Chapter 2
Using Medical Search Engines with a Special Focus on Google

There are billions of pages on the World Wide Web and search engines help us find the information we are looking for. Having so many pages with a huge amount of information online is the biggest advantage and disadvantage of the Internet at the same time. Instead of manually browsing the websites, search engines point us to the pages that contain the key words we are looking for.

How Do Search Engines Work?

Search engines use robots (also known as spiders or crawlers) which are programs that can automatically follow hyperlinks from one document to another one sending information about new sites back to its main repository to be indexed. Due to the dynamic nature of many websites, robots keep an index of words they find and the place where they find them; and do this indexing process regularly even on websites that had previously been catalogued. It means new websites or new content on known websites have to be indexed first in order to become accessible through search engines.

The Short History of Search Engines

It is not surprising using e-mails and search engines are the two most common practices online [1]. Due to this fact, the first widely adopted search engine was launched back in 1995

B. Meskó, *Social Media in Clinical Practice*,
DOI 10.1007/978-1-4471-4306-2_2,

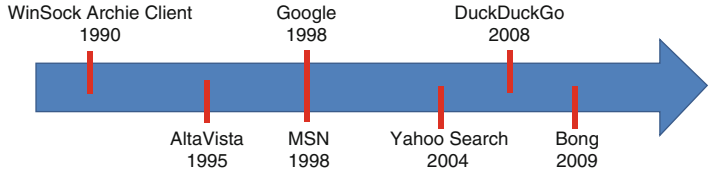

FIGURE 2.1 Timeline of the appearance of search engines

under the name AltaVista (http://www.altavista.com/) and other search engines were created soon after that (Fig. 2.1). As of 2013, the most popular search engines globally are Google (https://www.google.com/), Bing (http://www.bing.com/), Yahoo (http://www.yahoo.com/), Ask (http://ask.com/) and AOL (http://www.aol.com/) [2].

Out of these search engines, Google seemed to attract the most people (over one billion users as of 2012) [3]. It was founded by Larry Page and Sergey Brin on September 4, 1998 while they attended Stanford University. The name Google originates from a misspelling of the word "googol" accounting for the number one followed by one hundred zeros.

The Basics of Searching Online

On the main page, search queries can be inserted into the box and search results can be obtained by clicking on the Google Search button. A faster alternative is clicking on the "I'm Feeling Lucky" button which immediately takes the user to the first entry of the search results assuming that one is the best possible resource (Fig. 2.2.).

A typical page of search results features a sidebar on the left or on the top showing additional search functions such as searching only for images, maps, videos or news, among others. The number of search results and the time needed to come up with those are listed below the search box that contains the search query. For each search results, the title of the website, the URL and a short description of that particular website are shown (Fig. 2.3).

FIGURE 2.2 The main page of Google.com (Google and the Google logo are registered trademarks of Google Inc., used with permission)

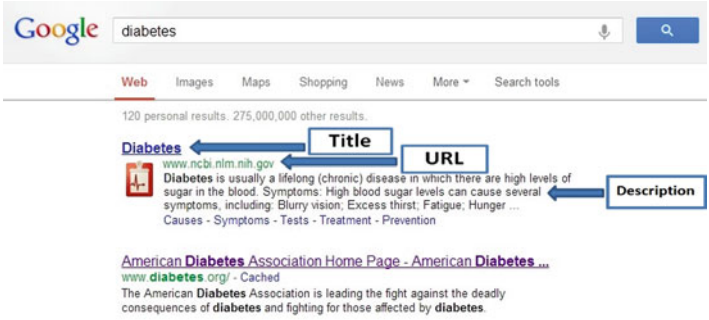

FIGURE 2.3 A search query on "diabetes" on Google.com (Google and the Google logo are registered trademarks of Google Inc., used with permission)

The order of the search results is determined based on the so-called page rank of the websites. Each website that is indexed by the robots of Google gets a score ranging from 1 to 10. This is called pagerank; and the more and better connections the website has with other websites, the higher the pagerank is making it appear in the search results.

