

Online Consumer Search Strategies for Smoking-Cessation Information

Nathan K. Cobb, MD

Background: For many Americans, the Internet has become a primary mechanism for locating information on healthcare and treatment options, including tobacco addiction. Detailed information on this behavior could inform design decisions for next-generation cessation interventions, but very little is known about how consumers search or what resources they locate.

Methods: A subset of a publicly available, anonymized record of the search behavior of 650,000 individuals over 3 months in 2006 was analyzed. Smoking cessation-related queries were extracted and coded via manual identification of terms and by back-identifying terms by matching them to the websites ultimately visited. Destination sites were coded as to whether or not they originated from a professional source based on the literature and known healthcare organizations.

Results: A total of 628 individuals (0.10%) made 1106 cessation-related searches during the observation period. Of these, 76% resulted in the individual reaching a website; professional sites were reached by only 34% of searchers. Complementary or alternative therapies were popular, with 10% of individuals searching for “laser” therapy.

Conclusions: A concerning disconnect exists between consumer demand (as demonstrated by search behavior) and the sites produced by researchers and health professionals. This “demand gap” may contribute to low overall participation rates and hamper the potential impact of such systems. Further research is needed to link online consumer preferences to intervention design decisions. (Am J Prev Med 2010;38(3S):S429–S432) © 2010 American Journal of Preventive Medicine

Introduction

Searching on the Internet has become a major mechanism of information-gathering for consumers, with an estimated 9% of all American adults having looked for smoking-cessation information online.¹ In theory this behavior should yield large numbers of participants for telephone and web-based treatment systems, yet a discord exists between these estimates and the much smaller number of users reported by even the largest evidence-based websites.^{2,3}

Past work on demand has concentrated on identifying existing sites by assuming that consumers use a series of common terms (e.g., *smoking cessation* or *quit smoking*) to locate information.⁴ As consumer literacy about effective cessation modalities is low and over 70% of smokers use no assistance,⁵ this assumption is questionable. Different choices in search terms may lead consumers to any

number of systems, ranging from effective to potentially harmful. This concern is bolstered by the fact that the most commonly recalled intervention among consumers was produced by a tobacco company.³ In order to inform the construction of the next generation of behavioral treatments and marketing strategies using a “design for demand” model⁵ (where consumer preferences and behavior are taken into account in designing behavioral interventions), there is a pressing need to identify the true mediators of consumer online-information seeking as it leads from demand to adoption of interventions.

Raw data from search engines would represent real-world, empirical evidence of these mediators: how individuals search in uncontrolled settings, what terms they use, what results they receive, and what links they ultimately click on. Unfortunately, little is published about consumer search patterns or outcomes as the data are proprietary and normally closely guarded by the large search companies.

In 2006, the Internet service provider America Online (AOL) released a large sample of search data in a good-faith effort to work with the research community. The sample was released on a AOL-controlled website for external researchers (from which we obtained our copy). While the data had been carefully anonymized, the

From the Schroeder Institute for Tobacco Research and Policy Studies, American Legacy Foundation, Washington DC

Address correspondence and reprint requests to: Nathan K. Cobb, MD, Schroeder Institute for Tobacco Research and Policy Studies, American Legacy Foundation, 1724 Massachusetts Avenue NW, Washington DC 20036. E-mail: ncobb@americanlegacy.org.

0749-3797/00/\$17.00

doi: 10.1016/j.amepre.2009.12.001

Table 1. Search queries by category

Category	Individuals (n=628)					
	Total making any search in category ^a		Found any website ^b		Found professional website ^b	
	n	%	n	%	n	%
Any	628	100.0	479	76.3	215	34.2
Medication	185	29.5	143	77.3	73	39.5
Quit/ing, stop/ing smoking^c	173	27.5	101	58.4	66	38.2
Complementary or alternative	87	13.9	71	81.6	3	3.4
Laser	62	9.9	50	80.6	1	1.6
Acupuncture	6	1.0	6	100.0	1	16.7
Hypnosis	12	1.9	9	75.0	0	0.0
Other	12	1.9	10	83.3	1	8.3
State- or city-specific	74	11.8	56	75.7	0	0.0
Internet resources	41	6.5	30	73.2	17	41.5
Quitline/telephone	6	1.0	3	50.0	0	0.0
Nonspecific/uncategorized	281	44.7	184	65.5	110	39.1

^aCategories may overlap, so totals exceed 100%.

