How Do General Public Search Online Health Information?

(Results of the survey conducted in March-April, 2011)

Natalia PLETNEVA, Alejandro VARGAS, Célia BOYER
Health on the Net Foundation, Geneva, Switzerland

July 2011
This project is supported by the European Commission under the Information and Communication Technologies (ICT) Theme of the 7th Framework Programme for Research and Technological Development.

The research leading to these results has received funding from the European Community's 7th Framework Programme (FP7/2007-2013) under grant agreement nº257528.
Khresmoi → Multi-lingual multi-modal search & access for biomedical information and documents.

- Information extraction from biomedical documents, including crowd sourcing and active learning improvements, estimation of the level of trust level and target user expertise.
- Analysis and indexing for medical images 2D (X-Rays) - 3D (MRI, CT) - 4D (MRI with time).
- Linking information extracted from un-/semi-structured biomedical texts and images to structured information in knowledge bases.
- Cross-language search including multi-lingual queries, and returning machine-translated pertinent excerpts.
- Adaptive user interfaces to assist queries and display results via ergonomic interactive visualizations.

The survey is part of the European KHRESMOI Project 2010-2014
Survey Overview

- **Time frame:** from the 9th of March till the 30th of March English and French versions available, and from the 31st of March till the 27th of April, 2011 English, French, Spanish and German versions available on HON web site.

- **Languages:** English, French, Spanish and German.

- **385 participants:** from 42 countries around the world, majority – France (23 %), Spain (14 %), the USA (10 %) speaking French (34 %), Spanish (19 %), English (17 %) and German (13 %).

- **Target audience:** Individuals looking for online health information both males (47%) and females (53 %) mostly aged between 20 and 59 (85 %), 79 % have completed higher education (43 % of them hold Master and 28% hold PhD) working in healthcare (30 %) or IT (21 %) area.

- **Methodology:** Non-probabilistic sampling, promotion and dissemination via HONcode certified web sites, social media and KHRESMOI partners. The main promotion started in the end of the March.
Executive Summary

- Wi-Fi access and a popularity of mobile devices is growing.
- The Internet is the second source of information after physicians, and it is more likely to be used due to its accessibility and easiness.
- A majority of e-patients are using a general search engine to find the answers to their health queries.
- The most important characteristics of a search engine are relevance and trustworthiness of the results. This appears to be the main problem, as current results do not satisfy these requirements.
- The ideal representation of the information is a categorization of results into different groups with the possibility to perform this automatically or manually.
- Tools that are highly appreciated by the respondents are advanced search, medical dictionary/thesaurus, suggested relevant topics, image search and risk assessment tools.
Detailed Findings
Use of the Internet

- **Internet history:** 90% of respondents declared that they have been using it for more than six years, thus 84% rate themselves as good or professional users.
- **Internet access:** 100% stated they have a regular access and 95% are using it on a daily basis, the remaining 5% several times a week.
- **Time spent online:** 45% spend from two to four hours a day, 25% spend from four to eight hours, and 13% spend even more than eight hours.
- **Connection to the Internet:** 62% of the respondents use Wi-Fi and 47% a modem.
- **Devices used:** 71% have PCs, 68% laptops, and 28% mobile phones, at their disposal; and 5% are using other devices.
- **Mobile devices use:** the most widely used are laptops (56%) and smartphones (36%); tablets and netbooks have not yet gained much popularity in our audience.
How Often the Internet Is Used to Search for Health Information

- 24%: Up to six times a day
- 25%: At least once a day
- 16%: Few times a week
- 16%: Once a week
- 11%: Few times a month
- 8%: Once a month
- 8%: Other
Sources Preferably Used for Health and Medical Information

The most important sources:
- 82% rate physicians as important or very important
- 71% rate the Internet

Mentioned among “others”:
- Patients organizations
- Medical/scientific/peer-review journals
When Internet Is Used Rather Than Other Sources of Information

Accessibility from home: some mentioned the office, and 85% considered it important or very important.
Focus of the Search Activity

- General information about health issues
- Long term, chronic disease
- Healthy lifestyle and nutrition
- Short-term (up to 2 weeks), acute disease
- Kids health
- Elderly health and care
- Other
## Types of Searched Online Health Information

<table>
<thead>
<tr>
<th></th>
<th>Always</th>
<th>Often</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment/Therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detailed disease description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General information about a disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drugs information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Therapy side effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scientific articles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical trial information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients and/or support groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-diagnostic tools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abbreviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Looking for Drug Information: What Kind of Content?

