

The BIOSIS® perspective on AIDS research

ANNE COWIE and PAULINE DUCKITT* *Vital Information Limited, Leighton Buzzard, Bedfordshire LU7 8AD*

The resources BIOSIS makes available to AIDS researchers and other interested people are examined. The online database BIOSIS Previews is a vast store of bibliographic biomedical information that can be searched in various ways to retrieve references to the AIDS virus and aspects of the disease—diagnosis, pathology, treatment, epidemiology, and public health. A printed bibliography on AIDS and AIDS-related disease (Collected Papers on AIDS Research 1976-1986) and a complementary current awareness service (AIDS Research Today) are produced as part of the BIOSIS Perspectives Series. The BioBusiness database extends coverage of the literature into those areas of AIDS research that have commercial implications.

Introduction

BIOSIS, a not-for-profit organization based in Philadelphia, has been serving the information needs of the world's life-science community for over 60 years. BIOSIS is perhaps best known as the publisher of *Biological Abstracts*® and *Biological Abstracts/RRM*® (reports, reviews and meetings) but in fact it produces a large range of information products and services in printed, machine-readable and microform media. We shall be considering four of these. The *BIOSIS Previews*® database is the world's largest collection of abstracts and citations for biology and the life sciences, and is widely available for online searching. From this database have been derived two specialist printed information services on AIDS. *Collected Papers on AIDS Research 1976-1986* and *AIDS Research Today*. We shall also look at *BioBusiness*®, an online database devoted exclusively to the economic implications and business applications of biological and biomedical research. Each of these services will be examined to show the perspective they give on AIDS research.

BIOSIS Previews

The *BIOSIS Previews* database, often abbreviated to *BIOSIS*, runs from 1969 to date. *BIOSIS* deals with all areas of life sciences—traditional, interdisciplinary and supporting areas—concentrating on original research and reviews. The source material is not only articles from primary journals, but also papers from meetings plus books, monographs, reports, reviews, US patents and research communications. More than 9000 serials are scanned for the database, 47% coming from Europe and the Middle East, 28% from North America, 16% from Asia and Australasia, 6% from Central and South America and 3% from Africa. Over half a million references (520 000)

*To whom correspondence should be addressed.

will be added to the database in 1988, of which half will be journal articles and half other types of literature, predominantly meetings papers. The total number of items searchable online in 1988 will exceed 6 million, approximately three-quarters of all BIOSIS records produced since 1926.

This outline of the database tells us several things. *BIOSIS Previews* covers the whole period of time during which AIDS research has been going on. It covers the basic sciences of virology and immunology, with all their supporting disciplines, plus the clinical aspects of AIDS. During the period 1981-85, it is estimated that 81% of this huge database related to experimental, clinical or veterinary medicine. *BIOSIS* is therefore the largest and most comprehensive source of bibliographic information in the biomedical sciences. The importance of meetings papers in rapidly conveying new information in a fast-moving area of research cannot be underestimated, and *BIOSIS* is committed to keeping pace with this source.

In order to carry out an effective search online, it is necessary to know the structure of a *BIOSIS Previews* record and the ways in which the subject information can be accessed within each record. A typical record is shown below.

ITNUM 84083075

TITLE THE TOXICITY OF AZIDOTHYIMIDINE AZT IN THE TREATMENT OF PATIENTS WITH AIDS AND AIDS-RELATED COMPLEX A DOUBLE-BLIND PLACEBO-CONTROLLED TRIAL

AUTHS RICHMAN D D; FISCHL M A; GRIECO M H; GOTTLIEB M S; VOLBERDING P A; LASKIN O L; LEEDOM J M; GROOPMAN J E; MILDVAN D; ET AL

AFFIL UNIV. CALIFORNIA, SAN DIEGO

JOURN N ENGL J MED

BIBLO 317 (4). 1987. 192-197.

