Information Processing & Management Vol. 28, No. 1, pp. 89–98, 1992
Printed in Great Britain.

0306-4573/92 $5.00 + .00

SUBJECT CONTROL OF THE LITERATURE OF ACQUIRED IMMUNODEFICIENCY SYNDROME (AIDS)

ESTHER GREEN BIERBAUM
School of Library and Information Science, University of Iowa, Iowa City, IA 52242, U.S.A.

TERRENCE A. BROOKS*
Graduate School of Library and Information Science, University of Washington, Seattle, WA 98195, U.S.A.

and

ROBERT M. BROOKS
Alameda Hospital, Alameda, CA 94501, U.S.A.

Abstract—This study analyses the *Medical Subject Headings* (MeSH) terms used to index the literature of Acquired Immunodeficiency Syndrome (AIDS). The AIDSLINE database for the years 1983–1989 was the source for approximately 20,000 bibliographic records, which yielded more than 190,000 instances of subject headings. Subject headings introduced in 1987, 1988, and 1989 dominate the present subject access to the literature of AIDS. Indeed, in the short span of six years, even the set of subject headings used by the National Library of Medicine (NLM) to construct the AIDSLINE database has been largely supplanted. The phenomenon of temporarily prominent subject headings reflects the uncertainty engendered by the AIDS phenomenon and the different approaches taken in the underlying biomedical investigations. The occurrence and distribution of subheadings suffixed to the subject "Acquired Immunodeficiency Syndrome" also reflect changes in focus and direction recorded in the literature of AIDS during its early turbulent history.

INTRODUCTION

This study investigates the subject access to the rapidly growing literature of Acquired Immunodeficiency Syndrome (AIDS). In an earlier paper Brooks *et al.*, (1990) examined the explosive growth and change in the scholarly journal cohort of the literature of AIDS. They found that the number of journals had increased rapidly, that specialty journals had replaced more general journals, and that the center of the literature was bifurcated into a core of highly productive biomedical journals and a pair of science news journals reporting developments in the field. Our focus in the present study is on the frequency and extent of variance occurring in the indexing of this new, active literature. Such variance reflects the turbulence and uncertainty in the early stages of the development of the underlying knowledge base, and is of interest to historians of science and investigators in the disciplines of communication and information science. The implications of this study also are of immediate concern to anyone seeking current or historical information about AIDS in online databases.

THE LITERATURE OF AIDS

The biomedical literature of AIDS is distinguished by a known and recent history; it burst upon the American medical scene in 1980. The disease was variously diagnosed and denoted by several research teams in the years 1980–1981 (Gallo & Montagnier, 1987). It was first officially termed "AIDS" in the September 24, 1982 issue of *Morbidity and Mortality Weekly Report* (*MMWR*): ". . . what is now referred to as Acquired Immune [*sic*] Deficiency Syndrome (AIDS)" (vol. 31, page 507). The name of the causative agent was formalized in late 1985 when a nomenclature committee recommended that the more generic

*Supported by a Graduate School Research Fund Grant from the University of Washington, 1990.

term, Human Immunodeficiency Virus (HIV), be adopted rather than continuing the Roman numbering of strains of Human T-Lymphocyte (HTL) retroviruses (Coffin *et al.*, 1986).

The literature of AIDS has had a short lifespan, but an extremely active one. The growth and diversity of the literature prompted the National Library of Medicine (NLM) through its Medical Literature Analysis and Retrieval System (MEDLARS) to develop a database for AIDS. Introduced in July 1988, AIDSLINE was "intended to be a bibliographic file of published literature on AIDS focusing on the biomedical, epidemiologic, and social and behavioral science literature" (AIDSLINE, 1988). Due to the unsettled nature of the terminology during the early years of AIDS research, NLM employed three different literature searches in order to retrieve AIDS-related citations from the MEDLINE database (AIDSLINE, 1988). The three retrieval searches focused on the following precursor, related, and current terms and their many syntactic variants:

- Acquired Immunodeficiency Syndrome
- AIDS Related Complex
- Cytomegalic Inclusion Disease
- HIV
- Homosexuality
- Human T Cell Leukemia Virus
- Immunoblastic Lymphadenopathy
- Pneumonia, Pneumocystis Carinii
- Sarcoma, Kaposi's

Records were loaded from MEDLINE files dating back to 1980 in the first phase of database development. Plans for future development include expansion beyond the MEDLINE journal literature to include proceedings, dissertations, monographs, governmental reports, and so on, and then to databases beyond MEDLARS.