^bPercentage of “Total making any search in category”

^cContains only exact matches to the “common” search terms

breadth of the search queries caused concern as the data spread rapidly across the Internet, leading AOL to take the files down.⁶ As a result, while the data set is widespread and can be easily downloaded or searched online from other sources, there have been few publications that reference it.⁷ In an effort to explore consumer search strategies without running the risk of inadvertently identifying a user, a tobacco-related subset was extracted from the AOL data and analysis restricted to those data.

Methods

The data consist of searches via the Google search engine, performed within the AOL website, by AOL members over the course of a 3-month period in 2006 (Table 1). It contains a randomly selected set of 657,426 users and their associated 21,011,340 searches. Users are identified by an anonymized numeric key only; there are no demographic or descriptive data and no way to tie them back to account data at AOL. Search data contain the unique user key, the query, the date and time the search was performed, and the URLs of any sites to which the user clicked through (i.e., clicked on a link to the site). An actual sample search session for a single user is presented in Figure 1. Analysis was performed in 2008 using a local copy imported into a relational database.

Two basic mechanisms were used to identify and then extract search terms that might be related to smoking

cessation. Searches containing the “common” terms *quit(ing) smoking* and *stop(ping) smoking* were extracted first, then supplemented with treatment terms specific to smoking cessation, such as *Zyban* (but not *Wellbutrin* or *bupropion*, which are not treatment-specific), *Chantix*, or *nicotine patch*. Telephone treatment-related terms were identified using the words *quitline* or *hotline*. To identify unrecognized terms, including potential misspellings, an arbitrary list of sites that were devoted to smoking cessation was compiled, independent of quality. The click-through results were then parsed to identify cessation-specific terms that led to one of these sites.

After extraction, the search terms were reviewed and coded; those that could not be associated exclusively with tobacco cessation were excluded (e.g., *nicotine* or *quit smoking weed*). Resulting terms were coded into broad, nonexclusive categories, and a list of all destination sites was extracted. As a proxy for evidence-based content, sites were coded with the assistance of previous review literature^{3,4,8} as “professional” based solely on their production by a recognized academic or healthcare entity.

Results

In all, 589 distinct search terms were identified. Five were medication names specific to smoking cessation and six specific to telephone treatment; 492 terms contained the common phrases (e.g., *quit smoking laser Sarasota*), and

