

Infodemiology – the epidemiology of (mis)information



Gunther Eysenbach MD, MPH

Associate Professor

Department of Health Policy,
Management and Evaluation, University
of Toronto;

Senior Scientist,
Centre for Global eHealth Innovation,
Division of Medical Decision Making and
Health Care Research;
Toronto General Research Institute of
the UHN, Toronto General Hospital,
Canada

Head

Research Unit for Cybermedicine & eHealth
Dept. of Clinical Social Medicine
University Hospital Heidelberg, Germany



R. Fraser
Kilmer
Building

Mission:

To study how to help people access the information and services they need, when and where they need them, regardless of who they are.

Centre for Global eHealth Innovation
Toronto General Hospital / University of Toronto

Usability labs at the Centre for Global eHealth Innovation







Epidemiology vs. Infodemiology (1)

- Epidemiology: the study of the different factors occurring along with diseases and their distribution in the population
 - Descriptive E.
 - Analytical E.
 - Interventional E.
- Infodemiology (information epidemiology): studying the determinants and distribution of health information and misinformation globally
 - Descriptive E.
 - Analytical E.
 - Interventional E.

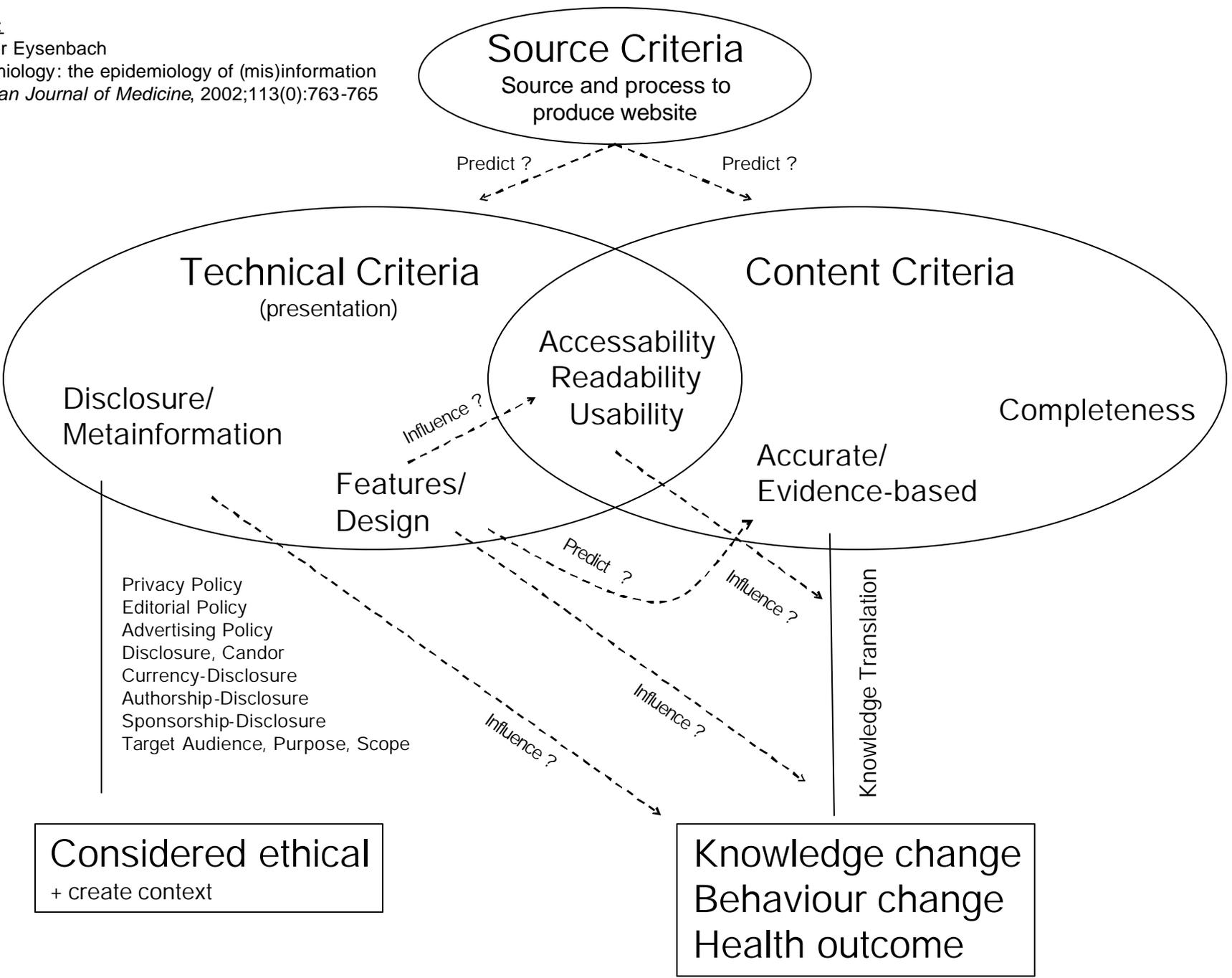
Descriptive Infodemiology

- Descriptive epidemiology: studying frequency (prevalence) and distribution of diseases and risk factors in populations,
- enables to identify possible disease outbreaks
- Descriptive Infodemiology: studying proportions (prevalence) of “inaccurate” or imperfect health information, and quality markers
- Raising a flag, pointing to areas where fraud / misinformation is prevalent, identifying gaps in information provision / knowledge translation

Analytical Infodemiology

- Analytical epidemiology: studying associations between risk factors and health outcomes using multivariate regression
- Analytical Infodemiology: studying associations between information characteristics or “quality” markers and outcomes or other variables (using multivariate regression)

Source:
 Gunther Eysenbach
 Infodemiology: the epidemiology of (mis)information
 American Journal of Medicine, 2002;113(0):763-765



Is a health website CREDIBLE?