DATA COLLECTION

The study data are citations collected from AIDSLINE during late 1989 and early 1990. In Fall, 1989, the Medical Subject Headings (MeSH) tree number "C20.673.483.23" identified the subject heading "Acquired Immunodeficiency Syndrome"; in early 1990 the equivalent MeSH number of "C20.673.483.440.40" was used.

As Table 1 illustrates, 21,114 bibliographic records were examined for the years 1983 through 1989. The subject headings affixed to each record were captured, producing 193,873 instances of subject heading usage. Thus our raw data are the subject headings affixed to the bibliographic records captured from AIDSLINE using these two MeSH tree numbers and during two data collection periods, late 1989 and early 1990. Table 1 also illustrates that, while the growth in the number of records in AIDSLINE has been furious, the average number of subject headings per record has been relatively steady at approximately nine subject headings per bibliographic record.

Table 1. Frequency of bibliographic records and subject headings

Year	Records	Subject headings	Average SH/record
1983	489	4,473	9.14
1984	1,041	10,946	10.51
1985	1,703	18,159	10.66
1986	2,542	26,482	10.41
1987	3,497	30,291	8.66
1988	5,357	45,196	8.43
1989	6,485	58,326	8.99
Totals	21,114	193,873	9.18

Computations and analyses were performed using The Bibliometrics Toolbox©, a computer program that permits the editing and bibliometric analysis of downloaded files.

MeSH INDEXING

Indexing is the attachment of a subject label to a bibliographic record so that the record may be retrieved from a database by subject rather than author or title. MeSH indexing as applied to MEDLINE has several features that make it a particularly flexible and functional means of subject retrieval. First, there is the advantage of variety and completeness of access in the online search. Whereas articles are indexed simultaneously for the NLM online system (MEDLINE) and the print formats, *Index Medicus* (*IM*) and *Current Catalog* (DuPont, 1990), prior to January 1991, *IM* (or print) terms were distinguished by an asterisk (*) from the more numerous nonprint (not *IM*) headings; but both were accessible in the online system. Since 17 December 1990 a new policy has ended the use of minor (not *IM* or "see under") descriptors, converting existing ones to major descriptors ("Minor Descriptors Discontinued . . . ," 1990, p. 18). During the period of our data, however, a larger number and variety of terms were available for online retrieval than for print format subject search.

Secondly, MeSH is geared to the current biomedical literature. As a subject heading authority, it has the advantage of not being tied to a physical collection; rather, it is applied to an active body of literature which quickly adopts new, functional terms. The terms are then introduced to the MeSH updating list of changes, and added to the next year's permanent roster (*Medical Subject Headings (MeSH) Annotated Alphabetic List*, 1989). For example, "Acquired Immunodeficiency Syndrome," having been sanctioned in its acronym form in MMWR in September 1982, entered MeSH in January 1983. MeSH terms, then, reflect current usage in the biomedical literature.

Third, MEDLARS seeks to make the subject labeling of a document coextensive, that is, to express all of, but not more than, the range of the document content. To this end, MeSH indexing is coordinated: "the concepts in the text of articles are expressed by the combination or coordination of two or more indexing terms" (*MEDLARS Indexing Manual, Part II*, section 4.1). Coordination is achieved variously: (a) with two or more main headings (i.e., terms in post-coordination); (b) with a main heading and a subheading; and (c) with a "pre-coordinated main heading, i.e., one created from a past frequent co-occurrence of two concepts" (*Manual*, 1983, 4.2). Thus "Liver" and "Rural," if brought together with an *and* command, are post-coordinated, while "Liver Glycogen" (Liver + Glycogen) is precoordinated. On the other hand, "metabolism of the liver" is expressed as a main heading and subheading, or aspect, for example, "LIVER/metabolism." (*Manual*, 1983, 4.2.)

