

An Overview of Biomedical Literature Search on the World Wide Web in the Third Millennium

Prince Kumar¹, Roshni Goel², Chandni Jain³, Ashish Kumar⁴, Abhishek Parashar⁵, Ajay Ratan Gond⁶

¹MDS. Former Postgraduate Student, Department of Prosthodontics.* ²MDS. Former Postgraduate Student, Department of Endodontics, Institute of Dental Studies and Technology (IDST), Ghaziabad, India. ³MDS. Private Practitioner, Jabalpur, India. ⁴BDS. Postgraduate Student, Department of Prosthodontics.* ⁵BDS. Postgraduate Student, Department of Orthodontics.† ⁶Undergraduate Student.†

* Institute of Technology and Science (ITS)—Centre for Dental Studies and Research, Ghaziabad, India.

† King George Medical College, Dental Wing, Lucknow, India.

Abstract

Complete access to the existing pool of biomedical literature and the ability to “hit” upon the exact information of the relevant specialty are becoming essential elements of academic and clinical expertise. With the rapid expansion of the literature database, it is almost impossible to keep up to date with every innovation. Using the Internet, however, most people can freely access this literature at any time, from almost anywhere. This paper highlights the use of the Internet in obtaining valuable biomedical research information, which is mostly available from journals, databases, textbooks and e-journals in the form of web pages, text materials, images, and so on. The authors present an overview of web-based resources for biomedical researchers, providing information about Internet search engines (e.g., Google), web-based bibliographic databases (e.g., PubMed, IndMed) and how to use them, and other online biomedical resources that can assist clinicians in reaching well-informed clinical decisions.

Key Words: Biomedical Literature, Internet, Bibliographic Database, PubMed, Search Engine

Introduction

In the present era of information and technology, staying up to date with the latest advances in biomedical sciences is a major challenge for clinical practitioners. Because the amount of biomedical information doubles every five years, clinicians must have free and easy access to the current literature database for easy and effective evidence-based clinical decision-making [1]. Traditionally, there have been several systems available that condense and dispense the medical intelligence in to easily absorbable forms (e.g. medical and dental textbooks and dictionaries). However, these are frequently based on the synopsis and ideas of established experts and may not be refreshed with current information. In our day-to-day practice, we often come across a single and specific clinical problem that may be explained well in a single article. Until recently, the problem for many clinicians

has been accessing this information.

The World Wide Web or Internet has resolved these dilemmas to a large extent. Its rapid growth has created a boom in the field of biomedical investigation and research, although there is a long way to go before its full potential is realised [2,3]. On the click of mouse/button, the Internet offers quick and economical access to medical literature in the form of databases, dictionaries, journals, textbooks, dental product information, continuing education resources, and e-journals [4,5]. There are several eminent interlinking sources that retrieve these results such as government organisations, publishers, online health libraries, and commercial agencies.

Searching biomedical literature is a very organised and specific procedure. It requires systematic planning so as to develop a well-constructed clinical question or precise keyword. Unplanned

Corresponding contributor: Dr. Prince Kumar, Department of Prosthodontics, Institute of Technology and Science (ITS)—Centre for Dental Studies and Research, Ghaziabad, UP-201206, India; e-mail: princekumar@its.edu.in

and messy efforts may result in the retrieval of several, apparently irrelevant articles thus discouraging the professional to look further [6-8]. Web-based search engines are tools designed specially to search for information in the form of images, databases, journals and dictionaries. As these search engines are computer operated, they mostly search algorithmically. With these points in mind, it is therefore very important to understand how to access the information that is being searched for.

Internet Search Engines

A web-based search engine (e.g., Google) is a specially designed tool that retrieves user-friendly data from the Internet. To achieve a more refined and efficient search result, researchers must keep in mind the strengths and drawbacks of a particular search engine/database. Search engines that are frequently used for web-based search of biomedical literature include Google, Google Scholar, Yahoo, MSN.

