A Multidimensional Approach to Web Evaluation
at the U.S. National Library of Medicine

by Fred B. Wood, DBA, MBA, and Becky J. Lyon, MLS

National Library of Medicine (NLM)/National Institutes of Health (NIH)

8th European Conference of Medical and Health Libraries, Cologne, Germany

September 19, 2002

UNITED STATES National Library of Medicine

Outline

A. Why web evaluation?

B. Usability feedback

C. User feedback

D. Usage data

E. Web/Internet performance data

F. Conclusions

--Note #1: This is an overview presentation. For an in-depth example of web evaluation, see Eve-Marie LaCroix's presentation Fri, 10:30am, 9/20/02, on MEDLINEplus.

--Note #2: For an overview of NLM, see Becky J. Lyon's presentation Fri, 9:30am, 9/20/02.
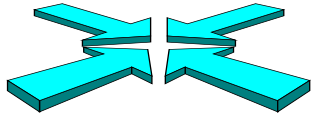
A. Why web evaluation?

- Transition to 99% web-based health information dissemination over last 5 years

- Expanded definition of "users"—general public, community-based organizations in addition to health providers, librarians, researchers

- Traditional "evaluation" methods less applicable, feasible in web environment

- More challenging desired "outcomes" or "performance measures"

H. Ross Ashby's Law of Requisite Variety

- Complexity of Monitoring/Evaluation and Management System must $\geq$

  Complexity of the Real World System

- Applies in spades to web-based health information systems

- Without new and updated methods, we run the risk of flying blind
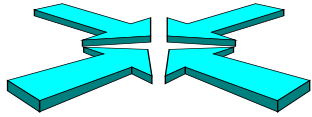
Why we care—desired outcomes

1. Increase awareness, accessibility, and usage of health information

2. Increase impacts of usage of health information

3. + Impact on patient/provider health decisions

4. + Impact on patient health-related behaviors

5. + Impact on patient and overall indicators of health

6. Reduce health disparities (among ethnic, cultural, geographic groups)

More desired outcomes

7. Reduce rolling digital divide (computer/Internet haves and have nots)

8. Improve quality of web-based health information

9. Encourage innovation in web-based medical informatics, telemedicine, Health Next Generation Internet

10. Promote successful application of innovations in health arena

NLM's response—Implement a multidimensional, diverse set of web evaluation methods & studies

B. Usability Feedback

- Heuristic/expert review of select web sites

  -- applies standard web design principles by recognized experts

  -- can include comparative search results

- Informal usability lab testing

  -- semi-structured testing in a computer lab

- Formal usability lab testing

  -- structured testing/feedback in formal lab setting

  -- use of videotape and transcripts
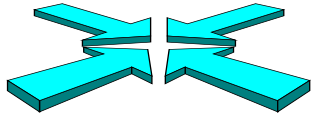
C. User Feedback

- Directed user feedback

  -- a group of users are asked to provide feedback, typically online, re a developmental or new web site in response to test exercises and questions

- Face-to-face focus groups

  -- a small group of users (typically 8-12) meet for an hour to provide feedback on a web site

- Online focus groups

  -- small group of users respond in an entirely online, interactive environment, with online moderator

User Feedback—Contd

- <u>Nationwide telephone surveys of online health information users</u>

  -- randomized sample of online users re health information-seeking practices

  -- e.g., syndicated survey CybercitizenHealth conducted annually in US (by CyberDialogue, Inc., now Manhattan Research)

  -- e.g., various surveys sponsored by the Pew Internet Project—see www.pewinternet.org

  -- provides useful contextual and comparative info

User Feedback—Contd

- Online "virtual" user survey

  -- survey of members of an external panel of web users, not NLM's own users

- Online "virtual" user competitive analysis

  -- survey of external panel members re comparative feedback on several web sites

- Online "real" user surveys

  -- randomized sample of actual users given opportunity to respond to online questionnaire

  -- NLM has completed surveys of 4 web sites

Cross Comparison of Survey Results

- MEDLINEplus (www.medlineplus.gov)— February 2001—2,969 respondents

- ToxNet (www.toxnet.nlm.nih.gov)—December 2001-February 2002—1077 respondents

- PubMed (www.pubmed.gov)—March-April 2002—15,826 respondents [3,839 subset for quality control]

- NLM Home [or Main] Page (www.nlm.nih.gov) —July 2002—4,163 respondents

Snapshot of Cross Site Comparisons

	NLM Main Page (n=4163)	PubMed (n=15,826)	MEDLINEplus (n=2969)	TOXNET (n=1077)
Study Dates	July 8th – July 17th, 2002	March 21st - April 5th, 2002	February 14th – 27th, 2001	December 6th, 2001 – February 11th, 2002
Country				
U.S.	**66%**	**45%**	**72%**	**60%**
Non U.S.	**34**	**55**	**28**	**40**
Site Usage				
Repeat visitors	**78**	**98**	**49**	**77**
Visit (nearly) daily	**32**	**56**	**14**	**15**
Visit at least weekly	**32**	**29**	**34**	**30**
Satisfaction Measures				
Extremely/Very/Satis-fied (Repeat Visitors)	**94***	**97**	**98**	**99**
Extremely/Very/Likely to Return	**94****	**98**	**100**	**97****

*Main Page survey did not offer "Extremely Satisfied" option

**TOXNET offers "Somewhat Likely", not included in this number

Cross Site Comparisons—Contd

	Main Page (n=4163)	PubMed (n=3839)	MEDLINEplus (n=2969)	TOXNET (n=1077)
Researcher/Scientist	24%	51%	7%	20%
Toxicological Professionals*	NA	NA	NA	43
Health Care Provider/Physician/Other Health Professional	19	20	13	5
Librarian/Information Professional	20	5	9	10
General Public (Net)**	31	21	66	15
College/Graduate Student	8	15	9	5
Patient/Patient w/Specific Condition	7	1	25	2
Family or Friend of Patient	5	1	17	1
General Health Consumer	5	-	12	NA

*Includes Toxicologist/Pharmacologist, Industrial hygienist, Environmental engineer, Regulator (federal, state, local), Public health official, Emergency responder, Poison control

**Includes Media, Legal, Other, and Misc. as well as categories noted here.
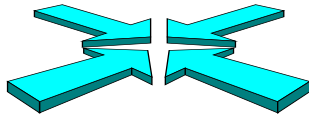
D. Usage Data

- Web log data analysis

-- collection and analysis of web log file data using commercial or custom software

-- provides basis for time series and drill down analyses of usage traffic, but must recognize error factors since IP addresses (computers), not identifiable users, are being measured

-- undercounts institutional users (e.g., libraries) with fixed IP addresses, overcounts individual users with dynamic IP addresses

Usage Data—Contd

- **Web log metrics**

  -- most common metrics include pages downloaded, total visits, unique visitors, and, where applicable, searches per month

  -- other metrics include repeat visitors, pages per visit, US vs non-US visitors (subject to limitations), originating and referring sites, most frequent pages visited

- **Survey-log data comparisons**

  -- can cross-validate accuracy of data sets

Usage Data—Contd

- <u>External Internet audience measurement</u>

  -- private companies assemble large virtual panels of users who agree to web usage monitoring

  -- panel size ranges from 50K to 1.5 million, and cover home, office, school, some international

  -- usage data extrapolated to US or global estimates based on demographics & census data

  -- exact methodologies vary by company, most use common metrics of pages viewed, unique visitors per month

Usage Data—Contd

- <u>NLM history with Internet audience companies</u>

  -- PCDataOnline, Inc., 1999-2001 [defunct]

  -- comScore Networks, Inc., 2001-present

  -- Nielsen/NetRatings, Inc., 2002-present

- <u>NIH and NLM drill-down data possible</u>

  -- drill-down studies indicate that NLM web sites account for ~50-55% of total NIH.gov traffic (as measured by unique visitors per month) and up to ~80% (based on page views)

Health Information Sector
## August 2002 (netScore data)

[excluding web portals with health channels]

	US unique visitors	Global unique visitors	US pages viewed	Global pages viewed
NIH.gov	3.2M	5.9M	42.7M	84.5M
WebMD	3.7M	4.6M	44.8M	53.0M
Medscape	0.54M	0.97M	6.7M	15.5M
MayoClinic	0.45M	0.64M	3.8M	5.2M
Intelihealth	0.35M	0.52M	2.5M	3.9M
DrKoop	0.28M	0.4M	1.5M	2.0M

US Government Health Sector
August 2002 (netScore data)

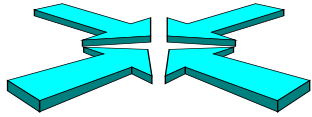|	US unique visitors	Global unique visitors	US pages viewed	Global pages viewed
NIH	3.2M	5.9M	42.7M	84.5M
CDC	1.8M	2.7M	9.0M	13.6M
FDA	0.76M	1.2M	7.5M	10.1M
HHS	0.97M	1.1M	8.5M	8.9M
VA	0.78M	0.85M	9.2M	10.2M
CMMS	0.24M	0.26M	1.5M	1.6M
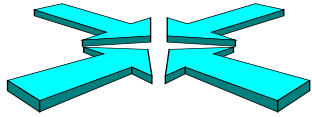
E. Web/Internet Performance Data

- Key question--How fast can users download web pages and conduct database searches via the Internet?

- Metrics—transport level/TCP (BTC, RTT, route stability, packet loss, route utilization)

- Metrics—applications level/HTTP (response time, download time)

- Timing—ad hoc, defined period, continuous

- Directional—asymmetric/symmetric, inbound/outbound

Performance Data—Contd

- <u>NLM's Internet connectivity performance testing</u>

   -- extended to higher bandwidth pathways (vBNS, Abilene) starting in 2000

   -- transition to operational status as part of NOSC (Network Operations & Security Center)

   -- planned NLM-centric monitoring network, e.g., with Regional Medical Libraries, Resource Libraries, select International MEDLARS Centers

   -- pilot project with University Corporation for Advanced Internet Development (UCAID) /Internet2 on critical incident high bandwidth uses

Performance Data—Contd

- External Internet performance testing

    -- Keynote measurement services, since 1998

    -- Average download times for NLM web sites as measured from Keynote's network of US and international nodes

    -- Comparisons with business, consumer, and government performance benchmarks

    -- Used for network management & comparison with internal testing

    -- Compare Keynote with internal and NLM-centric network measurements
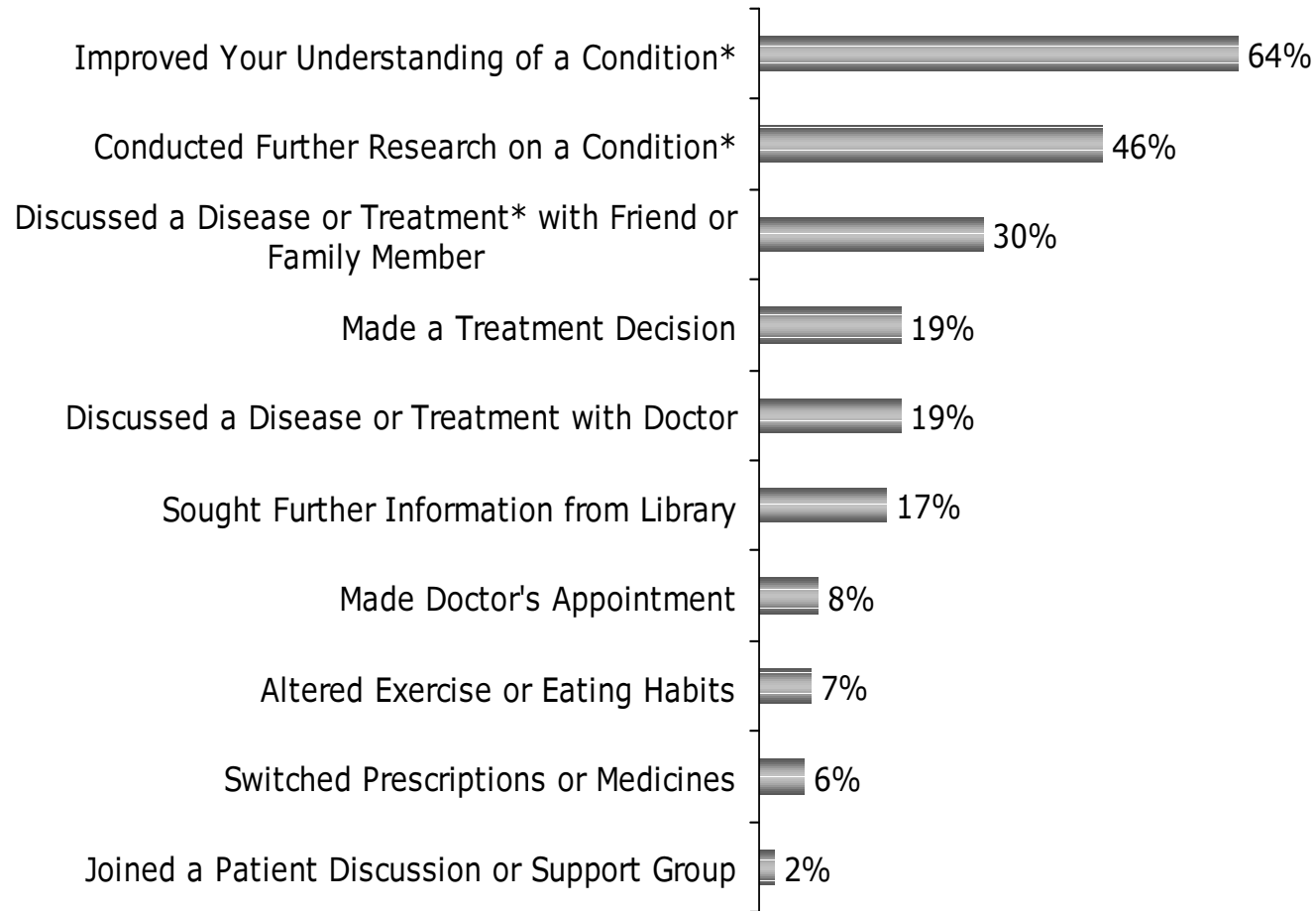
F. Some Conclusions

- NLM's experience since 1997 indicates that a diversified web evaluation strategy is warranted.

- Online user surveys provide useful information.

- For NLM, the four web sites surveyed have a solid base of highly satisfied core users.

- For these web sites, the predominant user groups appear to be well matched to the primary site mission(s).

- Taken together, these web sites appear to be meeting the needs of a broad spectrum of users.

Conclusions—Contd

- First time users may be less likely to fill out surveys—focus groups and special feedback projects can help here.

- Users learn about NLM web sites through diverse means and channels.

- A highly diversified communications strategy is warranted—e.g., search engines, links with other web sites, libraries, physicians, health groups.

- Are we having an impact? The user surveys suggest yes—witness MEDLINEplus.

Impact of Using MEDLINEplus Information

[Repeat Visitors Only, n = 1455; * = Disease, condition, diagnosis, or treatment]

Category	Percentage
Improved Your Understanding of a Condition*	64%
Conducted Further Research on a Condition*	46%
Discussed a Disease or Treatment* with Friend or Family Member	30%
Made a Treatment Decision	19%
Discussed a Disease or Treatment with Doctor	19%
Sought Further Information from Library	17%
Made Doctor's Appointment	8%
Altered Exercise or Eating Habits	7%
Switched Prescriptions or Medicines	6%
Joined a Patient Discussion or Support Group	2%

Conclusions—Contd

- In sum, a multidimensional, robust web evaluation strategy has helped NLM better understand its users and the health information sector, develop and improve its web sites, and gauge the impact of web site usage.

- This web evaluation strategy has helped the NIH family of web sites, and within NIH most predominantly powered by NLM, to emerge as a leader in both the USG and global health sector.

- This multidimensional strategy seems generally applicable to other health information web sites, including those represented by the EAHIL.
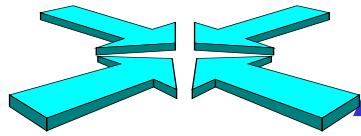
Additional Information:

- For a copy of this presentation or follow-up questions, e-mail Fred Wood at fred_wood@nlm.nih.gov

- NLM co-sponsored an April 17, 2001 Symposium on "Evaluating Our Web Presence" with CENDI, a U.S. interagency science information group

- Presentations available on the CENDI web site at http://www.dtic.mil/cendi/

- Also see www.Usability.gov, maintained by the U.S. National Cancer Institute, for information on a wide range of web usability and user feedback topics.

UNITED STATES
National Library of Medicine