Fourth, MeSH employs several devices to partition the bibliographic data and facilitate searches: (a) nonsubject "check tags" (e.g., Animal, Human, Adult), applied routinely to sort the literature according to the research or clinical population of interest (*Manual*, 1983, 18.8.1); (b) form and geographic terms, locating research by region or country or noting the form of presentation, such as a literature review or an English abstract of a foreign language article; and (c) subject subheadings, also a coordination device.

As described below, the subheadings of AIDSLINE presented a special difficulty in processing our raw data. The use of MEDLARS subheadings is complex: They signify aspects of subject headings, but they are used primarily as gathering devices, "convenient means of grouping together constantly recurring aspects of a subject," (*Manual*, 1983, 19.1.1). The application of subheadings in MeSH is closely structured; certain subheadings may be applied only to subject headings in specific tree categories. Indexers are instructed to qualify each main heading, if possible, and to qualify the same heading with more than one allowable subheading. But when more than two topical subheadings seem appropriate, indexers are instructed to use the main heading alone (*MeSH Annotated Alphabetic List*, 1989, I.125). As a result, the presence or absence of subheadings must be regarded with caution; a subject heading lacking subheadings may properly not deserve any, or perhaps should have three or more. Hence, the gathering function of the subheading may be

lost to the searcher. As noted later, these factors influenced the manner in which we processed our data.

TRIMMING SUBJECT HEADINGS

Our goal was to gain some insight into the frequency of use of subject headings indexing the literature of AIDS. Our initial efforts were frustrated by the fact that subject headings, such as "Acquired Immunodeficiency Syndrome," are festooned with a wide variety of subheadings in many different combinations. Furthermore, prior to 1991, headings and subheadings might or might not be not *IM*.

Our initial analyses treated each subject heading, subheading, and asterisk combination as distinct. These initial results proved to be misleading because frequently occurring subject headings were scattered over a wide range of more infrequent variants. Thus the impact of an extremely frequent subject heading was obscured. Consider, for example, the first five records of September 1989 that used these representations:

- Acquired Immunodeficiency Syndrome/*THERAPY
- Acquired Immunodeficiency Syndrome/*ECONOMICS
- *Acquired Immunodeficiency Syndrome
- Acquired Immunodeficiency Syndrome/*DRUG THERAPY
- Acquired Immunodeficiency Syndrome/*TRANSMISSION

To overcome such subject heading elaboration, we devised the strategy of trimming the asterisks and subheadings from the subject headings. In other words, we treated the preceding list as five occurrences of one subject heading, not as single occurrences of five different headings. Thus the many elaborations of a frequently subdivided subject heading were aggregated to reveal the real frequency of heavily used subject headings.

As described above, the factors affecting subheading practice influenced our decision to exclude subheadings from our analyses. However, in order to explore the effects of subheadings, we focused our last analysis solely on the subheadings of the most frequently occurring subject heading, "Acquired Immunodeficiency Syndrome."

LOOKING BACKWARDS IN TIME

Table 2 displays the ranks of the 30 most frequently occurring 1989 subject headings, form/geographic terms, and check tags. The prior year's frequency ranks of these subject headings, form/geographic terms, and check tags are also given. Thus by scanning the columns from left to right, one can determine whether a 1989 subject heading or check tag has been rising or falling in prominence.