Google search engine (<http://www.google.com>)

Google is the world's most popular and largest web-based search engine, receiving several hundred million queries each day through its various services [9]. A drawback is that it retrieves too much information, which is difficult to filter and which is not limited to biomedical research. It has the advantage, however, that it is easy to use and offers several user-friendly options including accessibility from cell phones. Another advantage is that it is updated regularly to include the new additions to the World Wide Web [10]. This allows users to retrieve very recently added biomedical information. In the Google home-page settings, JavaScript may be enabled to allow easy access to most unbounded or linked websites (Google Analytics™).

Online reprint request (ORR) via Google

Google enables users to make an online reprint request (ORR) for unavailable medical literature to the concerned author/authority. An ORR of the article of interest (explored via Google Search) can be made in the form of an e-mail containing the relevant information—name of journal, name(s) of authors and co-authors, year of publication, volume and page numbers.

Yahoo search engine (<http://www.search.yahoo.com>)

As reported in January 2012, Yahoo search was the second largest search engine available on the Web by query (volume: 5.82%) after its competitor

Google (volume: 83.43%) [11]. It offers numerous user-friendly features (e.g., shortcut key control, pull-down menus, automatic suggestions) that make it simple and popular.

Bing search engine (<http://www.bing.com>)

Bing (<http://www.bing.com>) is a novel web-based search engine that passes on Yahoo and Microsoft databases.

Google Scholar (<http://www.scholar.google.com>)

Google Scholar is the simple way to perform a broad search for scholarly articles. It searches across many resources and databases including articles, theses, textbooks, peer-reviewed papers, commercial academic publishers, universities, professional societies, and abstracts. Most of Google Scholar's database resources derive from a crawl of full-text journal content provided by both open source and commercial publishers [12].

Advantages of using Google Scholar

1. It has an advanced and quick search facility where literature/articles can be selected by author, publication, journal, date and subject. In contrast with a Google search, Google Scholar returns significantly fewer results because it filters the result by the stated search preference. For example, a Google search for "Implant and O ring attachment" would return about 204,000 results whereas a similar search in Google Scholar with the restricted date of 2011 yields only 6340 results (as on 6th February 2012).
2. It provides a single platform to explore research papers/articles directly or via online libraries, abstracts and citations, etc.).
3. It helps users to shortlist articles/papers on the basis of year of publication, journal, area of interest, citation, etc.
4. It ranks and lists results according to how relevant they are to the search query. The most relevant references should theoretically appear at the top of the page.
5. Federated search results: Web-based subordinates of Google Scholar incorporate the processed and indexed fresh information from the World Wide Web making it single pool for search hits.

Limitations of Google Scholar

Google Scholar has certain limitations:

1. It does not refer/index all commercial database agencies for biomedical research materials.

2. Like other medical search engines (e.g., PubMed), Google Scholar does not provide the list of journals indexed in its database. Furthermore, it does not identify where it finds its sources, so one cannot have confidence in its reliability.
3. It does not perform as well for older publications. Pauly and Stergiou (2005) [13] reported that Google Scholar had less than half of the citations for a specific set of scientific research papers published in a variety of disciplines between 1925-1989; however, for papers published in the period 1990-2004, both sources gave similar citation counts. Meho and Yang (2007) [14] found the majority of the citations from journals and conference papers in Google Scholar to date from after 1993.
4. Some articles are available only as abstracts; full access is possible only through a payment gateway.

Internet-Based Bibliographic Databases

Bibliographic databases are well-organised collections of descriptive information on author, journal, title of paper, source, date of publication. They have made searching the medical literature easier and more accessible to most researchers, clinicians and professionals [15]. The most commonly referred databases are Medline, PubMed, the Cochrane Library, and Embase.

National Institutes of Health (NIH)

The National Institutes of Health (NIH) is an agency of the U.S. Department of Health and Human Services (Government of United States). NIH is well known for its contribution to biomedical and health-related research. It comprises 27 separate institutes, centres, and offices, including the National Library of Medicine (NLM). The NLM collects, organises, and makes available biomedical science information to investigators, educators, and practitioners. It established the National Center for Biotechnology Information (NCBI), which is a central repository of biological information and includes the PubMed literature database. The electronic web-based reference system common to NIH, NLM, and NCBI is considered to be one of the “crown jewels” of medical research.