There are numerous tricks and operators that can facilitate the use of search engines such as removing words from the query or searching only on a specific website; or for a specific filetype. Examples are summarized in Table 2.1.

In cases when rules do not permit the use of search engines that track user information DuckDuckGo

Table 2.1 Search operators in Google.com

Search query	What it does
diabetes OR allergy	Search for websites containing either of the words, otherwise it would typically show sites that contain both
diabetes AND allergy	Search for websites that contain both words
-diabetes	It shows websites that do not contain the search term
site:www.example.com diabetes	It searches for the keyword on only the identified website
filetype:doc diabetes	It searches documents uploaded online featuring the search term. File type can be pdf or ppt, among others
"diabetes treatment options"	It searches for exactly this expression in this word order
5 kg in pound	Unit and money conversions can be performed
inurl:WHO diabetes	It searches for diabetes in all web pages that have "WHO" in their URLs
intitle:diabetes	It searches for diabetes in the titles of web pages and documents instead of the body of the text

(http://www.duckduckgo.com) serves as an alternative being a search engine that emphasizes privacy and does not record user information therefore all users receive the same search results for a given search query.

What Is Search Engine Optimization (SEO)?
SEO is a common term used all around the web, and it refers to the process of improving the visibility of a website in search engines by inserting the right key words in the right places in the content.

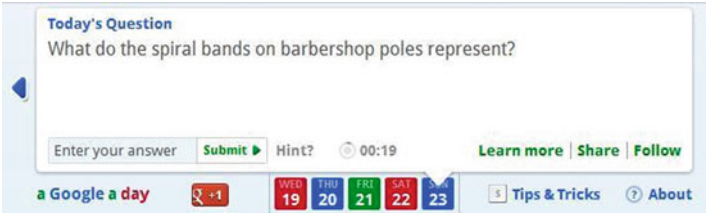

FIGURE 2.4 A Google A Day provides daily search tasks with solutions (Google and the Google logo are registered trademarks of Google Inc., used with permission)

How to Get Better at Searching Online?

A list of suggestions and tips that can help obtain better search results:

- Be as precise as possible.
- Focus only on the keywords of your query, the rarest words in the sentence or question.
- Fasten searches by using the right operators (e.g. AND, OR, intitle, author, etc.).
- Narrow your search results in more steps, first results do not necessarily lead to the information you are looking for.
- Think as if you published the information you are looking for and what kind of keywords you would have been using.

If you are looking for PDF documents that mention diabetes in their title but do not contain the word treatment in their content, the best way to do so is to search for "filetype:PDF intitle:diabetes –treatment" which will show only relevant results. This example underscores the importance of using different search operators at once with strategy.

An ideal strategy for getting better at online searches is practising more by using the right queries, keywords and operators in order to save time and effort every day. Methods for structured learning include "A Google A Day" (http://www.agoogleaday.com) that provides questions and tasks related to special search queries every day by showing the solution in the form of specific search queries and keywords as well (Fig. 2.4).

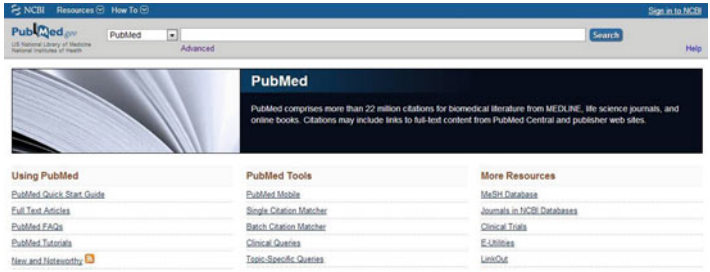

FIGURE 2.5 The main page of http://www.Pubmed.com

Search Engines in Medicine and Research

Pubmed.com

Certain search engines index the whole world wide web while others focus on topics and sub-topics such as medicine and research. The most prominent example is Pubmed.com which was launched by The United States National Library of Medicine at the National Institutes of Health in 1996 and includes over 22 million citations serving as a repository of the biomedical literature [4] (Fig. 2.5).