Since ranks merely reflect relative aggregation, we also examined intensity, that is, the proportion of the frequency of the subject heading of interest relative to a year's total of different headings. We found that the intensity of the headings is fairly stable year-to-year, and consistent with the ranking. For example, the check tag, Female, varies within the intensity range of .023–.031, with the rank value of 6 representing intensities of .025, .023, .026, and .031 from 1986 back to 1983. However, none of the year-to-year differences is statistically significant using the z test of difference of independent proportions. A shift in rank from 13 to 7, a considerable drop which occurred with HIV between 1989 and 1988, and which also showed an intensity loss from .016 to .009, was also not statistically significant at the .05 level. However, while Acquired Immunodeficiency Syndrome maintained top ranking, it did experience a statistically significant decline in intensity between 1987 and 1986 ($z = 2.375$, $p < .01$). This, as we note below, was a time of the fracture in the literature when many of the original headings dropped out of the top rankings and the focus turned more narrowly on AIDS, with less attention to concomitant issues. Since there was otherwise a dearth of statistically significant difference data, we describe the shift in subject headings frequency in terms of rank only.

Table 2. Frequency ranks by year of top 30 1989 trimmed subject headings

	1989	1988	1987	1986	1985	1984	1983
Acquired Immunodeficiency Syndrome**	1	1	1	1	1	1	1
Human#	2	2	2	2	2	2	2
Male#	3	3	3	3	3	3	3
Adult#	4	4	4	4	4	4	4
Female#	5	5	5	6	6	6	6
Review#	6	12	18	19	17	17	28
HIV Seropositivity	7	10	48	Introduced in 1988			
Review, Tutorial#	8	14	49	Introduced in 1988			
Risk Factors	9	9	46	Introduced in 1988			
United States#	10	6	9	14	16	10	8
Case Report#	11	8	8	8	10	7	7
Support, Non-U.S. Gov't#	12	13	10	13	12	11	15
HIV**	13	7	6	40	Introduced in 1987		
Middle Age#	14	11	11	12	9	8	9
Adolescence#	15	21	20	22	23	25	20
English Abstract#	16	16	16	18	27	19	31
Opportunistic Infections	17	18	17	73	Introduced in 1987		
Support, U.S. Gov't, P.H.S.#	18	22	15	15	14	15	16
HIV Antibodies	19	85	Introduced in 1989				
Child#	20	19	23	17	20	24	20
HIV-1	21	104	Introduced in 1989				
Homosexuality**	22	17	12	9	7	5	5
Animal#	23	26	24	20	21	20	20
Health Education	24	23	25	43	73	53	*
AIDS-Related Complex**	25	20	14	71	Introduced in 1987		
Sex Behavior	26	31	21	31	35	40	39
Zidovudine	27	98	Introduced in 1989				
Aged#	28	35	29	36	37	28	29
Sarcoma, Kaposi's**	29	24	19	16	11	9	10
AIDS Serodiagnosis	30	98	Introduced in 1989				

*No postings in this year.
**Term used in developing AIDSLINE.
#Check tags, or form/geographic term.

The subject headings of Table 2 can be categorized into two groups: (a) subject headings with a history of continuous high use, such as Acquired Immunodeficiency Syndrome, Homosexuality, Health Education, Sex Behavior, and Sarcoma, Kaposi's; and (b) newer subject headings that are usually specific to AIDS, such as HIV Seropositivity, HIV, Opportunistic Infections, HIV antibodies, HIV-1, AIDS-Related Complex, Zidovudine, and AIDS Serodiagnosis.

It is immediately apparent from Table 2 that the outstanding characteristic of the current AIDS literature is the predominance of new subject headings. Each of the newer subject headings has moved rapidly to a high frequency ranking. This characteristic is reinforced by the behavior of the older terms that have maintained a more distant connection to the core of the AIDS literature. For example, both Homosexuality and Sarcoma, Kaposi's drift over time into relative obscurity. Sex Behavior maintains a more or less constant distance from the center of the AIDS literature, while Health Education drifts, without vigorous action, closer to the center.