PubMed and MEDLINE

The portal <http://www.nlm.nih.gov> is the online gateway to the National Library of Medicine. It is

often referred to as the “PubMed link”; PubMed is the principal online bibliographic citation database and open access biomedical search engine for literature. It is free, user friendly, and the most commonly used online database for biomedical literature search [16]. It is part of Entrez™ information retrieval system [17]. This gateway allows concurrent search in the PubMed and MEDLINE (Medical Literature Analysis and Retrieval System Online) databases, as well as in other NLM databases in different fields of biomedical sciences (genetic search databases and clinical trials registered with the NLM). MEDLINE, a component of PubMed, contains journal citations and abstracts for biomedical literature from around the world. PubMed provides free access to MEDLINE and links to full text articles when possible. As the NLM website also searches in MEDLINE databases, it is imperative to attempt medical literature search in MEDLINE database via the NLM link [18]. When compared to the Google search, a PubMed search of medical literature is less time consuming. Additionally, it permits users to search for more refined results using a special tool “Clinical Queries and Systematic Reviews search filters”, present on the NLM NIH homepage. As the clinical queries are authenticated search strategies, they map out articles on the diagnosis, aetiology, or prognosis related to that question [19].

NLM database

The NLM database contains more than 21.52 million citations (as on 6th February 2012) from approximately 5,500 journals and research publications. Information about the journals indexed in PubMed is found in the NLM Catalog present on the PubMed homepage. On average, about 500,000 new records are added each year to NLM database. With the recent enhancement in the NLM database, it contains records as far back as 1951. References of publications prior to 1950 are available from the print version, *Index Medicus*. Amazingly, this contains records of the years 1865 and 1809. One may enter “1800:2100[dp]” into the PubMed search window to see the total size of the NLM database at that particular moment.

PubMed Central

PubMed Central (available at: (<http://www.ncbi.nlm.nih.gov/pmc>)) is a free digital database of full-text scientific literature in biomedical and life sciences. Developed and managed by the NCBI, it provides openly available peer-reviewed scientific

research. It does not include any unreviewed research papers. Participating publishers may, however, delay release of their articles for reasonable period after publication (often six months).

Medical Subject Headings (MeSH)

The search results of the NLM database are based on a vocabulary called Medical Subject Headings (MeSH). MeSH is a well designed vocabulary developed by NLM to enhance accuracy and efficiency in PubMed/MEDLINE searches. MeSH descriptors are available in both alphabetic and hierarchical format. The use of MeSH terms in conjunctions of certain limits (such as date of publication) improves the quality of search. The MeSH website (<http://www.nlm.nih.gov/mesh>) is the central access station for additional information regarding MeSH.

Inclusion criteria for journals in NLM database

Inclusion is based on the decision of a panel, the Literature Selection Technical Review Committee. The LSTRC evaluates such aspects as the scientific policy and quality of the publication, quality of content, quality of editorial work, and standard of editorial review.

PubMed tutorial

An online training programme is available, specially designed and recommended for new users. It demonstrates how to search articles in a quick, precise and effective manner.

PubMed/MEDLINE search strategy

For optimal search results from PubMed/MEDLINE, the following search strategy may be applied:

1. Identify the key concept related to query.
2. Find an alternative term (from MeSH) for this key concept.
3. Start your search using the selected MeSH.

When the key concept is unclear, the PubMed search engine may be used in the reverse manner to search one's concept from MeSH.

Key points of PubMed/MEDLINE Search

Stopwords and Superfluous Words

When attempting a search in PubMed/MEDLINE, it is important to remove "stopwords" or superfluous words from the search phrase. A list of such terms is available on the PubMed homepage. For example, if one was intending to search for "Overdenture in the management of partial edentulism", "in", "the", "of" are superfluous words. The correct key concept and MeSH words for the

above search would be "Overdenture management partial edentulism".