- Side effects
- Drugs contraindications
- Drug safety, toxicity
- Drugs interaction
- Drug description
- Drug dosage
- Drug prices
- Drug purchase

- Always
- Often
- Sometimes
- Rarely
- Never
How Often Online Health Information Sources Are Used

- Search engines (Google, Yahoo, MSN, etc.)
- Web sites providing health information (Hospital, University, Pharma, ...)
- Wikipedia
- Specialized search tool (HONselect, Medline Plus etc.)
- Links from a health web site
- Web sites suggested by a health care provider
- Forums and blogs
- Portal web sites (Yahoo health, Google Health, MSN Health, etc.)
- Web sites suggested by a friend
- Facebook or other social networks

Some participants mentioned in the comments also: PubMed
Easiness or simplicity is still important (67%) despite the fact that most participants consider themselves as experienced Internet users.
Words Typed Into a Search Bar

- **Quantity**: 60% type **two to three words** in a search bar, 29% type **up to five words**.
- **Terms**: the terms typed in a search bar are **medical terms** such as a diagnosis or drug name (in 82% of the cases). Some mentioned that they type a **physician’s name**.
- **Format**: only 14% prefer to type a query in a **question format** in the same way they would have asked their **physician**.
- **Options**: the most widely used advanced options are **language** (70%), **date range** (57%) and **country** (50%).
When Search Results Appear ...

- **62%** scroll down and click on some of the results, **25%** click on the first five links.
- A few users state that they **do not click on first results, considering them sponsored**, thus biased.
- **71%** check the second or third page of results sometimes or often.
- **81%** stated that they **rarely or never access restricted information that requires paying membership fees etc.**
- Most respondents check **web sites that are clearly intended for healthcare professionals**: 21% do it always, 27% often, and 27% sometimes.
- **Kinds of preferred contents**: **web 1.0**, i.e. informative, or **web 2.0**, i.e. collaborative and interactive platforms such as forums, blogs, social networks etc. **57%** give preference to the information itself, but another **40%** have chosen a mix of both webs.
When an Answer Is Found, Do Users Double-Check It?

- 42 % verify results on trustworthy websites (based on personal perception) found by themselves.
- 21 % ask their medical doctor after performing the search.
- 13 % use another search engine.
- 10 % do not verify the answer and trust the search results.

Other options mentioned include:

- Checking books and other information channels
- Presence of HONcode seal
- Adherence to a similar ethical code
- Verification on a governmental web site

Many mentioned that they prefer to use common sense to evaluate different websites and other sources of information, including physicians, before drawing the final conclusion.
In Case of “Complex” Question

- 36% spend **usually more than 10 minutes to explore** it and to find an answer.
- 30% spend **5 to 10 minutes**.
- 20% spend **3 to 5 minutes**.

In such cases, **54% change the terms** of their search 2 to 3 times before they get the desired answer, and **36% do it more than 3 times**.
TOP 10 Difficulties When Searching for Health Information

- Distracting Ads
- Lack of quality filter
- Relevance of matches
- Questionable trustworthiness
- Quality and explicitness/fullness of description
- Overload with information quantity
- Poor organization of search results
- Search results lack specificity
- Evaluation of the link for each search result
- Lack of information in my mother tongue

Bar chart showing the frequency of these difficulties:
- Always
- Often
- Sometimes
- Rarely
- Never
Influence of Advertisement Banners

- 50% state that advertising banners do not influence their search; however 42% say they are distracted by ads.

Mentioned in the comments of some participants:

- Some block ads banners, others doubt the credibility of the web site with many ads banners.
- In some cases respondents state they do not return to a website with many advertising banners.
- «Too much ads kill ads»
Actions When an Answer Is NOT Found

**Answer is not found “sometimes” for 51 % of respondents** and “often” for 9 % of respondents. The solutions are “usually” or “always” to modify search terms (80 %), ask their medical doctor (48 %) or verify the information on a website they trust (46 %)
Reasons to Fail to Retrieve the Answer

Possible reasons for failing to retrieve a satisfying answer: search results do not guide the user towards an answer (32% of the responses are “always” or “usually”), or the user is overwhelmed by the quantity of results (30%).
Other Reasons to Fail to Retrieve the Answer

- **Contradicting and conflicting data**: different web sites provided contradicting and conflicting data, thus a final conclusion cannot be drawn.