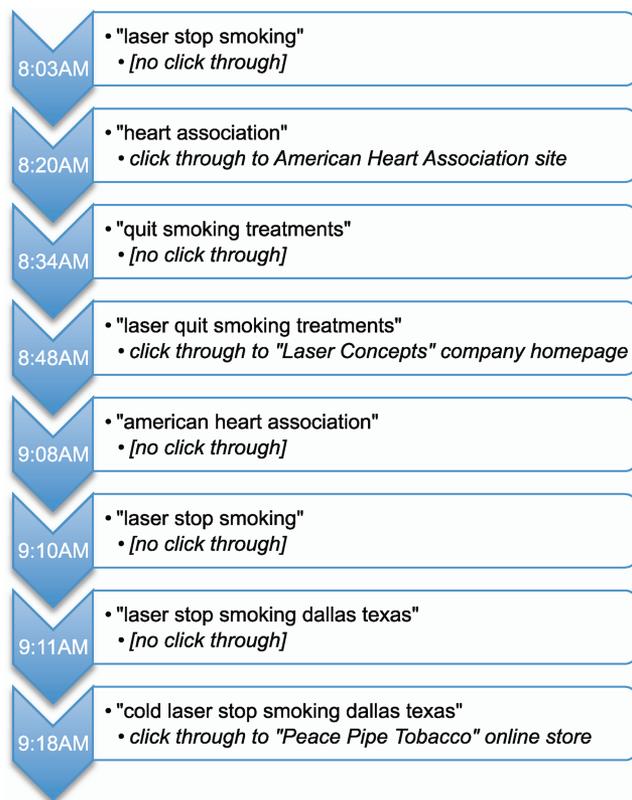


Figure 1. Sample search session

86 additional terms were identified via back-extraction from click-through data.

During the 3-month period, 628 users (0.10% of the total population) made 1106 queries using these terms. Nonspecific searches were common (e.g., *stop smoking tips*), with

nearly half (44.7%) of individuals using such terms at some point. A total of 173 users (27.5%) used the exact stem of the "common" search terms; 29.5% of users searched for identified medications; 13.9% for alternative therapies (meaning not evidence-based [e.g., herbal supplements or acupuncture]); and only 1.0% for quitline terms (six individuals).

Of the 1106 searches in the data set, 741 (67%) were successful in the sense that they led to one or more click-throughs, in this case to one of 619 websites. In only 24% of the searches did the user click through to one of the 28 identified professional sites. These sites were ultimately accessed by 34.2% of users (many of whom made multiple searches). *Alternative therapy* searches infrequently (3%) led to professional sites. Interestingly, the most common alternative therapy modality was cold laser acupuncture (where low-intensity laser light is thought to stimulate acupuncture points), with 62 users (10%) using keywords containing the word *laser*.

Top destinations are shown in Table 2. Most common was the CDC (www.cdc.org), yet the second was a site forcefully opposed to nicotine replacement therapy (whyquit.org). Six of the ten were commercial sites; two promoted laser therapy, and an additional two promoted primarily herbal therapies. Use of tobacco company sites appeared rare: four individuals clicked through to the Philip Morris QuitAssist site (www2.philipmorrisusa.com/en/quitassist/), three of whom searched on variants of the common terms, and one who searched specifically for the site.

Discussion

The discord between the large number of individuals estimated to be looking online for information on smok-

Table 2. Top ten search destinations

Website URL (average results list position)	Searches (%; n=1106)	Description and source
1. www.cdc.gov (2) ^a	7.7	Governmental site; CDC
2. whyquit.org (5)	6.8	Personal site; promotes nonpharmacologic cessation
3. www.quitnet.com (2) ^a	5.1	Commercial site; owned by large disease management company
4. quitsmoking.about.com (4)	4.9	Commercial site; single lay author, part of larger network of small health sites
5. www.quitsmoking.com (4)	4.3	Single lay author; surrounding advertisements for commercial alternative therapies
6. www.cancer.org (4) ^a	3.3	Nonprofit site; American Cancer Society
7. www.freedomlasertherapy.com (2)	2.5	Company home page; provides laser therapy
8. www.ash.org.uk (4) ^a	2.1	Nonprofit site; Action on Smoking or Health
9. www.finalsmoke.com (3)	2.0	Company home page; sells herbal therapies
10. www.laserconcept.com (2)	2.0	Company home page; provides laser therapy

^aCoded as "professional" site

ing cessation¹ and the relatively low utilization rates for online evidence-based treatments imply that consumers are not finding such interventions. Results from this brief analysis suggest that the disconnect may be due in part to the fact that online consumers use a broad array of search terms, many of which lead to non-evidence-based therapies. Given that online searching is an important mechanism of health-information seeking for American consumers,¹ these findings may have broad implications.