Instead of searching for a specific field of interest or research topic regularly, making these search queries automatic facilitates our work flow for which there are basically two methods.

Registering a free account on Pubmed.com and clicking on "Save Search" below a search query sets up e-mail alerts to receive the latest publications focusing on these keywords automatically (Figs. 2.6, 2.7, and 2.8).

The second method includes doing searches for keywords and clicking on the RSS button below it which makes it possible to create RSS feeds out of the search query which can be added to RSS readers therefore all the upcoming papers focusing on our field of interest will be syndicated to our RSS reader (Fig. 2.9).

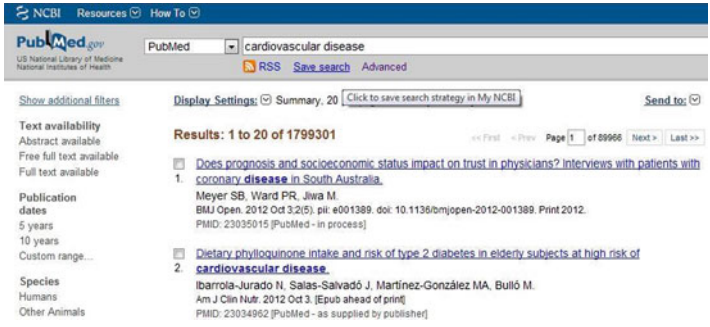

FIGURE 2.6 A common search results page on Pubmed.com. Filtering options can be found in the left sidebar

FIGURE 2.7 Dashboard after clicking on "Save search" on Pubmed.com

Third-Party Pubmed Tools

As the US National Library of Medicine released the application programming interface (API) through which other services and online platforms can access the database, third-party Pubmed tools were introduced such as GoPubmed (http://www.gopubmed.org), BibliMed (http://www.biblimed.com) or Pubget (http://pubget.com). These