CODEN NEJMA

LANG ENGLISH

ABST We concluded a double-blind, placebo-controlled trial of oral azidothymidine (AZT) in 282 patients with the acquired immunodeficiency syndrome (AIDS) or AIDS-related complex. Although significant clinical benefit was documented (N Engl J Med 1987; 317: 185-191), serious adverse reactions, particularly bone marrow suppression, were observed. Nausea, myalgia, insomnia, and severe headaches were reported more frequently by recipients of AZT; macrocytosis developed within weeks in most of the AZT group. Anemia with hemoglobin levels below 7.5 g per deciliter developed in 24 percent of AZT recipients and 4 percent of placebo recipients ($P < 0.001$). Twenty-one percent of AZT recipients and 4 percent of placebo recipients required multiple red-cell transfusion ($P < 0.001$). Neuropenia (< 500 cells per cubic millimeter) occurred in 16 percent of placebo recipients ($P < 0.001$). Subjects who entered the study with low CD4 lymphocyte counts, low serum vitamin B12 levels, anemia, or low neutrophil

counts were more likely to have hematologic toxic effects. Concurrent use of acetaminophen was also associated with a higher frequency of hematologic toxicity. Although a subset of patients tolerated AZT for an extended period with few toxic effects, the drug should be administered with caution because of its toxicity and the limited experience with it to date.

KEYWD HUMAN IMMUNODEFICIENCY VIRUS RETROVIRUS ORAL ADMINISTRATION ACETAMINOPHEN SERUM VITAMIN B-12 HEMOGLOBIN LYMPHOCYTE ERYTHROCYTE TRANSFUSION ACQUIRED IMMUNODEFICIENCY SYNDROME BONE MARROW SUPPRESSION NAUSEA MYALGIA INSOMNIA HEADACHE MACROCYTOSIS ANEMIA NEUTROPENIA

CONCP *13018 Metabolism-Water-Soluble Vitamins *15004 Blood, Blood-Forming Organs and Body Fluids-Blood Cell Studies *15006 Blood, Blood-Forming Organs and Body Fluids-Blood, Lymphatic and Reticuloendothelial Pathologies *15008 Blood, Blood-Forming Organs and Body Fluids-Lymphatic Tissue and Reticuloendothelial System *22005 Pharmacology-Clinical Pharmacology (1972-) *22504 Toxicology-Pharmacological Toxicology (1972-) *34504 Immunology and Immunochemistry-Bacterial, Viral and Fungal *34508 Immunology and Immunochemistry-Immunopathology, Tissue Immunology *36006 Medical and Clinical Microbiology-Virology *38506 Chemotherapy-Antiviral Agents 02508 Cytology and Cytochemistry-Human 07004 Behavioral Biology-Human Behavior 10063 Biochemical Studies-Vitamins 11304 Chordate Body Regions-Head (1970-) 14006 Digestive System-Pathology 15001 Blood, Blood-Forming Organs and Body Fluid-General; Methods 20506 Nervous System-Pathology 21002 Psychiatry-Psychopathology; Psychodynamics and Therapy 22100 Routes of Immunization, Infection and Therapy

BSYST 02242 Retroviridae-Lentivirinae (1979-) 86215 Hominidae

The indexing policies for *BIOSIS Previews* are explained here very briefly, and more detail may be obtained by attending one of the free training courses held regularly in many countries throughout the world and from the *BIOSIS Previews Search Guide*.

The title, keyword and abstract data elements all contain subject information arising from the author's own terminology. The words in the title (translated into English where necessary) are augmented by keywords, chosen from the body of the original document by the indexers to enhance the information in the title. Together, these two data elements provide a stock of up-to-date and specific vocabulary for searching. Abstracts exist for 55% of the database, providing extra text where available. Titles and keywords, however, are available for the entire database and indicate the subject coverage of the original document. The keywords are not controlled except in a handful of specific areas, of which drug action is the most relevant to AIDS. There are around 70 'drug affiliation' terms that are assigned by indexers to indicate the drug action being studied in a particular document.

In addition to this textual information, two systems of codes are used. Concept codes are added to indicate the broad subject areas covered by a document. About 500 of these exist. An additional advantage is that concept codes are weighted to show which concepts are of major importance within a particular document. Biosystematic codes are also used to show which broad groups of organisms are being studied.

