

4. Hagerman RJ. Medical follow-up and pharmacotherapy. In: Hagerman RJ, Hagerman PJ, eds. *Fragile X Syndrome: Diagnosis, Treatment and Research*. 3rd ed. Baltimore, Md: Johns Hopkins University Press; 2002:287-338.
5. Bailey DB, Skinner D, Hatton D, Roberts J. Family experiences and factors associated with the diagnosis of fragile X syndrome. *J Dev Behav Pediatr*. 2000;21:315-321.
6. Jacquemont S, Farzin F, Hall D, et al. Aging in individuals with the *FMR1* mutation. *Am J Ment Retard*. 2004;109:154-164.

## RESEARCH LETTER

### Health-Related Searches on the Internet

**To the Editor:** There is ongoing uncertainty about how often consumers use the Internet to search for health-related information. For instance, although a 2000 survey found that 55% of individuals with Internet access have used the Web to find health or medical information,<sup>1</sup> Phillipov and Phillips<sup>2</sup> reported that 17 health-related search terms they found in the first 300 search terms of the Wordtracker Top 500 keyword report comprised fewer than 1% of all queries. It is possible, however, that people may use many different health-related keywords that are so rare that they never appear on the Top 500 Keywords list. To address this, we analyzed a random sample of search terms entered into a search engine over a longer period of time.

**Methods.** We used the Metaspyp Exposed Web site (<http://www.metaspyp.com/>), which lists 10 unfiltered search strings people are currently entering on Metacrawler, and which refreshes every 15 seconds with new terms. We developed a computer program that visited this page every 30 minutes between February 2001 and April 2002, extracted the search terms, and added them to a database. From a sample of 298512 query strings in the database, a computer program randomly selected 2985 (1%) search queries that were independently classified by 2 physicians as either "non-health-related" or "health-related." Discordant ratings were resolved by discussion.

**Results.** Of the 2985 search expressions, 2827 (94.7%) were concordantly classified as non-health-related, 108 (3.6%) were coded by both raters as health-related, and 50 (1.7%) received a discordant classification ( $\kappa=0.80$  for interobserver variability, indicating substantial agreement). Thus, our estimate of health-related search terms as a proportion of all search queries conducted is in the range of 3.6% to 5.3%. Examples of search expressions that were coded discordantly included "stop thumb sucking" and "calcium crystals." After discussion, 135 (4.5%) of all search terms were classified as health-related.<sup>3</sup> We also coded health-related search queries by broad categories (partly corresponding to clinical specialties) and found that the searches for which people entered the name of a medical organization or looked for specific health care services, as well as searches for medicines and diet, were the most frequent queries (TABLE).

**Comment.** Our data suggest that health information is not among the most commonly sought topics on the Internet. Nonetheless, given the millions of searches performed each day, neither could they be described as uncommon. Studies suggest that more than 95% of consumers use a search engine rather than

**Table.** Non-Health-Related Queries and Health-Related Queries, by Category

	No. of Queries (%)
Total No. of queries	2985
Non-health-related queries	2850 (95.5)
Health-related queries	135 (4.5)
Health-related subcategories (n = 135 queries)*	
Miscellaneous†	28 (20.7)
Health care services and organizations	13 (9.6)
Medicines	11 (8.1)
Diet, nutrition, weight loss	9 (6.7)
Infectious diseases	8 (5.9)
Cancer	7 (5.2)
Rheumatology	7 (5.2)
Women's health (except pregnancy)	7 (5.2)
Addiction	6 (4.4)
Mental health (except addiction)	6 (4.4)
Cardiovascular	5 (3.7)
Pregnancy, fertility	5 (3.7)
Dermatology	5 (3.7)
Natural/alternative medicine	5 (3.7)
Organs, anatomy	4 (3.0)
Injury	3 (2.2)
Neurology	3 (2.2)
Laboratory tests and results	3 (2.2)
Gastrointestinal	3 (2.2)

\*The sum of percentages for health-related subcategories is greater than 100% because 16 search terms were coded under 2 categories (eg, "cervical cancer" coded as "women's health" and "cancer").

†Health-related subcategories with less than 2% of health-related queries were combined in the "miscellaneous" category.

going directly to a Web site when confronted with specific health-related questions.<sup>4,5</sup> While our data were harvested from a single search engine, we doubt that the pattern of searches in other search engines would be different.

Gunther Eysenbach, MD, MPH  
 geysenba@uhnres.utoronto.ca  
 Centre for Global eHealth Innovation  
 University Health Network  
 Toronto, Ontario  
 Christian Köhler, MD  
 Department of Clinical Social Medicine  
 University of Heidelberg  
 Heidelberg, Germany

**Acknowledgment:** Jim Lai programmed the scripts and David Mason provided technical input.

1. Pew Internet and American Life Project. The Online Health Care Revolution: How the Web Helps Americans Take Better Care of Themselves. November 26, 2000. Available at: <http://www.pewinternet.org>.
2. Phillipov G, Phillips PJ. Frequency of health-related search terms on the Internet. *JAMA*. 2003;290:2258-2259.
3. Eysenbach G, Köhler C. What is the prevalence of health-related searches on the World Wide Web? qualitative and quantitative analysis of search engine queries on the Internet. *Proc AMIA Symp*. 2003:225-229.
4. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the World-Wide-Web? qualitative study using focus groups, usability tests and in-depth interviews. *BMJ*. 2002;324:573-577.
5. Hansen D, Derry H, Resnick P, Richardson C. Adolescents searching for health information on the Internet: an observational study. *J Med Internet Res*. 2003;5:e25.