- **C**urrent and frequently updated
- **R**eferences cited
- **E**xplicit purpose and intentions of the site
- **D**isclosure of developers and sponsors
- **I**nterests disclosed and not influencing objectivity (e.g. financial interests)
- **B**alanced content, list advantages and disadvantages
- **L**abeled with metadata
- **E**vidence-level indicated

Infodemiology: The Epidemiology of (Mis)information

Gunther Eysenbach, MD

Much of the health information on the Internet has been described as being discordant with information from evidence-based sources (1). A new research discipline and methodology has emerged—the study of the determinants and distribution of health information and misinformation—which may be useful in guiding health professionals and patients to quality health information on the Internet. Information epidemiology, or infodemiology, identifies areas where there is a knowledge translation gap between best evidence (what some experts know) and practice (what most people do or believe), as well as markers for “high-quality” information.

The first infodemiological study was published in 1996 (2), but this type of research only became widely known with a subsequent publication in a prominent journal (3). A recent review identified 79 infodemiological studies (1), and as of today more than 100 articles have been published. Most of the early studies were descriptive, reporting the percentage of websites that had inaccurate or otherwise imperfect health information (1). Such studies are also useful in identifying where the evidence is conflicting, where fraud is prevalent, or where misleading advertisements prevail over balanced health education. For example, studies have concluded that up to 90% of information on diet and nutrition is unreliable, compared with only 5% for cancer (1).

Descriptive studies, however, do not explain how indicators for quality and website characteristics are related, and whether these characteristics are associated with utility to the consumer. Analytical studies, which employ statistical methods such as multivariate regression to explore how quality criteria and other variables are related (Figure), are more useful in addressing questions such as “Are government sites more accurate than commercial sites?” or “Are websites that disclose the authorship and include a date of last update more accurate?” One impor-

tant question is which markers or characteristics of a website are “valid” quality criteria to discriminate or predict a “good” health website. A valid quality criterion would be a feature (or a combination of features) that predicts effective health communication in terms of improving knowledge or changing health behavior, or which is associated with a measurable effect on health outcomes. However, analyses of site characteristics and health outcomes can be complicated, since measuring the very small and indirect effect that a single site may have on health outcomes is not feasible. Furthermore, users rarely use only one website and often use other sources of information, and health outcomes may take years to develop.

Perhaps a more realistic aim of analytical studies is to answer the question of whether technical or formal site characteristics that suggest accuracy of content can be identified, as was done by Martin-Facklam et al. in this issue of the *Journal* (4). The authors found citing references and an absence of financial interest to be associated with content accuracy. Indeed, websites targeted at medical professionals are more likely to provide references (5) and may be more consistent with evidence-based guidelines. Disclosure of the authors’ names was not found to be a predictor of site reliability, perhaps because government organizations (6), other organizations, and drug companies (7) do not usually provide the names of authors on their websites. Disclosure of when the website was last updated may also be an independent predictor of site accuracy, although this association was not observed in an earlier analysis (8). One study suggested that displaying the HONcode logo, having an organization (.org) domain, and displaying a copyright are predictors of content “reliability” (9). These authors, however, evaluated comprehensiveness rather than accuracy (1). For all these results, the question of whether these findings can be generalized arises.

Technical (or formal) markers are called “valid” quality criteria if they predict accurate content, as accurate content is assumed to lead to positive health outcomes. Conversely, failure to establish an association between a quality criterion and accuracy of site content does not mean that the marker is an “invalid” quality criterion, because some criteria such as disclosure of sponsorship or appropriate privacy policies are ethical tenets and quality criteria per se. They have face validity, independent of whether they predict an accurate website or a positive

Am J Med. 2002;113:763–765.

From the Centre for Global eHealth Innovation, University Health Network, and Department of Health Policy, Management and Evaluation, University of Toronto, Ontario, Canada.

Requests for reprints should be addressed to Gunther Eysenbach, MD, Centre for Global eHealth Innovation, Toronto General Hospital, R. Fraser Elliott Building, 4th Floor, Room 45435, 190 Elizabeth Street, Toronto, Ontario M5G 2C4, Canada, or geysenba@uhures.utoronto.ca.

Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web

A Systematic Review

Günther Eysenbach, MD

John Powell, MSc, MRCPsych, MFPHM

Oliver Kuss, PhD

Eun-Ryoung Sa, MS

THE INTERNET HAS BECOME AN important mass medium for consumers seeking health information and health care services online.¹ A recent concern and public health issue has been the quality of health information on the World Wide Web. However, the scale of the problem and the "epidemiology" (distribution and determinants) of poor health information on the Web are still unclear, as is their impact on public health and the question of whether poor health information on the Web is a problem at all.² Many studies have been conducted to describe, critically appraise, and analyze consumer health information on the Web. These typically report proportions of inaccurate or imperfect information as estimates of the prevalence of flawed information or the risk of encountering misinformation on the Web.

However, to date no systematic and comprehensive synthesis of the methodology and evidence has been attempted. Two previous systematic reviews focused on compiling quality criteria and rating instruments, but did not synthesize evaluation results. Jadad and Gagliari³ reviewed non-research-based rating systems (eg, cri-

Context The quality of consumer health information on the World Wide Web is an important issue for medicine, but to date no systematic and comprehensive synthesis of the methods and evidence has been performed.

Objectives To establish a methodological framework on how quality on the Web is evaluated in practice, to determine the heterogeneity of the results and conclusions, and to compare the methodological rigor of these studies, to determine to what extent the conclusions depend on the methodology used, and to suggest future directions for research.

Data Sources We searched MEDLINE and PREMEDLINE (1966 through September 2001), Science Citation Index (1997 through September 2001), Social Sciences Citation Index (1997 through September 2001), Arts and Humanities Citation Index (1997 through September 2001), LISA (1969 through July 2001), CINAHL (1982 through July 2001), PsychINFO (1988 through September 2001), EMBASE (1988 through June 2001), and SIGLE (1980 through June 2001). We also conducted hand searches, general Internet searches, and a personal bibliographic database search.

Study Selection We included published and unpublished empirical studies in any language in which investigators searched the Web systematically for specific health information, evaluated the quality of Web sites or pages, and reported quantitative results. We screened 7830 citations and retrieved 170 potentially eligible full articles. A total of 79 distinct studies met the inclusion criteria, evaluating 5941 health Web sites and 1329 Web pages, and reporting 408 evaluation results for 86 different quality criteria.