Table 2 indicates that the developing literature of AIDS is driven by new, specific headings, and a continuing association with a select few other headings. It is particularly notable that subjects used to create the AIDSLINE database, such as Cytomegalic Inclusion Disease, Human T Cell Leukemia Virus, Immunoblastic Lymphadenopathy, and Pneumonia, Pneumocystis Carinii, are missing from the 1989 list of 30 top-ranked headings. Check tags are also indicators of research developments. For example, the check tags Human, Male, Adult, and Female maintain their positions, while Middle Age and Adolescent show

Table 3. Frequency ranks by year of top 20 1983 trimmed subject headings

	1983	1984	1985	1986	1987	1988	1989a	1989b
Acquired Immunodeficiency Syndrome[a]	1	1	1	1	1	1	1	1
Human[b]	2	2	2	2	2	2	2	2
Male[b]	3	3	3	3	3	3	3	3
Adult[b]	4	4	4	4	4	4	4	4
Homosexuality[a]	5	5	7	9	12	17	20	22
Female[b]	6	6	6	6	5	5	5	5
Case Report[b]	7	7	10	8	8	8	11	8
United States[b]	8	10	16	14	9	6	10	7
Middle Age[b]	9	8	9	12	11	11	14	9
Sarcoma, Kaposi's[a]	10	9	11	16	19	24	24	29
Hemophilia	11	18	24	23	29	46	51	47
Pneumonia, Pneumocystis Carinii[a]	12	16	26	27	27	33	29	30
T-Lymphocytes	13	14	15	16	22	34	38	44
Risk	14	13	13	11	13	87	92	82
Support, Non-U.S. Gov't[b]	15	11	12	13	10	13	12	10
Support, U.S. Gov't, P.H.S.[b]	16	15	14	15	15	22	18	16
Haiti[b]	17	31	52	69	87	112	94	93
Cytomegalic Inclusion Disease[a]	18	18	25	28	28	40	43	39
Blood Transfusion	19	20	19	21	25	37	39	40
Adolescence[b]	20	25	23	22	20	21	16	13
Animal[b]	20	20	21	20	24	26	22	21
Child[b]	20	24	20	17	23	19	19	18

1989a includes January through June 1989.
1989b includes July through December 1989.
The "T-Lymphocytes" data include "T Lymphocytes" for 1986 through 1989b.
[a]Term used in developing AIDSLINE.
[b]Check tag, form/geographic term.

changes in their prominence. This movement suggests that the research populations have remained stable through the history of AIDS, except for the recent (1989) surge in studies of adolescents and a slight decline in studies of the middle-aged.

LOOKING FORWARD IN TIME

Table 3 presents the 20 subject headings and check tags most frequently used in 1983, and their rankings in subsequent years.[†] Except for Acquired Immunodeficiency Syndrome, all are drifting away from the core of the AIDS literature. Table 3 demonstrates the extraordinary time compression factor operating in the rapidly expanding literature of AIDS. Within six years, subject headings such as Hemophilia, Pneumonia, Pneumocystis Carinii, T-Lymphocytes, and Cytomegalic Inclusion Disease, which were above the 20th rank in 1983, have faded into obscurity. Two of these headings, Cytomegalic Inclusion Disease and Pneumonia, Pneumocystis Carinii, were used to construct AIDSLINE. Of the top ten 1983 subject headings, only three continued to be in the top 30 ranks in 1989: Acquired Immunodeficiency Syndrome, Homosexuality, and Sarcoma, Kaposi's.

It would appear that by the time the nomenclature committee was readying its 1985 report on selecting "HIV," research findings had become more focused. A fracture in subject heading use reflects this change: Subject headings pertinent in 1989 were introduced in 1987, 1988, and 1989. Except for the less specific headings, Health Education and Sex Behavior, and the two more distant associates, Homosexuality and Kaposi's Sarcoma, every other medical subject heading achieving prominence in 1989 was introduced into the literature in those three years. Conversely, subject headings prominent in 1983 are ignored in 1989; all of the subject headings that characterized the database in 1983 are drifting away

[†]Table 3 displays only the 20 highest frequency ranks because subsequent ranks become crowded with numerous ties resulting from the wide-ranging research and unsettled terminology of the early years of AIDS.

from the current core of the literature. In other words, the process of subject maturation that may take decades to occur in other subjects took place in the literature of AIDS within the span of six years.

SUBJECT HEADINGS OF TEMPORARY PROMINENCE

The preceding analyses chart the utility of those subject headings that were prominent in either or both 1983 and 1989. These analyses, however, miss the subject headings that rose to prominence and then fell back into obscurity between 1983 and 1989.

Table 4 is an alphabetical presentation of the top-ranked 20 subject headings in each of the years from 1983 to 1989. Unlike the 1983 subject heading that fell into obscurity, and also unlike the newer headings that emerged and maintained rank in 1989, a few subject headings broke the 20th-rank barrier only in the intervening years. The prominence of these subject headings is primarily due to extraordinary single source events, such as conferences or special journal issues.

Antibodies, Viral, was introduced to MEDLINE in 1973, and was particularly active in the AIDS literature during 1984–1988. The major sources of this heading were *Lancet* (9.3%), *Journal of the American Medical Association* (4.2%), and the *New England Journal of Medicine* (3.7%). One extraordinary source event, the *International Conference on*

Table 4. Alphabetical listing by year of 20 top-ranked trimmed subject headings

	1983	1984	1985	1986	1987	1988	1989
Acquired Immunodeficiency Syndrome**	1	1	1	1	1	1	1
Adolescence#	20				20		15
Adult#	4	4	4	4	4	4	4
AIDS-Related Complex**					14	20	
Animal#	20	20		20			
Antibodies, Viral**		17	8	7	7	15	
Blood Transfusion**	19	20	19				
Case Report#	7	7	10	8	8	8	11
Child#	20		20	17		19	20
Cytomegalic Inclusion Disease**	18	18					
English Abstract#		19		18	16	16	16
Female#	6	6	6	6	5	5	5
Haiti#	17						
Hemophilia**	11	18					
Homosexuality**	5	5	7	9	12	17	
HIV**					6	7	13
HIV Antibodies**							19
HIV Seropositivity**						10	7
HTLV Viruses**		12	5	5			
Human#	2	2	2	2	2	2	2
Male#	3	3	3	3	3	3	3
Middle Age#	9	8	9	12	11	11	14
Opportunistic Infections**					17	18	17
Pneumonia, Pneumocystis Carinii**	12	16					
Retrovirus Infections**			18	10			
Review#		17	17	19	18	12	6
Review Tutorial#						14	8
Risk**	14	13	13	11	13		
Risk Factors**						9	9
Sarcoma, Kaposi's**	10	9	11	16	19		
Support, Non-U.S. Gov't#	15	11	12	13	10	13	12
Support, U.S. Gov't, P.H.S.#	16	15	14	15	15	18	18
T-Lymphocytes**	13	14	15	16			
United States#	8	10	16	14	9	6	10

The "T-Lymphocytes" data include "T Lymphocytes" for 1986.
**Term used in developing AIDSLINE.
#Check tags or form/geographic term.

Acquired Immunodeficiency Syndrome, April 14–17, 1985, Atlanta, Georgia, accounted for 4.0% of the total occurrences of this subject heading.

HTLV Viruses is an example of the terminological adjustment that has occurred in the AIDS literature. While HTLV Viruses became an established subject heading only in 1989, it can now also be used retrospectively as a search term equivalent to Human T-Cell Leukemia Virus in the literature between 1984–1988. The use of HTLV Viruses peaked in April 1986, the month the International Conference on Acquired Immunodeficiency Syndrome was indexed. This conference contributed 7.5% of the uses of this subject heading; other major contributors were *Science* (6.7%) and *Lancet* (6.6%).

The Atlanta Conference also contributed to 23% of the uses of Retrovirus Infections. The frequency of this heading, established in 1981, was further influenced by the September 1985 issue of *Cancer Research* (5%) and the *XXIIth Symposium for Comparative Research on Leukemia and Related Diseases*, Hamburg, July 7–11, 1985 (3%).

SUBJECT SUBHEADINGS

The preceding analyses of trimmed headings provide an overview of the activity of the most frequently used indexing terms in the AIDS literature and amply illustrate the "Acquired Immunodeficiency Syndrome" is the leading subject heading. Table 5, by contrast, analyzes the 10 most frequent occurrences of the single subject heading, AIDS, and the subheadings chosen to specify aspects of that subject. The table cumulates occurrences across time; thus the data for 1989 subsume all other years. Reading this table from left to right enables one to chart the frequency of use of subject-subheading combinations over time.

Table 5 shows the top nine subheadings selected from the *C* category MeSH tree and their small degree of movement among frequency rankings. As Table 5 also demonstrates, AIDS appears most frequently precoordinated with a topical subheading. The increasing rank of the single term suggests that more recent articles have dealt more specifically with particular aspects of AIDS, calling for several subheadings which resulted in the default term, AIDS. In only two instances was the subheading non-*IM*.

Because, as previously noted, "subheadings afford a convenient means of grouping together constantly recurring aspects of a subject" (*Manual*, 1983, 19.1.1), the analysis of subheading frequencies reveals the varying degrees of importance ascribed to different aspects of a broader topic. The major focus of AIDS research has consistently been the physiological ramifications of this devastating disease—its complications and sequelae and the diseases coexisting or associated with it—as shown not only in the primary ranking of /Complications, but as well in its cumulated total frequency of 2620 occurrences. From 1985 on, /Prevention & control and /Transmission are close in ranking and linked in their increasingly frequent dual posting; here is a second research focus. A third focus in the literature is shown by the subheadings /Immunology and /Diagnosis, which are linked in the scope notes (*MeSH Annotated List*, 1989, p. 419). Both terms are linked back to /Prevention & control, creating an overlapping set of concerns. /Epidemiology is linked to /Trans-

Table 5. Cumulative ranks of topical subheadings of the subject heading "Acquired Immunodeficiency Syndrome" (AIDS)

	1989	1988	1987	1986	1985	1984	1983
AIDS/*Complications	1	1	1	1	1	1	1
AIDS/*Prevention & Control	2	2	4	4	4	6	2
AIDS/*Transmission	3	3	3	3	3	4	6
AIDS/*Immunology	4	4	2	2	2	2	5
*AIDS	5	5	9	12	10	12	6
AIDS/*Epidemiology	6	6	5	7	4	3	4
AIDS/Complications	7	8	7	6	5	7	9
AIDS/*Diagnosis	8	7	6	8	8	8	8
AIDS/*Prevention & Control/Transmission	9	10	13	15	16	16	16
AIDS/*Psychology	10	12	14	17	14	14	unused

mission in another overlap. From its 1984 introduction, /Psychology steadily increased in frequency, and represents another, more recent research area, that of the affective consequences of the disease. /Therapy does not figure in this top-ten ranking of AIDS frequencies; the early literature of AIDS was largely devoted to identifying and coping, and until transmission mechanisms were verified and the causative agent identified, there was little hope of preventive vaccine or drug therapies. As shown in Table 3 by the 1989 entry of "Zidovudine" (a possibly promising drug), therapeutic essays came later.

Our investigation of the use of subheadings of AIDS demonstrates that major research concerns centered on the manifestations and characteristics of the disease, then its prevention and control and transmission, and finally—within the time frame of the study—the diagnostic and immunological aspects which would eventually lead to vaccines and therapy.

DISCUSSION

The indexing of the literature of AIDS reflects the explosive growth and astonishing development of the field. Except for the primary subject heading itself, virtually all the terms that were frequently used in the beginning years of the literature had been supplanted in 1989 by new terms or were no longer in the high-frequency group. Superseded terms include most of those used to create the AIDSLINE database.

Conferences, single-topic journal issues, and review articles create identifiable ripples in the frequency of use of specific subject headings. The extent of the variance in several instances is considerable, resulting in a rise, fall, and disappearance of the heading from the ranks of heavily used terms. Examples of this variance are seen in the sole appearance of Haiti in 1983, the early appearance and loss of Pneumonia, Pneumocystis Carinii, the introduction midway of Retrovirus Infections, and the late appearance of HIV.