The recommended approach to searching the *BIOSIS Previews* database is to use title words and keywords to represent specific subjects or organisms (genera or species) and to use concept codes and biosystematic codes as appropriate for broad subject areas and groups of organisms.

As the disease AIDS is a specific concept, it should be retrieved by free text searching of title and keyword data elements. In order to ensure that the word 'AIDS' appears only in the desired context, the most appropriate concept code (CC34508 Immunopathology) should be included in the strategy*.

1. AIDS (TI, KW)
2. ACQUIRED ADJ IMMUN\$1 ADJ DEFICIENCY (TI,KW)
3. ACQUIRED ADJ IMMUNODEFICIENCY (TI,KW)
4. 1 OR 2 OR 3
5. 4 AND CC34508

The truncation in search statement 2 allows retrieval of 'immune' or 'immuno', since 'immunodeficiency' or 'immunodeficiency' may have appeared as 'segmented' words before 1985. Between 1969 and 1984, compound terms occurring in the title or keywords were broken into two or more parts when they produced useful search access points. Each part of the word, or 'segment', must be searched for as a separate word. Compound words are no longer segmented, but search strategies must take this policy into account when the pre-1985 portion of the database is being searched.

A similar strategy is adopted to retrieve references on the virus causing the disease. The biosystematic codes for the retroviruses (BC02240-BC02246) are included to ensure relevance when using abbreviations.

6. HTLV (TI,KW)
7. HUMAN ADJ T SAME VIRUS (TI,KW)
8. IMMUN\$1 ADJ DEFICIENCY SAME VIRUS (TI,KW)
9. HIV (TI,KW)
10. LYMPHADENOPATH\$1 SAME VIRUS (TI,KW)
11. LYMPH ADJ ADENOPATH\$ SAME VIRUS (TI,KW)
12. 6 OR 7 OR 8 OR 9 OR 10 OR 11
13. 12 AND BC0224\$

*The strategies shown in this paper do not reflect the search language of any specific online host and need to be modified to work on the particular online search system used. Search statements are numbered consecutively throughout the paper. TI=title field; KW=keyword (descriptor) field; CC=concept code; BC=biosystematic code; ADJ=occurring next to; SAME=occurring in the same field ; \$=unlimited truncation; \$1=truncation limited to one extra character.

Using these strategies, we found the number of references to AIDS as a disease and the AIDS virus in the *BIOSIS Previews* database. They are shown in Table 1 by year of publication. The number for 1987 represents those entered into the database by the beginning of 1988.

Table 1. References to AIDS in the *BIOSIS Previews* database

Year	Disease	Virus
1981	2	8
1982	20	25
1983	233	59
1984	967	300
1985	1204	534
1986	1769	1056
1987	1576	995

Particular aspects of the disease may be retrieved by including relevant concept codes in the search strategy. Where necessary the concept code can be limited to retrieve references where the desired aspect receives major emphasis. Some of the concept codes most pertinent to AIDS are listed here, along with extracts from their scope notes in the *BIOSIS Previews Search Guide*.

Diagnosis

CC12504—DIAGNOSTIC PATHOLOGY

“studies and techniques used to determine the cause of disease or methods used to distinguish one disease from another”.

Epidemiology

CC37052—COMMUNICABLE DISEASES

“studies on the incidence, distribution and prevalence of infectious diseases caused by bacteria, viruses, fungi and parasites which infect large numbers of humans”.

Pathology

CC34508—IMMUNOPATHOLOGY (TISSUE IMMUNOLOGY)

“studies of cellular and tissue antigens, lymphocyte and RES immune reactivity, immune responses against tissue and tumor cells and depression of the immune response as a result of infectious disease, drugs, radiation and thermal burns. It also accesses information on autoimmune and immune deficiency diseases as well as transplantation immunology”.

Public health

CC37010—PUBLIC HEALTH ADMINISTRATION AND STATISTICS

“studies on public health planning; legislation; administration & education”.

CC37012—HEALTH SERVICES AND MEDICAL CARE

“studies on public health programs and services including programs for home care, patient rehabilitation, community health, nutritional health, mental health, geriatric health and multiphasic screening”.

Treatment

CC12512—THERAPY

“studies on all chemical, biochemical and physical techniques and methods of treatment for the disorders of animals and humans”.

CC38506—ANTIVIRAL AGENTS

“studies on the chemotherapeutic treatment of viral infections”.

Virology

CC36006—MEDICAL AND CLINICAL MICROBIOLOGY: VIROLOGY

“studies of pathogenic animal and human viruses including oncogenic viruses and infections of known or suspected viral etiology”.

Possible search strategies for retrieving aspects of AIDS are given below. Some of the numbers refer back to previous search statements.

14. (5 OR 13) AND CC36006—virology
15. (5 OR 13) AND CC37052—epidemiology
16. (1 OR 2 OR 3 OR 4 OR 13) AND *CC34508 (MAJOR LEVEL)—immunopathology
17. (5 OR 13) AND CC12504—diagnosis
18. 5 AND (CC37010 OR CC37012)—public health
19. (5 OR 13) AND (CC12512 OR CC38506)—treatment

Studies in humans may be retrieved by including the biosystematic code for Hominidae (BC86215), as in the example here.

20. (5 OR 13) AND BC86215 (Hominidae)

It should be noted that some concept codes, for example CC37052 (communicable diseases) and CC37012 (health services and medical care), imply a ‘human’ aspect in the scope of their application.

Studies of the therapeutic actions of antiviral drugs may be retrieved by using the appropriate controlled drug affiliation term, as below:

21. (5 OR 13) AND (ANTIVIRAL ADJ DRUG\$1) (TI,KW)

All the suggested strategies here are valid for the database from 1981 to date. More specific aspects may, of course, be retrieved by including appropriate free text strategies.

Specialist printed information sources

The BIOSIS Perspectives Series is a new type of information service that links past and present research findings on specific topics. The service provides a complete bibliography on the topic, which is then linked to a current awareness service with monthly additions. AIDS has been selected as the first research problem to receive this composite treatment.

Collected Papers on AIDS Research 1976–1986 is a retrospective bibliography providing references to research on AIDS from the *BIOSIS Previews* database. It is particularly valuable as it covers the period before 1981 when the first description of AIDS as a recognizable entity was published. Articles describing neoplasms and opportunistic infections frequently associated with AIDS in populations at risk are included from these years. While these reports do not necessarily indicate infections by the AIDS virus, they are included to ensure its comprehensiveness. The bibliography contains over 4500 items and 34% have abstracts. The items are arranged alphabetically by first author's surname and there are author and subject indexes. The subject index is of the KWIC (Key Word In Context) type and allows searching of all significant words in the title and keyword data elements.

AIDS Research Today, a monthly current-awareness service, has the same arrangement and similar indexes. It is designed to provide the latest information about AIDS research and to complement the bibliography. Each monthly issue is provided ready punched for easy cumulation in a ring binder. The last issue of the year includes a cumulated author and subject index for the year.

BioBusiness

Although the scientific aspect of AIDS research is the main subject of this presentation, AIDS research has considerable implications for business. BIOSIS covers these in its *BioBusiness* database, which will be described very briefly here.

BioBusiness is available only in machine-readable form and can be searched online through several hosts. Introduced in 1985, it provides information on the commercial and industrial applications of research and development in the life sciences. Items are selected from a broad spectrum of life-science, business and management publications. By the end of 1987 there were over 104 000 records on subjects with a combined life sciences and commercial focus. They provide information on discoveries, products, evaluations, trends and forecasts, companies and personnel, and legal aspects. *BioBusiness* records include controlled and non-controlled descriptors and codes for broad

subject areas. People and companies may also be indexed. Abstracts are included for nearly 60% of the database. A typical record is shown below.

AN 0103145

TI AIDS DRUG BEING DROPPED FROM NCI STUDY

AU ANON

JN CHEMICAL MARKETING REPORTER

SO VOL. 232, NO. 9, AUG. 31, 1987, P. 7

PY 1987

LA ENGLISH

AB National Cancer Institute (NCI) has decided to cancel its marketing plans for an acquired immune deficiency syndrome (AIDS) therapeutic, which it hoped would compete with Burroughs Wellcome Co.'s azidothymidine (AZT) or Retrovir. The drug, cyanothymidine, cannot be produced from the intended process, and its closely related analogue, dehydrothymidine, is probably not patentable.

DE PHARMACEUTICAL INDUSTRY; PHARMACEUTICALS; MARKETING; COMPETITION; ACQUIRED IMMUNE DEFICIENCY SYNDROME; CYANOTHYIMIDINE; AZIDOTHYIMIDINE; RETROVIR; DEHYDROTHYIMIDINE; ANTIVIRAL-DRUG; PATENTABILITY

CC 15900—IMMUNE SYSTEM; 21100—PHARMACOLOGY & CHEMOTHERAPY; 52100—INFECTIOUS DISEASE; 52500—VIROLOGY

NC NATL CANCER INST; BURROUGHS WELLCOME CO

Readers are referred to the free training courses and to the *BioBusiness Search Guide* for further details on content and methods of searching.

Some idea of the content of *BioBusiness* can be obtained from looking at the coverage of two subjects that have been of particular interest to AIDS researchers recently. Sample titles retrieved from the *BioBusiness* database on these topics are shown below.

Dispute over AIDS antibody test patent and the discovery of the virus thought to cause AIDS

GALLO PATENT ON ACQUIRED IMMUNE DEFICIENCY SYNDROME ANTIBODY TEST IS ISSUED.

U.S. PATENT AND TRADEMARK OFFICE TO ALLOW FORMAL PROCEEDINGS ON THE CONTESTED AIDS ANTIBODY TEST.

U.S. DENIES STEALING FRENCH AIDS VIRUS.

PASTEUR INSTITUTE SUES USA OVER ACQUIRED IMMUNE DEFICIENCY SYNDROME TEST ROYALTIES.

PASTEUR INSTITUTE AND NIH SEEK SETTLEMENT IN DISPUTE OVER DISCOVERY OF AIDS VIRUS.

DISPUTE OVER DISCOVERY OF AIDS VIRUS MAY SOON BE SETTLED.

U.S. AND FRANCE REACH SETTLEMENT AGREEMENT IN DISPUTE OVER DISCOVERY OF AIDS VIRUS.

AIDS PACT JOINS US, FRANCE.

Treatment of AIDS with zidovudine (AZT, azidothymidine)

BW-A 509 U.

INHIBITION OF REPLICATION AND CYTOPATHIC EFFECT OF HUMAN T CELL LYMPHOTROPIC VIRUS TYPE III/LYMPHADENOPATHY-ASSOCIATED VIRUS BY 3'-AZIDO-3'-DEOXYTHYMIDINE IN VITRO.

A QUIET DRUGMAKER TAKES A BIG SWING AT AIDS.

AN AIDS PILL THAT OFFERS HOPE.

DRUG COMPANIES FIGHT THE DRAGON.

ANTIVIRAL CHEMOTHERAPY AGAINST HUMAN T CELL LYMPHOTROPIC VIRUS TYPE III/LYMPHADENOPATHY ASSOCIATED VIRUS INFECTIONS.

AIDS DRUGS MAY PROVE WELCOME TONIC.

AZT MAY REVERSE SOME OF THE DAMAGE DONE BY AIDS.

WELLCOME PLC RECEIVES FRENCH GOVERNMENT CLEARANCE FOR RETROVIR.

AIDS DRUG APPROVED.

BURROUGHS AIDS DRUG SET FOR MARKET BY EARLY SPRING.

DEVELOPMENT OF VACCINE WILL DECREASE NEED FOR THERAPEUTICS.

Conclusion

Each of the BIOSIS services described has a different perspective on AIDS research. The total information resource is unique and both printed and online services provide powerful tools in identifying relevant information in this new area, which is moving and expanding with a rapidity unparalleled in the history of medical literature.

Acknowledgments

This article is based on a paper presented by Anne Cowie at the LAMHWLG's meeting 'AIDS information: needs and sources' on 29 October 1987.

Vital Information Ltd provides a help desk for users of BIOSIS products in Europe and the Middle East. The help desk gives advice on search strategy and the content and coverage of BIOSIS databases.

(Accepted February 1988)