Data Extraction Two reviewers independently extracted study characteristics, medical domains, search strategies used, methods and criteria of quality assessment, results (percentage of sites or pages rated as inadequate pertaining to a quality criterion), and quality and rigor of study methods and reporting.

Data Synthesis Most frequently used quality criteria used include accuracy, completeness, readability, design, disclosures, and references provided. Fifty-five studies (70%) concluded that quality is a problem on the Web, 17 (22%) remained neutral, and 7 studies (9%) came to a positive conclusion. Positive studies scored significantly lower in search ($P=.02$) and evaluation ($P=.04$) methods.

Conclusions Due to differences in study methods and rigor, quality criteria, study population, and topic chosen, study results and conclusions on health-related Web sites vary widely. Operational definitions of quality criteria are needed.

JAMA. 2002;287:2691-2700

www.jama.com

Author Affiliations: Unit for Cybermedicine and E-Health, Department of Clinical Social Medicine, University of Heidelberg, Heidelberg, Germany (Dr Eysenbach); Health Services Research Unit, London School of Hygiene and Tropical Medicine, London, England (Dr Powell); Department of Medical Epidemiology, Biostatistics and Informatics, University of Halle-Wittenberg, Halle/Saale, Germany (Dr Kuss); and Global Health

Network Group, Department of Epidemiology, University of Pittsburgh, Pittsburgh, Pa (Ms Sa). Dr Eysenbach is now with the Centre for Global eHealth Innovation, Toronto General Hospital, Toronto, Ontario. **Corresponding Author and Reprints:** Günther Eysenbach, MD, Centre for Global eHealth Innovation, Toronto General Hospital, 190 Elizabeth St, Toronto, Ontario, Canada M5G 2C4 (e-mail: ey@t.com).

For editorial comment see p 2713.

Meta-analysis of information quality on the web

Eysenbach G, Powell J, Kuss O, Sa ER.

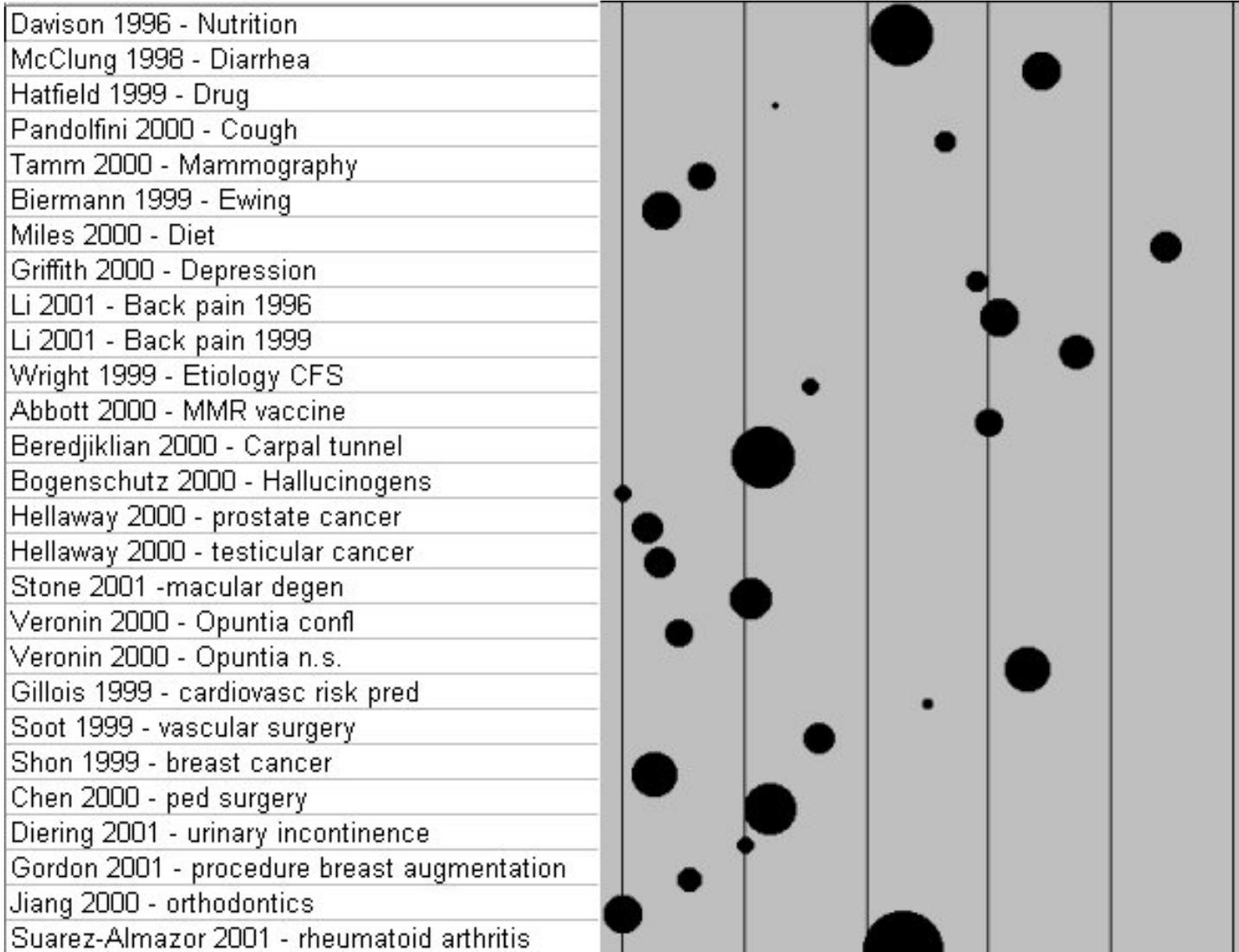
Empirical studies assessing the quality of health information for consumers on the World Wide Web: A systematic review.

JAMA 2002; 287: 2691-2700

Systematic review of studies evaluating health information on the web

(Eysenbach et al., 2002. **JAMA** 2002; 287: 2691-2700)

0% Inaccurate / non-evidence based information on the web 100%



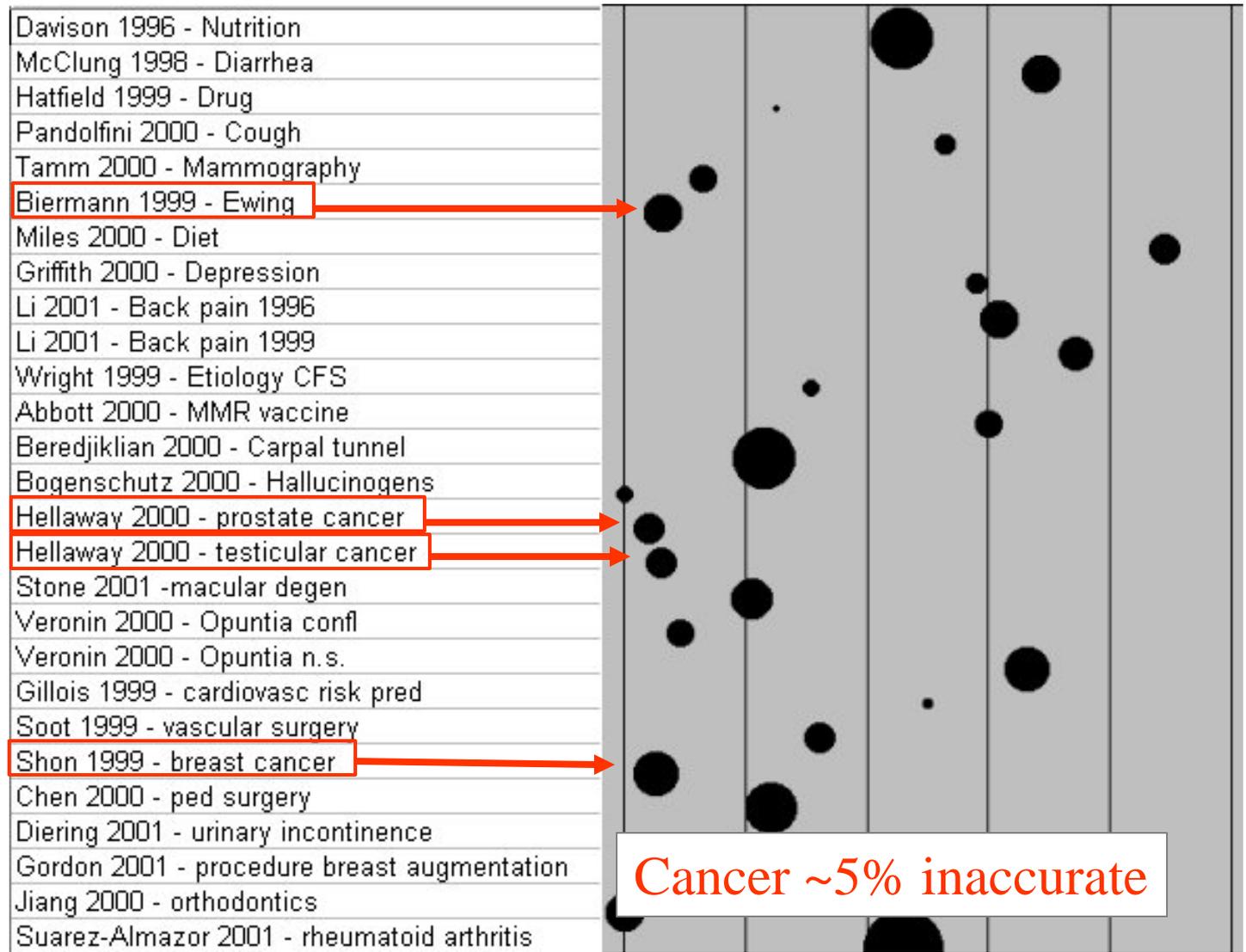
n=1781 websites
27 studies

Systematic review of studies evaluating health information on the web

(Eysenbach et al., 2002. **JAMA** 2002; 287: 2691-2700)

Inaccurate /
non-evidence based
information on the web

0% 100%



Cancer ~5% inaccurate

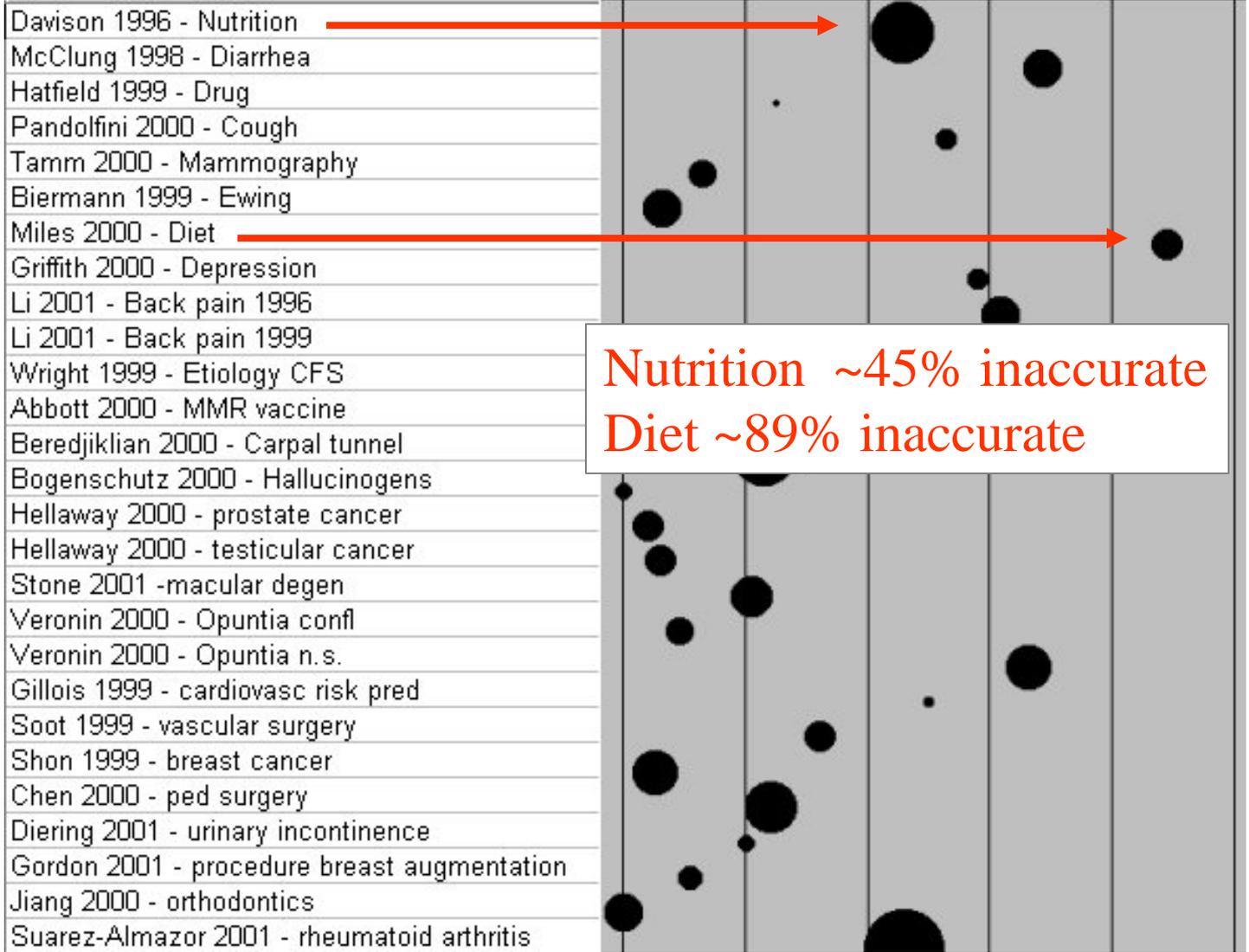
n=1781 websites
27 studies

Systematic review of studies evaluating health information on the web

(Eysenbach et al., 2002. **JAMA** 2002; 287: 2691-2700)

Inaccurate /
non-evidence based
information on the web

0% 100%



Nutrition ~45% inaccurate
Diet ~89% inaccurate

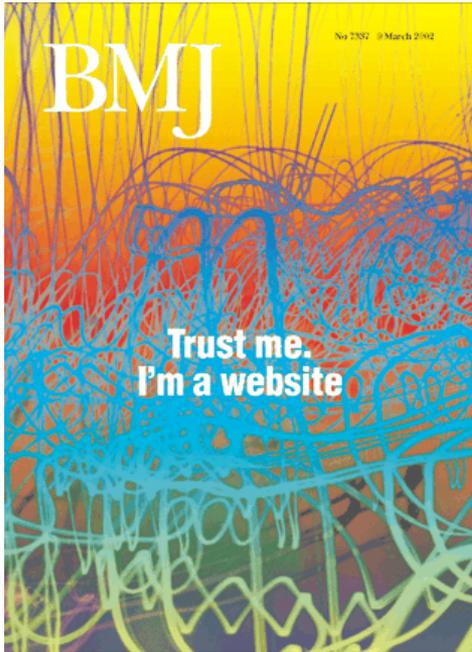
n=1781 websites
27 studies

„Technical“ (disclosure / transparency) consensus quality criteria for health websites

	Quality criterion	No of studies	Total number of websites/ webpages evaluated	Proportion (%) of sites/pages not complying		p-value
a	Disclosure of authorship	19	1636	(57.8)	**	<0.001
b	Disclosure of ownership	5	196	1.0		0.279
c	Sources clear	4	110	(44.5)	**	<0.001
d	Disclosure of sponsorship	7	738	(93.4)	*	0.006
e	Disclosure of advertising	3	119	30.3		0.113
f	Statement of Purpose	4	230	(51.7)	**	<0.001
g	General disclosures	3	298	(49.7)	**	0.001
h	Date of creation disclosed	5	284	83.1		0.564
i	Date of last update disclosed	7	801	(58.8)	*	0.049
j	Date of creation or update disclosed	12	1366	(63.7)	**	<0.001
k	Authors credentials disclosed	9	1030	(70.6)	**	<0.001
l	Credentials of physicians disclosed	3	81	97.5		0.899
m	Authors affiliation disclosed	5	779	(42.5)	**	<0.001
n	Easy navigation (subjective rating)	4	326	(10.1)	**	<0.001
o	Internal search engine present	3	91	79.1		0.164
p	Links provided	4	238	(27.7)	*	0.005
q	References provided	30	2135	(68.9)	**	<0.001
r	Balanced evidence	3	182	(47.9)	**	<0.001
s	Writing style appropriate (subjective rating)	4	136	16.2		0.840
t	Feedback mechanisms provided	4	157	(14.0)	**	<0.001
u	Fax number provided	5	1322	(38.0)	*	0.024
v	Email provided	8	1642	(24.3)	**	0.001
w	General disclaimers provided	6	390	(75.9)	*	0.047
x	Copyright notice	4	318	40.9		0.093
y	Editorial review process	5	166	87.3		0.754
z	Hierarchy of evidence clear	4	89	76.4		0.747

Limitations of *descriptive* “infodemiology” studies

- Impossible to draw a “random” sample of websites (search engines -> selection bias)
- Limited comparability, no control group
- Evaluation of “accuracy”: The higher you set the bar (the more evidence-based the gold standard), the more “inaccurate” information you will find
- “Accuracy” often confused with “completeness” (sites not containing all elements from a guideline referred to as “inaccurate”)
- unit of evaluation often unclear (webpage vs website)
- “so what”?



BMJ Theme Issue
„Quality of health
information“
9 March 2002
(Volume 324, Issue
7337)

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

Gunther Eysenbach, Christian Köhler

Abstract

Objectives To describe techniques for retrieval and appraisal used by consumers when they search for health information on the internet.

Design Qualitative study using focus groups, naturalistic observation of consumers searching the world wide web in a usability laboratory, and in-depth interviews.

Participants A total of 21 users of the internet participated in three focus group sessions. 17 participants were given a series of health questions and observed in a usability laboratory setting while retrieving health information from the web; this was followed by in-depth interviews.

Setting Heidelberg, Germany.

Results Although their search technique was often suboptimal, internet users successfully found health information to answer questions in an average of 5 minutes 42 seconds (median 4 minutes 18 seconds) per question. Participants in focus groups said that when assessing the credibility of a website they primarily looked for the source, a professional design, a scientific or official touch, language, and ease of use. However, in the observational study, no participants checked any “about us” sections of websites, disclaimers, or disclosure statements. In the post-search interviews, it emerged that very few participants had noticed and remembered which websites they had retrieved information from.

Conclusions Further observational studies are needed to design and evaluate educational and technological innovations for guiding consumers to high quality health information on the web.

Introduction

Little is known about how consumers retrieve and assess health information on the world wide web. Some

surveys have elicited data by using semistructured questionnaires or focus groups,¹⁻⁵ but little (if any) unobtrusive observational research has been done to explore how consumers are actually surfing the web. Although several criteria for quality of health websites have been proposed—including disclosure of site owners, authors, and update cycle^{4,5}—little or nothing is known about whether and to what degree such markers are recognised or even looked at by consumers or what other credibility markers consumers are looking for. We aimed to obtain qualitative and semiquantitative data to generate some hypotheses on how consumers might search for and appraise health information.

Methods

We used multiple methods of data collection that are commonly used in studies of human-computer interactions,⁶ combining focus groups,⁷ naturalistic observation of consumers searching the internet, and post-search in-depth interviews. Two researchers independently analysed transcripts by using N5 (NUD*IST 5.0; QSR International, Melbourne) with the grounded theory approach.⁸

Participants in the focus groups and the observational study were mostly healthy volunteers recruited through newspaper advertisements seeking people who had already searched for health information on the web. We selected them on a first come first served basis.

Focus groups

We held three focus group sessions with 6-8 participants each (21 participants in total: five men, 16 women; mean and median age 37, range 19-71 years) in March 2001. Self reported internet experience of the participants ranged from 17 to 84 months (mean 46 months; median 42 months). Each session was facilitated by a skilled moderator and lasted about two

Unit for
Cybermedicine and
eHealth,
Department of
Clinical Social
Medicine,
University of
Heidelberg,
Berghelmer Str 58,
69115 Heidelberg,
Germany

Gunther Eysenbach
senior researcher
Christian Köhler
researcher

Correspondence to:
G Eysenbach
ey@yi.com

BMJ 2002;324:573-7

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



- Focus Groups

- How do you search for information?
- Which markers for quality / credibility do you use?



- Usability lab:

- Giving health questions to consumers
- Let participants search the web and try to come up with an answer from the web
- Sessions logged, videotaped



- In-depth interviews:

- post-interview with consumer
- qualitative analysis

Credibility criteria applied by consumers (from focus groups)

- **Authority of source** (not confirmed in usability tests)
- **Layout and appearance**
- **Advertising**
- **Readability**
- **Outbound links**
- **Picture of the site owner**
- **Email**
- **Credentials and qualifications**
- **Updating of content**
- **Quality seal and third party endorsements**

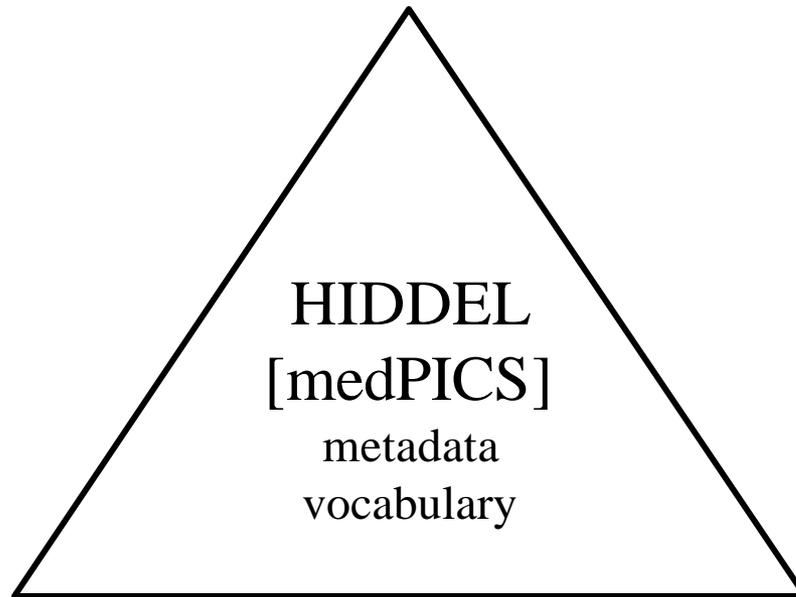
Role for WHO/PAHO

- Facilitate education and training
- Help creating a “web of trust” (semantic web)
- Foster and disseminate research, create collaborating centres in the field of consumer health informatics
- Promote best practices, incl. use of metadata
- Facilitate and disseminate standards, eg
HIDDEL

HIDDEL - XML/RDF

Health Information Disclosure Description & Evaluation Language

Health information provider
describes and discloses properties of the service



HIDDEL
[medPICS]
metadata
vocabulary

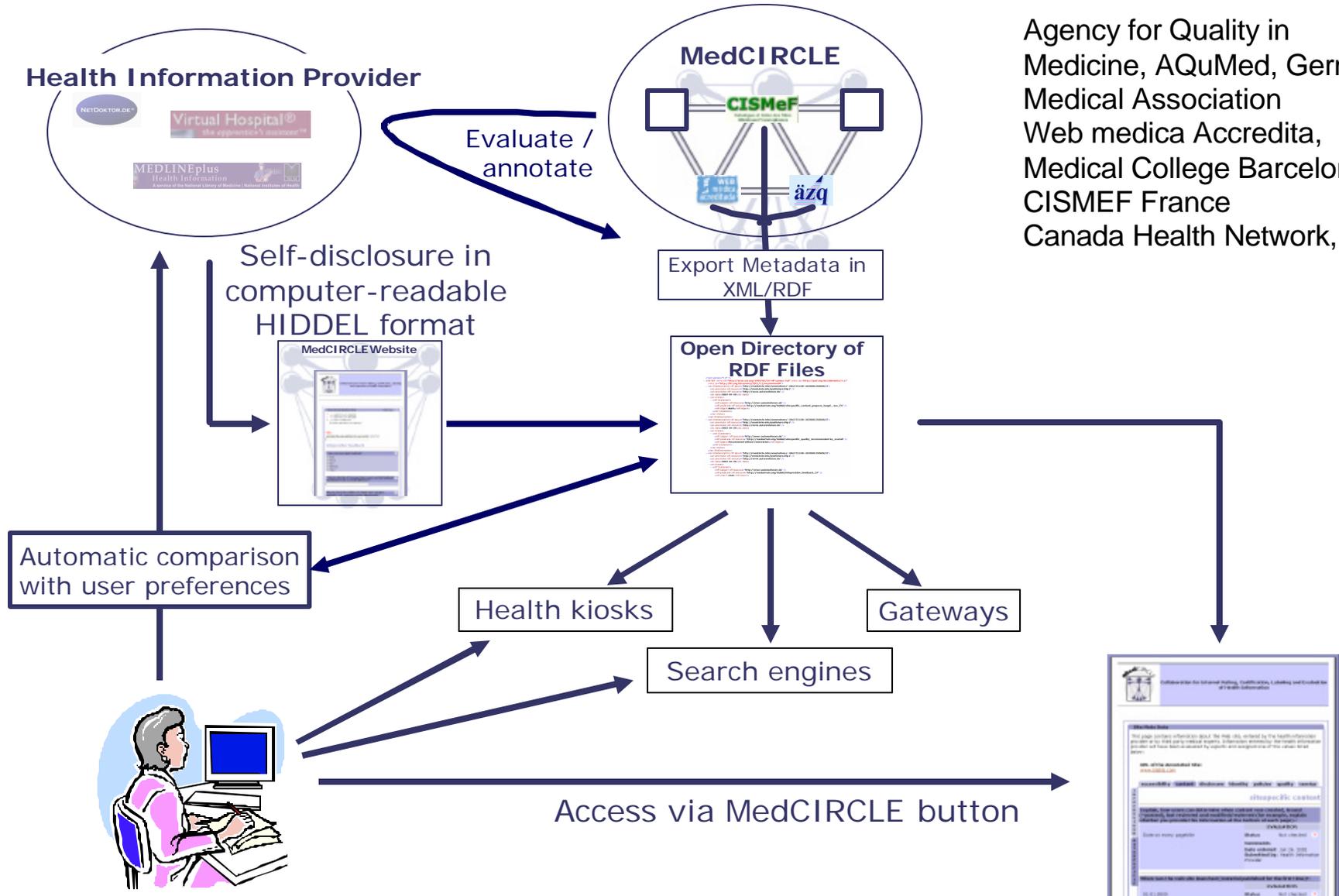
User
describes needs and expectations

Third party
(expert, rater)
describes and annotates
information and services

MedCIRCLE: Collaboration for Internet Rating, Certification, Labelling and Evaluation



Agency for Quality in Medicine, AQuMed, German Medical Association
 Web medica Accredita,
 Medical College Barcelona
 CISMEF France
 Canada Health Network, ...





A collection of
Search the archive:

[about](#) | [search options](#) | [browse](#) | [help](#) | [feedback](#) | [registration](#) | [privacy and usage rules](#)



Metadata coming from health information provider

Metadata / certification info from third parties



Website Annotation System

Site Meta Data

This page contains information about the Web site, entered by the health information provider or by third party medical experts. Information entered by the health information provider will have been evaluated by experts. For more information about the displayed information click on the help icon.

URL of the Annotated Site: <http://www.aqs.de>

- content
- disclosure
- identity
- policies
- quality
- service

INFORMATION sitespecific content

The resource was last modified (including minor changes such as correcting spelling errors or changes in layout) on:

✓	22.01.2003	---
	(Health Info Provider, Jan 22, 2003)	(fiene, Jan 22, 2003)

The content of the site for the need to update/revise is reviewed:

?	---	---

The content was last revised substantially (i.e. beyond correcting spelling errors or changes in layout) on:

✓	22.01.2003	---
	(Health Info Provider, Jan 22, 2003)	(fiene, Jan 22, 2003)

Following entities (persons, organisations, institutions) are making contributions to the content of the resource:

✓	Ärztekammern Ärztliche Zentralstelle Qualitätssicherung (ÄZQ), Köln Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Düsseldorf Bundesgeschäftsstelle Qualitätssicherung (BQS), Düsseldorf Bundesministerium für Gesundheit Gesundheitsministerkonferenz (GMK)	---
---	--	-----

INFORMATION SPECIFIC

Client (user)

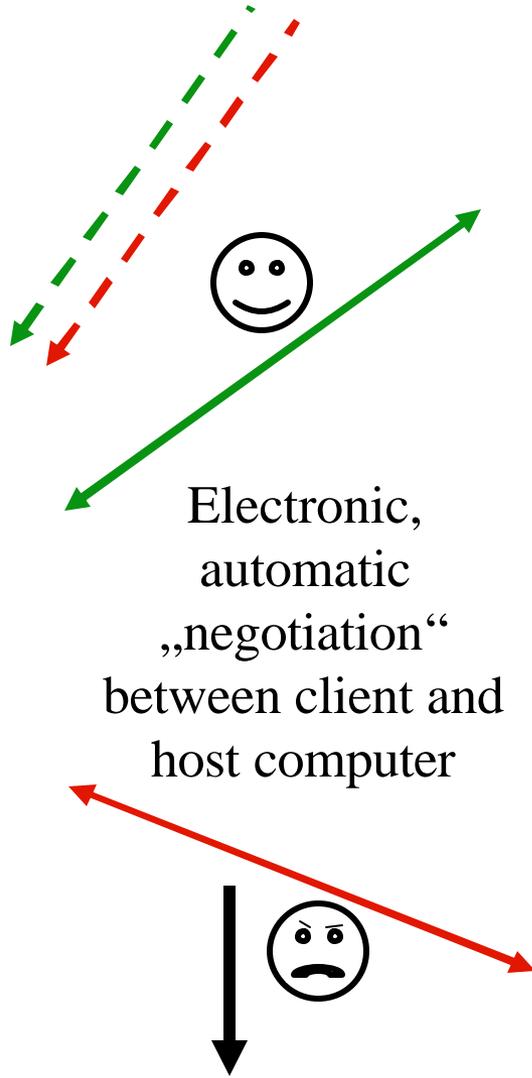


Target_group =
adult layperson

Funding_Source ¹
pharmaceutical
company

User sets preferences

Third party



Host (info provider)

Target_group =
adult layperson

Funding_Source =
public funding

Target_group =
physicians

Funding_Source =
pharmaceutical
company

Metadata encoded
disclosure information of
the information provider +
third party statements

Alerts,
Advice



MedCIRCLE collaboration

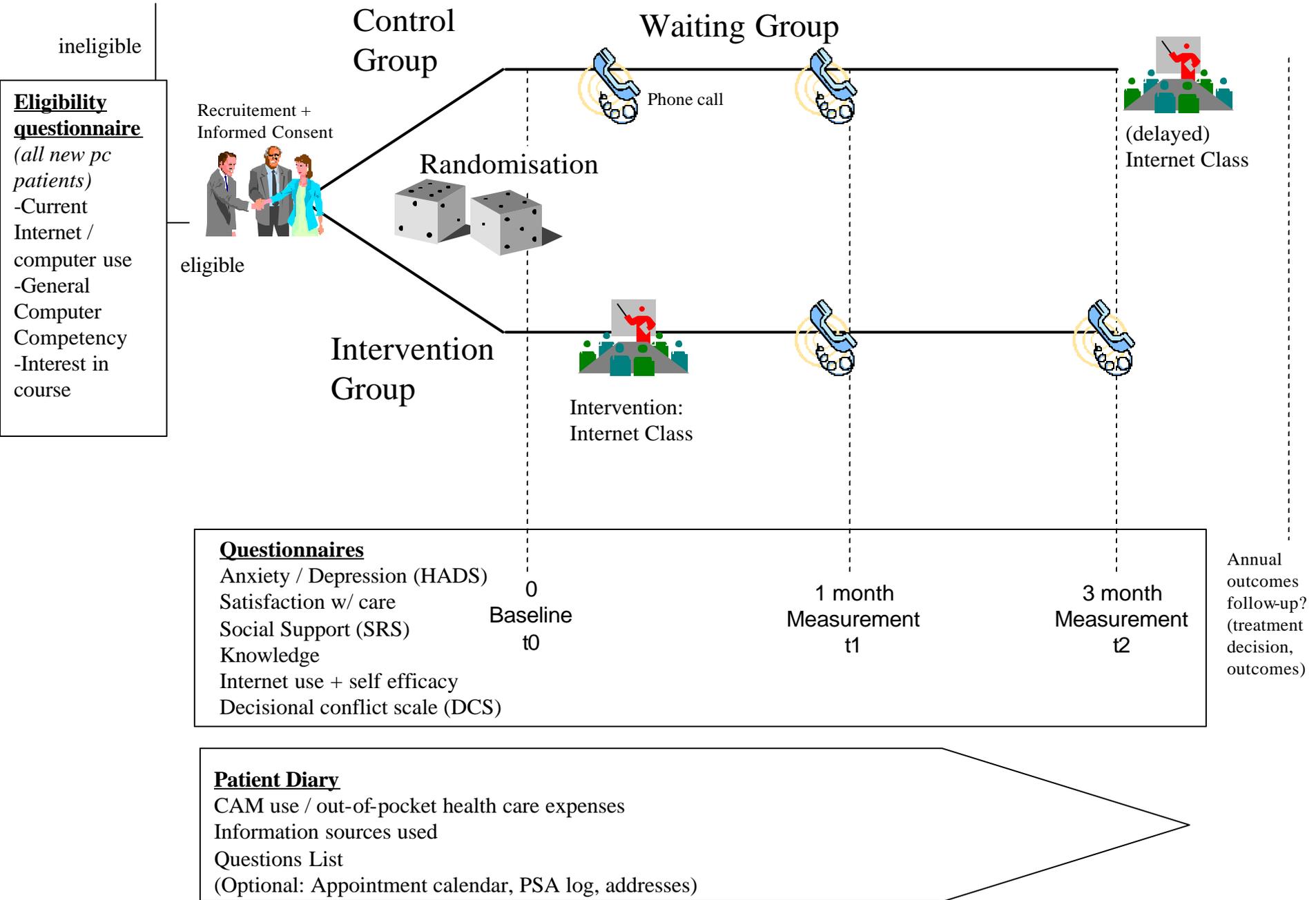
- A model for global collaboration with the aim to guide consumers to trustworthy health information
- Could be endorsed/adopted by WHO

Impact of the Internet?

- Few high quality studies have assessed the effectiveness of health care intervention delivered over the internet
- It can be difficult to distinguish between the effectiveness of the intervention and the influence of the medium
- Interventions that have been shown to be effective through other media, such as cognitive behavioural therapy and decision aids, have been effectively provided via the internet
- There is little information as yet whether use of the internet by consumers affects the interaction between consumers and health professionals and whether use of the internet could enhance or interfere with the relationship.
- Knowledge of disease and self management of disease may be positively influenced by internet interventions, but this effect can be limited by the high rates of drop-outs from studies
- Qualitative studies indicate that participants feel that there is an increase in social support and enablement as a result of the use of the internet, but this has not been evaluated fully in quantitative studies.
- Although there is the potential for the internet to increase access to information and some forms of health care interventions, at present there remain many barriers to access.

Doust J, Del Mar C, Pietzrak E, Brown R. The Impact of the Internet on Consumers Health Behavior: a Critical Literature Review. *J Med Internet Res* (in press)

The I3MPACT study



Thank you!

Dr Gunther Eysenbach,
Email: geysenba@uhnres.utoronto.ca,
Personal Homepage: <http://yi.com/ey/>

Acknowledgements: The work presented here is funded by the European Union (Safer Internet Action Plan), the Change Foundation Toronto, and the Canadian Institutes for Health Research (CIHR)

Conflicts of interest/Disclosures:
see <http://yi.com/home/EysenbachGunther/disclosures.htm>