The scope of this data set is unprecedented; a rough estimate based on a national smoking prevalence rate of 19.8%⁹ suggests it includes over 130,000 smokers. It is still limited by its origins from a single provider, one with its roots in inexpensive, dial-up access. The data lack demographics or tobacco-use status, thus it is impossible to be sure that it is truly representative of the larger population of Internet users. AOL is presumed to serve a more rural, lower-SES population, a population that tends to have higher smoking rates, but receives evidence-based treatment at a lower rate.⁵ The fact that searching for laser acupuncture therapy was an order of magnitude more common than for telephone quitlines (despite substantial media coverage¹⁰) may indicate a failure to reach or engage this high-priority population.

There are many implications of these findings for developers, policymakers, and practitioners to better leverage existing consumer demand. Simplest is that buyers of keyword-based advertising within search engines should expand their lists to include terms related to location, alternative therapies, and pharmacotherapy. More importantly, the use of “design for demand”⁵ techniques is one potential means to bridge the gap in expectations between consumers and professionals. For example, the integration of information on alternative therapy treatment modalities (even with negative conclusions) into online interventions and marketing campaigns will cause these sites to appear higher in results lists for a broader array of searches. As consumers preferentially click on higher-ranked results, this strategy has the potential to increase utilization and improve consumer satisfaction and adherence.

The fact that only 34% of searchers reach “professional” sites suggests that relatively low levels of use of evidence-based interventions may not be entirely due to insufficient awareness of evidence-based therapies or lack of desire for cessation support, but rather a disconnect between consumer interests and the content of evidence-based therapies. This work with raw search data merely scratches the surface; more exploratory and formative research is urgently needed to address this “demand gap” between online consumer behavior and the dissemina-

tion of evidence-based treatments. Such research correlating demographics, knowledge, attitudes, and beliefs with real-world observed behavior is possible with industry collaboration¹¹ and will lead to interventions that can leverage existing online demand patterns to efficiently deliver evidence-based tobacco-cessation treatments and improve population health.

This research was supported in part by NHLBI under a training grant to the Massachusetts General Hospital (T32HL00787). The author appreciates the support of Dr. Nancy Rigotti in conceptualization of the project and comments from Drs. Amanda Graham and David Abrams on early drafts.

NKC was a stakeholder in QuitNet Inc. prior to its acquisition in 2005, and is a consultant to MeYouHealth, a behavioral health technology company.

References

1. Fox S. Online health search 2006. Washington: Pew Internet and American Life Project, 2006.
2. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. *Nicotine Tob Res* 2005;7:207–16.
3. Etter JF. A list of the most popular smoking cessation websites and a comparison of their quality. *Nicotine Tob Res* 2006; 8(1S):S27–34.
4. Bock BC, Graham AL, Whiteley JA, Stoddard JL. A review of web-assisted tobacco interventions (WATIs). *J Med Internet Res* 2008;10:e39.
5. Orleans CT. Increasing the demand for and use of effective smoking-cessation treatments reaping the full health benefits of tobacco-control science and policy gains in our lifetime. *Am J Prev Med* 2007;33:S340–8.
6. Barbaro M, Zeller T. A face is exposed for AOL Searcher No. 4417749. *The New York Times* 2006, Aug 9. Available at <http://www.nytimes.com/2006/08/09/technology/09aol.html>. Accessed August 9, 2006.
7. Hafner K. Researchers yearn to use AOL logs, but they hesitate. *The New York Times* 2006, Aug 23. Available at <http://www.nytimes.com/2006/08/23/technology/23search.html>. Accessed August 23, 2006.
8. Pike KJ, Rabius V, McAlister A, Geiger A. American Cancer Society's QuitLink: randomized trial of Internet assistance. *Nicotine Tob Res* 2007;9:415–20.
9. CDC. Cigarette smoking among adults—U.S., 2007. *MMWR Morb Mortal Wkly Rep* 2008;57:1221–6.
10. Hurd AL, Augustson EM, Backinger CL, Deaton C, Bright MA. Impact of national ABC promotion on 1-800-QUIT-NOW. *Am J Health Promot* 2007;21:481–3.
11. Tancer B. Click: what millions of people are doing online and why it matters. New York: Hyperion, 2008.