[b]Peer-reviewed journals.
[c]Periodicals, trade press, and newsletters.
[d]General press.
[e]Association and advocacy organization publications.
[f]Think tank/foundation publications.
[g]Government agency publications.
[h]Books.
[i]Proprietary information.
[j]Legal references.
[k]Nonpublished sources.

single source, NLM's system has ambitions at comprehensive coverage in this field. Its HealthSTAR database is self-described as containing:

> citations to the published literature on health services, technology, administration, and research. It focuses on both the clinical and non-clinical aspects of health care delivery. The following topics are included: evaluation of patient outcomes; effectiveness of procedures, programs, products, services and processes; administration and planning of health facilities, services and manpower; health insurance; health policy; health services research; health economics and financial management; laws and regulation; personnel administration; quality assurance; licensure; and accreditation.

As the forum on assessing useful information was being planned, Marjorie Cahn of NLM offered to have NLM review the citations to determine the extent to which they were included in HealthSTAR, MEDLINE, and CATLINE. Ione Auston and Marlyn Schepartz undertook this task. The results are summarized in Table III, which includes 9 of the 10 publications that are covered in Table II.

The data in Table III show that about 68% of the cited publications were included in NLM's databases. (The "not true citation" line is not counted in this

TABLE III Coverage of Citations in National Library of Medicine Databases*

	Publication Number†										
Citation Analysis	1	2	3	4	5	6‡	7	8	9	10	Total
Total citations	93	25	69	19	42		20	17	119	484	888
Not true citation§	1	2	7	10			2	8	7	48	85
Not currently in scope‖		1	6	2	9		6	3	16	27	70
Proprietary¶		1	11	1					15	6	34
Adjusted total citations	92	21	45	6	33		12	6	81	403	699
CATLINE	14		6		6		2	2	16	132	178
HealthSTAR#	2	9	10		9		1		12	14	57
Other	7	3	12					2	5	21	50
Total indexed/held by NLM	85	13	36	4	31		9	6	56	313	553
Percentage available from NLM	92	62	80	67	94		75	100	69	78	79

*This Table is based on a table prepared by Ione Auston of the NLM staff.

†As indicated in Table I.

‡Publication 6 not included in NLM analysis.

§Not true citation: telephone interview, press release, speech, data tape, explanatory footnote, Web URL, etc.

‖Not currently in scope: general economics, newspapers/letters (not indexed comprehensively), legal/testimony/hearings.

¶Proprietary: commissioned papers or studies, briefings to agencies, economic studies and forecasts, and consulting firm reports.

#Numbers exclude CATLINE and MEDLINE records appearing in HealthSTAR.

calculation; the "proprietary" and "not in scope" lines are.) Of the 250 items not in the database, 28% were to publications deemed out of scope, and another 14% were to publications from proprietary organizations and are apparently not available to libraries. Of the remaining publications (the "adjusted total cites" in Table III), 79% of the citations were in NLM's databases. Whether one uses the 68% or 79% number (a case can be made for either), the coverage is far lower than in the biomedical literature, for which, I am told, 95% is the norm.

NLM's system is the single most comprehensive resource available to researchers and practitioners, but clearly not everything that falls within its own coverage goal (as defined above) is included in the database. A list of the sources of the citations that were not included in these databases is shown on Table IV. The list includes some publications from important government agencies (Health Care Financing Administration, GAO, etc.), major advisory entities (MedPAC and its ancestors, ProPAC and PPRC), think tanks, professional and trade associations, and consulting firms. Coverage also is limited regarding legislative, regulatory, and judicial activity. (There is also the matter of publications in disciplinary outlets that are not primarily concerned with health care; I have not explored the extent to which this is a problem.*)

In commenting on the results of my research of the 10 publications, staff at NLM noted several types of exclusions from the NLM system. Thus, for example, GAO reports are indexed, but published GAO congressional testimony is not.† Annual compilations of information (e.g., the *InterStudy Competitive Edge*) are collected by NLM, but are not necessarily indexed. Although researchers may obtain (and cite) "proprietary reports," these are generally not available to libraries or their databases.

NLM staff also noted that certain materials that are available in other databases may not be included fully in NLM's database. Behavioral health material, they note, also is covered in Mental Health Abstracts, the Center for Mental Health Services Publications Database, the Information about Drugs and Alcohol Database, and Psychological Abstracts. It also was noted that NLM does not index general newspapers, and that newsletters are not indexed (although some are collected by NLM) because the contents are often ephemeral (e.g., predictions

*One method for doing this would be to search disciplinary databases in the social sciences, law, public administration, management, and the like using some key words from health care (e.g., health, illness, disease, death, mortality, medicine, doctors, physicians, nurses, drugs, mental illness, hospital, HMO, etc.) and to examine the extent to which the publications so identified are now covered in HealthSTAR.

†It is worth noting that GAO's own monthly list of publications that are available for ordering includes both reports and testimony.

TABLE IV Sources for Citations Not in the National Library of Medicine Database*

Organizations	Government Agencies	Publications	Legal References	Other Sources
Anderson & Co.	Agency for Health Care Policy and Research	*Federation Bulletin* (National Council of Medical Examiners)	Bills	Speeches
AARP	Bureau of Labor Statistics	*Fortune* Magazine	Case law	Press conferences
AMA Council on Ethics & Judicial Affairs	Department of Health & Human Services	*Houston Law Review*	Conference reports	Memoranda
Blue Cross & Blue Shield Assoc.	Congressional Budget Office	*Medicine and Health* (newsletter)	Enacted legislation	Conference materials
Forster-Higgins	Federal Hospital Insurance Fund	*Managed Care Week* (newsletter)	Regulations	Software
GWU Center for Health Policy Research	General Accounting Office	*New York Times*		Documentation
GWU Health Insurance Reform Project	Health Care Financing Administration	*National Journal*		Data tape
GWU National Health Forum	Office of Inspector General	*State Health Notes*		Letters
Health Insurance Assoc. of America	Physician Payment Review Commission	*Wall Street Journal*		Telephone interviews
Institute of Medicine	US Census Current Population Survey	*Washington Post*		Personal interviews
InterStudy	US House of Representatives			
Kaiser Family Foundation	US Senate Special Committee On Aging			
Kaiser Commission on Medicaid				
KPMG Peat Marwick				
Lewin Group				
Mathematica Policy Research				
Minnesota Health Data Institute				
National Academy of State Health Policy				
National Committee on Quality Assurance				
National Conference of State Legislatures				
National Governors Association				
Price Waterhouse				
Robert Wood Johnson Foundation				
U. of Michigan School of Public Health				
U. of Minnesota				
Wyatt Associates				

*Eight publications.

of what might happen) and because newsletters themselves are moving toward on-line availability. It also was noted that the Library of Congress assumes major responsibility for collecting, organizing, maintaining, and providing access to legal/legislative material—particularly on the federal level, and that these materials can be accessed by researchers through that library's Web resources.

Duplication of effort and cost-effectiveness clearly are legitimate concerns in a world of scarce resources. It also is clear that users need to be aware that no one system is comprehensive and to understand that there inevitably are quirks regarding what is and is not included. The further one gets away from standard modes of publication and core topics, the wider are the gaps in coverage. Experienced reference librarians recognize and understand these limitations, and they know the locations of alternative sources of information. Less experienced users would benefit either from broader coverage or from the inclusion in the database of pointers to sources of information for topics not covered.

If a library would like to enhance its coverage of the so-called grey literature, it should not be difficult to be placed on the mailing list of organizations that regularly generate reports and other publications that are pertinent to health policy and to add those materials to the database. As new sources come to light in newsletters and the news media, they can be contacted and added.

THE QUALITY AND INTERPRETATION OF INFORMATION

The more that sources other than peer-reviewed journals are covered by databases, the more they will include sources that are not "scientific." Some publications essentially report current events. Other publications come from organizations for which research is not the primary agenda. Some originating organizations clearly have an interest in the outcome of policy debates, an interest that may bias their publications in obvious and subtle ways.

This raises the question of the quality of information. A consequence of expanding databases beyond the peer-reviewed literature is that the screen of peer review, imperfect as it is, is not there. Publications from some organizations may be prepared with an eye toward support of a particular private interest or ideology. Sophisticated users understand this and apply discounts if needed. Although nothing will substitute for the sophistication of an experienced user, a database could deal with the problem in several ways.

One would be to augment the "publication types" typology that is now available to users of the MEDLARS databases. It is now possible to focus on certain types of materials, including journal articles and many other types of

information, such as bibliographies, literature reviews, and clinical trials. However, the publication type list reflects the categories in the grey literature imperfectly; many of the categories are not available to ordinary users of HealthSTAR. The publication types include "news," "newspaper articles," and "legal briefs" (although not in HealthSTAR), but it is not possible to flag trade press and newsletters, government reports and publications, or reports from think tanks, trade associations, or consulting firms. Revision and updating of the publication types, and making these types more readily available in HealthSTAR, would enhance the system's utility for the health policy and public health worlds.

The use by journals of formatted (or structured) abstracts also may make it easier to identify articles that are based on sound research methodologies without having to review dozens or hundreds of articles that appear in databases despite including no new research findings.

In addition, for the publications produced by think tanks, trade associations, consulting firms, and so forth, links could be set up to a source that briefly describes the organizations, how they are funded, and what interests, if any, they represent.

FINAL OBSERVATIONS

The access to information that has been made available through electronic databases is remarkable, but many publications and other information are not captured in any one source. The more one moves beyond the journal literature, the more one encounters publications with a pedigree that is uncertain or dubious. Findings and conclusions that have little basis can gain credibility by appearing on printouts from databases that include publications from recognized journals. Segregating materials by source may help with the problem, but is not a full solution.

In addition, the availability of electronic databases allows naive users to jump into new topics quickly, perhaps acquiring a veneer of expertise, without having the knowledge needed to locate the topic in both its historical and broader context. Thus, the role of literature reviews by expert individuals and organizations may gain importance in the future and is an activity worthy of encouragement and support.

The field of public health and health policy is extraordinarily broad and increasingly specialized. Whether in updating the vocabulary or in casting the net for items to be included in databases, the collaboration of advisers with diverse backgrounds and interests will be necessary. In a field such as public

health and health policy, considerable diversity by discipline and substantive interest is needed. None of us is aware of how much we do not know.

ACKNOWLEDGEMENT

Appreciation is due to Gail Hendler, of the library of the New York Academy of Medicine, for answering many questions; and to Cathy Rowe, Division of Health and Science Policy, New York Academy of Medicine, and Marjorie Cahn, Ione Austin, and Marlyn Schepartz, National Library of Medicine, for assistance that is reflected in the tables used herein. The interpretation of these materials is the responsibility of the author.