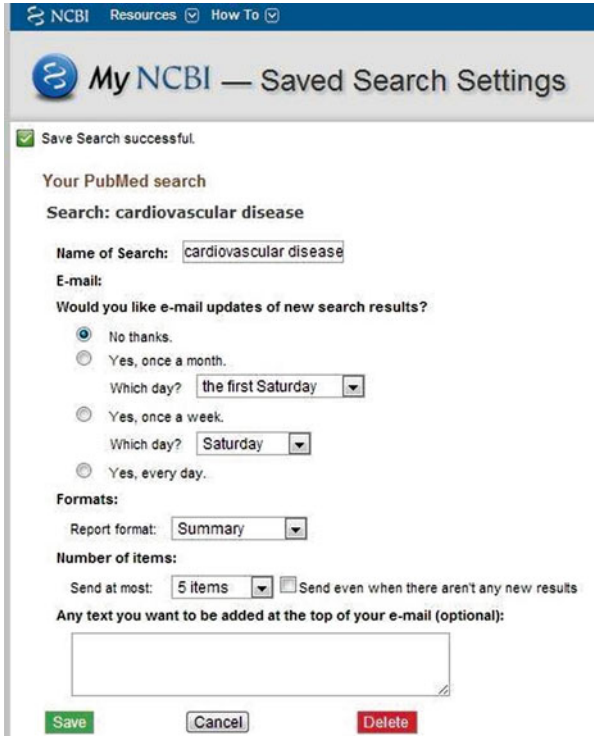

FIGURE 2.8 Settings using the "Save search" button

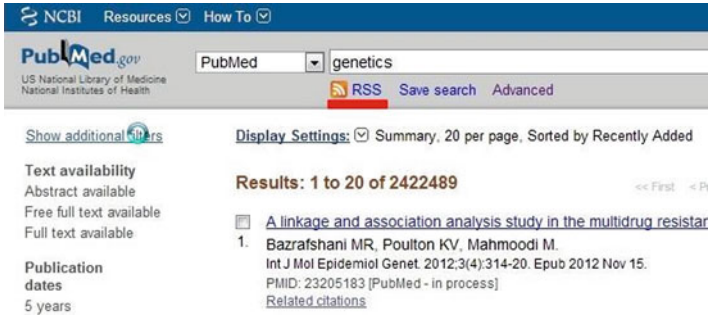

FIGURE 2.9 The RSS button is below the search box on Pubmed.com

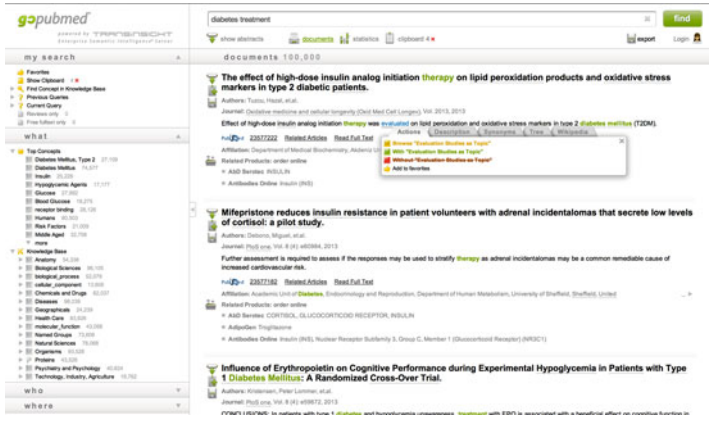

FIGURE 2.10 The page of search results on Gopubmed.com with the additional filtering options in the left sidebar. A third-party Pubmed tool mixes the database of NCBI and the technological advantages of other search engines

are knowledge-based search engines specifically designed to search for biomedical texts significantly faster than Pubmed (Fig. 2.10).

These make it possible to narrow search results by different identifiers including authors, sub-topics or publication date in an easy-to-use interface.

Google Scholar

Launched in 2004, Google Scholar (http://scholar.google.com) is a search engine that indexes the full text of scholarly literature, most peer reviewed journals, scholarly books and other non-peer reviewed journals. Google Scholar allows users to search for digital or physical copies of articles, technical reports, preprints, theses, books, and patents (Fig. 2.11).

It works similarly to other search engines but there are several tools for customization of the search results such as

FIGURE 2.11 The search box on Google Scholar. Basic settings include searching in patents as well besides research articles (Google and the Google logo are registered trademarks of Google Inc., used with permission)

filtering by date, relevance or document type. It also shows how many times a paper were cited by other papers or whether there are other versions available of a particular article (Fig. 2.12).

Similarly to the search engine of Google, search operators can be used in Google Scholar as well in order to make more customized search queries (Table 2.2).

In order to receive automatic updates of the newest additions focusing on a specific search query, alerts can be set up by using the "Create alert" function in the bottom left corner. By providing a valid e-mail address, updates will be sent from time to time (Fig. 2.13).

WolframAlpha

Wolfram Alpha (http://www.wolframalpha.com) developed by Wolfram Research is an online search engine that answers factual queries by computing the answer from structured data. Numerous medical databases have been uploaded in different fields of medicine from public health to laboratory results [5]. Instead of providing a list of relevant websites and articles, WolframAlpha aims at giving the final answer right away.

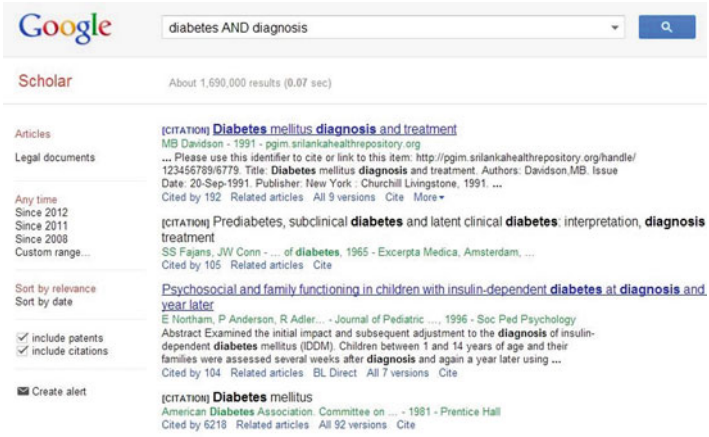

FIGURE 2.12 The search results' page on Google Scholar. Filtering options are shown in the left sidebar and citations of each paper can be accessed below the results (Google and the Google logo are registered trademarks of Google Inc., used with permission)