The extent of variance is also shown by the relative handful of stable terms, most of which are check tags such as Human, Male, and Adult. The sudden rise of form headings such as Review, and Review, Tutorial, indicates that by 1989 the literature had attained a complexity and breadth that required literature and study reviews.

The structure and development of the literature of AIDS and its subject control are of interest to the medical historian and the student of scientific communication. On the more immediately practical level, the growth and change in AIDS subject access is of interest to anyone undertaking research requiring a retrospective literature survey. The retrieval value of current terminology is indisputable in searching recent literature and ongoing studies. But in a rapidly developing literature in which subject headings come into temporary prominence and then fade, or in which many competitor subject headings crowd the list before the literature settles down, research indexed to those headings may be lost to searchers who are not careful and diligent. Hence the very characteristics of a rapidly expanding knowledge base like AIDS—intense but diffuse research activity compressed into a short time—introduce often unexpected degrees of mutability in the use of subject headings; subject access tends to deteriorate over time (Bierbaum, 1989). Clearly, a retrospective database such as AIDSLINE, one capturing literature indexed to noncurrent terms, is of inestimable value to the researcher, and illustrates the potential pitfalls of exploring a rapidly developing knowledge base without a working knowledge of the history of the field.

A study of subject access is one approach to the history of a field. But the mutability of access suggests tantalizing imponderables. For example, tentative evidence suggests that AIDS is not a sudden and recent phenomenon after all (Nahmias, 1986). However, access to prior studies is masked by the inadequacy of precursor headings. The researcher must thus in effect ask, "What was the subject heading *then*?" As Swanson (1989) has pointed out in describing his encounters with deteriorated subject access, subject searching in online databases is a more subtle operation than most information scientists have been willing to admit.

The subject access to the literature of AIDS mirrors the turbulence and uncertainty experienced by investigators in this new field of biomedical knowledge. It can also teach us, on a time-compressed scale, what happens to subject access as the knowledge base and environment of a discipline grow and change.

REFERENCES

AIDSLINE: A new MEDLARS database (1988, July). *The NLM Technical Bulletin, 231*, 1, 11.

Bierbaum, E. (1989, March). The rhetoric of retrieval: Culture and the subject search. Project on the Rhetoric of Inquiry (POROI). Paper presented at the Conference on Discourse, Rhetoric, and Culture. Iowa City: The University of Iowa.

Brooks, T.A., Bierbaum, E.G., & Brooks, R.M. Journal frequency analysis of the literature of acquired immunodeficiency syndrome (AIDS). (1990, November). *American Society for Information Science (ASIS), Proceedings of the 53rd conference*, Toronto, 197–203,

Centers for Disease Control (September 24, 1983). Atlanta, Georgia. *Morbidity and Mortality Weekly Reports, 31*, 507.

Coffin, J., Haase, A., Levy, J.A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P., & Weiss, R. (May 9, 1986). Human Immunodeficiency Viruses [Letter to the editor]. *Science, 232*, 697.

DuPont, G. (December 26, 1990). MEDLARS Management Section. Personal correspondence.

Gallo, R.C., & Montagnier, L. (1987). The chronology of AIDS research. *Nature, 326*, 435–436.

Medical Subject Headings Annotated Alphabetic List. (1989). Washington, D.C.: NLM.

MEDLARS Indexing Manual, Part II. (1983). Washington, D.C.: National Library of Medicine; National Technical Information Service.

Minor descriptors discontinued effective 1991. (1990, September–October). *The NLM Technical Bulletin, 255*, 18.

Nahmias, A.J., Weiss, J., Yso, X., Lee, F., Kodsi, R., Schanfield, M., Matthews, T., Bolognesi, D., Durack, D., & Motalsky, A. (1986). Evidence for human infection with an HTLV III/LAV-like virus in Central Africa, 1959. *Lancet, 8429*, 1279–1280.

Swanson, D.R. (1988). Historical note: Information retrieval and the future of an illusion. *Journal of the American Society for Information Science, 39*, 92–98.