- **Lack of time for surfing**: some respondents stated that they do not check the first results because they are sponsored, and do not have patience and time to look through all the pages to reach the answer.

- **Very precise questions do not have an answer in a plain language**: some admitted they were asking very precise questions, and after having been redirected to physicians’ websites it was difficult for them to understand the information there, so they had to go back to more basic questions to understand the terminology. In other cases a professional site required membership to access the data.

- **Very precise questions do not yield any answer at all.**

- **No results relevant to a user’s country.**

- **Social media obscure scientific results.**
Ideal Citizen-Centred Health Search Engine

- **54 %** of respondents have chosen the **categorization approach** when all the links are grouped into scientific, clinical, commercial, advertising, forums and blogs.

- **24 % would prefer a summary** referencing the different sources, and **20 % like the “conventional” form of search result** presentation as a list of links (such as Google®).

- Some proposed a **combination of two options**, i.e. all links being categorized, and for each category having a summary.

- Another option is to have a list of all links with a possibility of **manual categorization**, including also date, certification, geolocation, language and text size.

- Users also want **to know if the information they search for exists on the Internet**, and **whether it is explained in the same way their doctor would**.

- Regarding the amount of results presented, **36 % would like to see between five and ten trustworthy results**, **30 % more than ten results**, and **24 % would like to have all possible results**.
Top 10 Helpful Tools

- Advanced search (country, language, date range)
- Medical dictionary/Thesaurus
- Suggested relevant topics
- Spelling correction
- Search of images
- Risk factors assessment tools
- Body 3D anatomy visualization
- Suggested filling of query
- Automatic translation of results
- Tutorial on strategy of successful online health...
Other Suggestions Proposed by Respondents (1)

- **Ability to search in social media content, but with clear separation from scientific content:** “I would appreciate separating public forums, non-professional blogs and social networks from my search.”

- **User’s customization of search results** based on the intended audience (e.g. in case a page offers information for professionals, there should be an interpretation service for the general public) dates, language, certification marks, ranking, funding and (inter)national authorities’ recognition etc.

- **Clustering** of the search results and **highlighting** of the key terms.

- **Interpretation of the question in context.**
Other Suggestions Proposed by Respondents (2)

- Natural language search engine.
- Free access to scientific and medical journals and literature (to be able to search in a literature of the last 70 years).
- A search engine as a self-diagnosing tool.
- Regarding the content: not only disease information, but also information on prevention, food, tests, physiological and psychological changes that can occur during disease and therapy.
- Highlight the specialized web sites vs general ones.
- Saving a search history and organizing search results.
- A correct display and functioning of a search engine with all browsers.
- No advertising banners and no commercial influence on the ranking of search results.
Advertisement In an Ideal Search Engine

- 84% of the respondents tolerate ads in order to have a search engine that is free of charge, 9% are prepared to donate money, 5% prefer membership subscription and 2% would prefer a fee-for-service, meaning that the user would pay each time s/he uses a search engine.

- Nevertheless, 67% of the respondents think that it is important or very important that ads undergo a quality control, thus are compliant with a certain ethical policies.
Emerging Topic of Self-Diagnosis

- 46% across 12 countries use the Internet for self-diagnosing (Online Health: untangling the web. McDaid D, Park AL. Bupa. 4 Jan 2011).

- The consequences of a wrong self-diagnosis are difficult to estimate if actions are taken without a consultation with a qualified physician. However, the desire of patients to be more educated and able to research their own health symptoms is understandable, especially in countries where access to healthcare facilities is uneven or expensive.

- Thus, risk-assessment tools can be proposed, which implies that upon answering a set of questions about a specific disease, a patient will be given a range of potential disease developments, preventive recommendations and a notification of urgency to have a medical consultation.