TABLE 2.2 Additional search operators that could be used in Google Scholar as well

Search operator	What it does
+word	Adds a word to the search
-word	Removes a word from the search
crohn OR diabetes	Searches for articles containing either of the words
author:Brown	Searches for authors named Brown
intitle:diabetes	Searches for articles that feature the word diabetes in their titles

Examples of medical search queries on WolframAlpha (with the exact search query):

- compute estimated risk of heart disease ("heart disease risk 50 year-old male")
- get gender and age-specific information about a test result ("creatinine = 0.9 mg/dL, adult male")

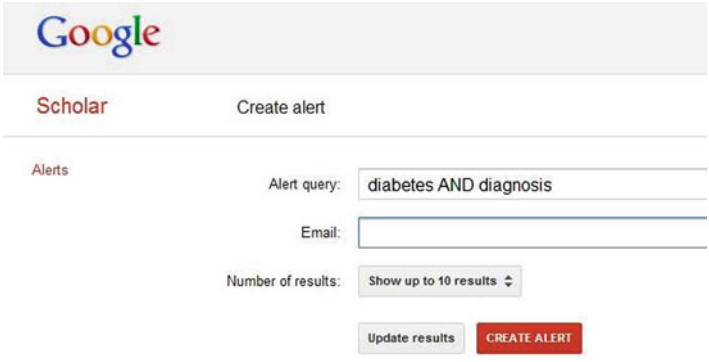

FIGURE 2.13 E-mail alerts can be set up using Google Scholar (Google and the Google logo are registered trademarks of Google Inc., used with permission)

- compute various body statistics based on height, weight, etc. ("BMI 5′10″, 165 lb")
- get an overview of health care costs in a country ("Canada healthcare expenditures")

Self-Test

1. Which search engine should I use?
   There are minor differences between the most popular search engines, give the top three a try and you will be able to choose. Speed and safety can be the major features.
2. How can I search for PDF documents in diabetes?
   Use the relevant operators such as "filetype:pdf diabetes".
3. How can I receive updates in my academic field of interest automatically?
   Both Google Scholar and Pubmed offer e-mail alerts or RSS feeds.

Next Steps

1. Do a couple of search queries for your name, workplace and field of interests.
2. Do the same search queries in the three major search engines (Google, Bing, Yahoo).
3. See what kind of search operators you could use [6].
4. Pratice your skills for a few days or weeks on "A Google a Day".

Key Points

- Search engines help find the content, websites or pieces of information you need.
- Focus on keywords and use the search query operators.
- Set up automatic alerts on Pubmed in order to receive the latest updates in a field of interest.
- Use Google Scholar and Pubmed together for the best results in academic search.
- WolframAlpha provides excellent computational solutions in medicine and healthcare.

References

1. Pew Internet: the social life of health information. 2011. http://pewinternet.org/Reports/2011/Social-Life-of-Health-Info/Summary-of-Findings.aspx. Accessed 8 Aug 2012.
2. Top 15 Most Popular Search Engines. http://www.ebizmba.com/articles/search-engines
3. Google's new record, 1 billion visitors in May. http://itsalltech.com/2011/06/22/googles-new-record-1-billion-visitors-in-may/. Accessed 4 Sept 2012.
4. http://en.wikipedia.org/wiki/PubMed. Accessed 28 Jan 2013.
5. http://www.wolframalpha.com/examples/HealthAndMedicine.html. Accessed 24 Jan 2013.
6. http://www.googleguide.com/advanced_operators.html. Accessed 20 Mar 2013.