Boolean Logic

PubMed uses "Boolean logic" to search for MeSH terms [20]. It refers to the logical relationships among search terms. When a search query is unrestricted, it assumes that all the terms that are mentioned need to be present. Boolean operators (AND, OR, NOT, always in capital letters) can be used to combine or exclude search terms in PubMed to get more structured results.

Associates of PubMed/MEDLINE

GoPubMed (<http://www.gopubmed.com>)

GoPubMed is a knowledge-based online (medical) search engine for biomedical texts. The Gene Ontology (GO) and MeSH serve as "Table of contents" in order to structure the millions of articles in the MEDLINE database. GoPubMed database allows its users (biologists and medical professionals) to locate related search results significantly faster than by using PubMed.

MedlinePlus (<http://www.medlineplus.org>)

MedlinePlus is a free, open access electronic portal that provides consumer health information for patients, families, and health workers. It is an online medical encyclopaedia that is managed and powered by NLM and NIH. MedlinePlus includes over 4,000 articles about diseases, tests, symptoms, injuries, and surgical procedures. Over 150 million people from around the world use MedlinePlus each year [21]. It also provides an online medical dictionary and news service for medical news reports.

Indian Internet-Based Bibliographic Databases

Medknow Publications (<http://www.medknow.com>)

Medknow Publications publish nearly 170 print and online journals. They provide a "fee-less-free" model of open access publishing, enabling immediate free access to electronic editions of the journals. Researchers just have to access the web site, fill the key word that they require, and search. Every journal published by Medknow has its independent website. The open access strategy has resulted in more than a half a million in a month all Medknow journals.

IndMed (<http://www.medind.nic.in>)

IndMed database has been designed and developed

by the Indian Medlars Center, an autonomous body under the joint control of National Informatics Centre and Indian Council of Medical Research (NIC and ICMR, Government of India). The main endeavour of the IndMed bibliographic/indexing service is to facilitate easy and free access to the Indian biomedical literature. The IndMed database includes peer-reviewed popular medical and dental journals published from India. It provides a full-text access web link to about 77 Indian journals, indexed from 1985 onwards.

Other Non-Government Internet-Based Bibliographic Resources

Cochrane Library (<http://www.thecochranelibrary.com>)

The Cochrane Library is a collection of databases that contain different types of high-quality independent evidence to inform health care decision-making in medical, dental and other health care specialties. They include systematic reviews and a central register of controlled clinical trials. The collection may be accessed through the Wiley Online system.

Embase (<http://www.embase.com>)

Embase is an online information source of published literature that includes over 24 million indexed records (from 1947) and more than 7,500 indexed peer-reviewed journals (more than PubMed/MEDLINE). It has exclusively subscribed biomedical database that may be accessed online through a number of database vendors, hosted by Elsevier. Using Embase, it is possible to track, for example, drug-specific adverse events published in the literature.

WebMD (<http://www.webmd.com>)

WebMD is an American corporation that provides valuable health information, tools for managing health, and support to those who seek information. WebMD is primarily known for its health portal, which has information regarding health and health care professionals. Other services from WebMD include creating and maintaining up-to-date medical reference content databases and providing health news.

Medscape (<http://www.medscape.com>)

Medscape is an online resource for medical students and health care professionals. It features peer-reviewed original research articles in the biomedical sciences, including dentistry. Medscape offers

clinicians medical information and user-friendly educational tools. It requires a one-off registration. Additional features include daily medical news, journal commentary, expert columns, patient education articles, major conference coverage, and drug information including a drug database (Medscape Drug Reference or MDR).

Medical Matrix (<http://www.medmatrix.org>)

Medical Matrix is a web-based guide to medical and dental resources, available free of cost. It also provides external links to other websites of particular interest and a comparison chart to evaluate the websites. Apart from medical journals and literature, its database contains more than 70 dental journals.

Commercial Web-Based Resources: Scholarly Research Databases

Several commercial resources, such as web-based libraries and specialist websites (dentistry, educational, continuing education), provide very valuable additions to the ever-growing list of online resources. Scholarly research databases are the commercial aggregators of the premium full-text content of a vast range of biomedical and health literature, enabling users to get access on a single click. They provide a dropdown list or external web link to various peer-reviewed indexed journals via online commercial merchants (e.g., Blackwell Publishing, Elsevier, John Wiley & Sons, Harcourt, Wolters Kluwer, Springer Science+Business Media, Science Direct). On a subscription basis, they provide access to national universities, research institutes, professional colleges (medicine, dentistry, nursing, pharmacy, and public health), teaching hospitals, government offices and national medical libraries. Some of the prominent scholarly research databases include: Caspur, DOAJ, EBSCO, Health & Wellness Research Center, HINARI, Index Copernicus, Journal Seek, MANTIS, OpenJGate, Primo Central, ProQuest, SCOLAR, SIIC.

Other Online Resources

Free online journals and books

(<http://www.freemedicalbooks.org>)

Medical texts are the foremost basic and traditional means that lay down the basic concepts of the subject. Several hundred medical/dental journals and e-books provide free and full-text access to their contents via the Internet. Other information regard-

ing e-textbooks, CD-ROMs, software, animations and videos are available on the web.

How To Cite A Web Page Reference

Any electronic data/material that is accessed or downloaded from the Internet may be used as citation in research documents. The Vancouver style of referencing/citing has been one of the worldwide accepted and standard protocols in academic papers since 1978 [22]. For preparing a web citation according to the Vancouver guidelines, the following details should be gathered:

- Name of author/s
- Title of article
- Name of website
- Name of publisher and place of publication (if available)
- Date accessed
- Date updated
- Online URL

The above data are arranged in a logical manner to frame the elements of web citation, namely:

Author. Title of publication [type of medium: Internet]. Place of publication (if available): Publisher (if available); Date of publication: year month day (supply year if month and day not available) [updated year month day; cited year month day]. Available from: web page URL.

Example:

Malathi M, Mohan Thappa DM. The intricacies of impact factor and mid-term review of editorship [Internet]. India: Medknow; 2012 [updated 2011 Feb 05; cited 2012 Feb 06]. Available from: <http://www.ijdv1.com/article.asp?issn=0378-6323;year=2012;volume=78;issue=1;spage=1;epage=4;aulast=Malathi>

Conclusion

This paper has sought to review the search engines

References

1. Mattox DE. Welcome to ARCHIVES CME. *Archives of Otolaryngology—Head and Neck Surgery*. 2000; **126**: 914.
2. Jenner A. How has the Internet affected dentistry? [Research summary of: Chestnutt IG, Reynolds K. Perceptions of how the Internet has impacted on dentistry]. *British Dental Journal* 2006; **200**: 149.
3. Schleyer TK, Spallek H, Torres-Urquidy MH. A profile of current Internet users in dentistry. *Journal of the American Dental Association*. 1998; **129**: 1748-1753.
4. Doyle DJ, Ruskin KJ, Engel TP. The Internet and medicine: Past, present, and future. *Yale Journal of Biology and Medicine*. 1996; **69**: 429-437.

and databases that are likely to be of use to colleagues working in the fields of biomedicine and more specifically dentistry, and are available via the World Wide Web. It has not attempted to explain how or why they can be used in any detail. A number of other publications (both papers and books) provide significant aid in these areas [23,24]. It should be remembered that, for copyright and political reasons, there is not always universal, unlimited access to all material available via the Internet and that copyright applies to material published on the Internet in the same way that it applies to printed material. Consequently, a researcher must be aware of the strengths and confines of a particular search engine and database while exploring the literature in order to apply their best possible use in the field of medical and allied science researches.

Contributions of each author

- PK planned and designed the paper, wrote it and reviewed it.
- RG co-planned and designed the paper and assisted in its writing and review.
- CJ conducted a literature review and assisted in the writing, review and editing of the paper.
- AK conducted a literature review, acquired data and assisted in the editing and review of the paper.
- AP conducted a literature review, acquired data and assisted in the editing and review of the paper.
- ARG edited and reviewed the manuscript.

Statement of conflict of interest

As far as the authors are aware, there is no conflict of interests.

5. Jethwani KS, Chandwani HS. The internet: Revolutionizing medical research for novices and virtuosos alike. *Journal of Postgraduate Medicine*. 2008; **54**: 49-51

6. Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. *Journal of the American Medical Association*. 1993; **270**: 2093-2095.

7. Bin Ghouth AS. Using computer and Internet for medical literature searching among medical students in Hadramout University, Yemen. *Online Journal of Health and Allied Sciences*. 2008; **7**(1): 6. Available from: <http://www.ojhas.org/issue25/2008-1-6.htm>

8. Stewart MG, Moore AS. Searching the medical literature. *Otolaryngologic Clinics of North America*. 1998; **31**: 277-287.
9. Chang P, Hou IC, Hsu CL, Lai HF. Are Google or Yahoo a good portal for getting quality healthcare web information? *AMIA Annual Symposium Proceedings*. 2006: 878.
10. Al-Ubaydli M. Using search engines to find online medical information. *PLoS Medicine*. 2005; **2**: e228.
11. Netmarketshare. Search Engine Market Share [Internet]. United States: Net Applications; 2012 [updated 2012 Feb 06; cited 2012 Feb 06]. Available from: http://marketshare.hitslink.com/search-engine-market-share.aspx?qprid=4&qpcustom=*2&qpcustomd=
12. Henderson J. Google Scholar: A source for clinicians? *Canadian Medical Association Journal*. 2005; **172**: 1549-1550.
13. Pauly D, Stergiou KI. Equivalence of results from two citation analyses: Thomson ISI's Citation Index and Google Scholar's service. *Ethics in Science and Environmental Politics*. 2005; **5**: 33-35.
14. Meho LI, Yang KA. New era in citation and bibliometric analyses: Web of Science, Scopus, and Google Scholar. *Journal of the American Society for Information Science and Technology*. 2007; **58**: 2105-2125.
15. Marlborough HS. Accessing the literature: Using bibliographic databases to find journal articles. Part 1. *Primary Dental Care*. 2001; **8**: 117-121.
16. Watwood C. NLM Gateway. *Journal of the Medical Library Association*. 2006; **94**: 361-363.
17. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, *et al*. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*. 2009; **37**: 5-15.
18. Bernstam EV, Herskovic JR, Aphinyanaphongs Y, Aliferis CF, Sriram MG, Hersh WR. Using citation data to improve retrieval from MEDLINE. *Journal of the American Medical Informatics Association*. 2006; **13**: 96-105.
19. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *Journal of the American Medical Informatics Association*. 1994; **1**: 447-458.
20. Kiley R. Medical databases on the Internet: Part 1. *Journal of the Royal Society of Medicine*. 1997; **90**: 610-611.
21. MedlinePlus. MedlinePlus Statistics [Internet]. United States: National Institutes of Health; 2012 [updated 2012 Feb 05; cited 2012 Feb 06]. Available from: <http://www.nlm.nih.gov/medlineplus/healthstatistics.html>
22. International Committee of Medical Journal Editors (ICMJE). Uniform requirements for manuscripts submitted to biomedical journals [Internet]. 2010 [updated 2010 Apr; cited 2012 Feb 06]. Available from: http://www.icmje.org/urm_full.pdf
23. Farbey R, Eaton KA, Santini A. An introduction to research for primary dental care clinicians. Part 1. The first three stages in a research project. *Primary Dental Care*. 2010; **17**: 168-172.
24. Eaton KA, Wiles L, O'Malley D. An introduction to research for primary dental care clinicians. Part 9: Stage 10. Writing and disseminating the results. *Primary Dental Care*. 2012; **19**: 85-91.