- However, such tools cannot be created automatically by using information retrieval technologies; the help of qualified physicians with extensive experience is required.
Limitations

The results presented in this report reflect the views of educated people. A vast majority of participants graduated from a university, many of them have a Master or PhD. This group of people cannot represent the entire e-patient community or Internet community. Nevertheless, having such sample of highly educated individuals allowed us to collect different ideas more elaborated and conscious.
Conclusions (1/2)

1. As most respondents are connected to the Internet via Wi-Fi, the use of mobile devices will keep growing. Thus, if the problem of **web site accessibility from a mobile device is not a priority for now**, this might change.

2. The **Internet is the second source of information after physicians**, and it is more likely to be used due to its **accessibility and easiness**.

3. Most researched topics are **general health, chronic diseases and lifestyle**. Illness-related **information can be prioritized** in the following way: treatment/therapy, detailed and general disease description, drugs and scientific articles.

4. A majority of **e-patients are using a general search engine** to find the answers to their health queries.

5. The most important characteristics are **relevance and trustworthiness** of the results. This appears to be the main problem, as current **results do not satisfy these requirements**.

6. Additionally, **advertising banners**, contradicting information found on the different web sites and difficulties related to very precise queries are **among other barriers**.
Conclusions (2/2)

7. The ideal representation of the information is a categorization of results into different groups with the possibility to perform this automatically or manually.

8. Tools that are highly appreciated by the respondents are advanced search, medical dictionary/thesaurus, suggested relevant topics, image search and risk assessment tools.

9. However, taking into account the low Internet literacy of the general population we need to develop tutorials to explain how to use all these options.

10. Taking into account varying health literacy of patients, it is important to assign the levels of reading difficulty for each search result.

11. Most users will tolerate ads in order to keep a search engine free, but would like to be ensured that the ads are in adherence with ethical policies.

12. Some of the respondents proposed the possibility of creating a profile with a search history for follow up investigation of the topic.

13. Some respondents have also mentioned they would like to have free access to the medical literature.

14. Many users would like to see search results appear in social networking web sites but it is important to clearly identify the original source.
Our Approach

We cannot ...

... prevent users from finding irrelevant and misleading information on Internet.

We can and should ...

... highlight quality information proved by research and current medical practice.
Acknowledgements ...

... to all the respondents who dedicated the time necessary to this survey;

... and to the webmasters and websites who decided to promote and supported the survey, particularly:

www.galenotech.org       www.studiodontisticobalestro.com       www.espondilitis.eu
www.santeromande.ch      http://031d07e.netsolhost.com       www.stop-tabac.ch
www.healthdirect.co.uk   www.brainandspine.org.uk       www.ir-facility.org

(The survey was also promoted on the project partners’ websites, and on social networking websites)
Acknowledgements ...

... to all Khresmoi’s partners who contributed to the survey development (primarily Marlene Kritz and Manfred Gschwandtner from the Society of physicians of Vienna, Austria), validation, translation and promotion of this survey:

- www.hes-so.ch
  University of Applied Sciences of Western Switzerland

- www.tuwien.ac.at
  Vienna University of Technology

- www.cuni.cz
  Charles University in Prague

- www.sheffield.ac.uk
  The University of Sheffield

- www.atosorigin.es
  Atos Origin

- www.dcu.ie
  Dublin City University

- www.meduniwien.ac.at
  Medical University of Vienna

- www.billrothhaus.at
  Society of Physicians in Vienna

- www.uni-duisburg-essen.de
  University of Duisburg-Essen

- www.elda.org
  Evaluations and Language resources Distribution Agency
Thank you for your participation and interest!

If you have questions, please email to:

nataly.pletneva@healthonnet.org

The full report is available:

Please, remember also that:

All the information (including the design and graphics of the Web site) contained on the http://www.HealthOnNet.org/ site is copyrighted © Health On the Net (HON), and is the property of the Foundation Health On the Net. Material found on this site (information, text and image) may be used or distributed for non-commercial purposes only under the condition that this note of Copyright remains intact in each document where HON is stated as the source and a link to the HON website is provided. The use, reproduction, or distribution of the material of the surveys conducted by HON (information, text, image) is subject to the following restrictions:

- request for use of any part of the surveys must be submitted to the HON at HONinfo@healthonnet.org.
- any copies or distribution must be accompanied by the following copyright notice. If in HTML, it must also feature the hypertext links as shown in this copyright notice: