Making Sense of Cancer Risk Calculators on the Web

Andrea Gurmankin Levy, PhD, MBe[1,2], Seema S. Sonnad, PhD[3,4], Jibby E. Kurichi[3], Melani Sherman, BA[3], and Katrina Armstrong, MD, MSCe[3,4,5]

[1]Center for Community-Based Research, Dana-Farber Cancer Institute, Boston, MA, USA; [2]Department of Society, Human Development and Health, Harvard School of Public Health, Boston, MA, USA; [3]School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [4]Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA; [5]Abramson Cancer Center, University of Pennsylvania, School of Medicine, Philadelphia, PA, USA.

OBJECTIVE: Cancer risk calculators on the internet have the potential to provide users with valuable information about their individual cancer risk. However, the lack of oversight of these sites raises concerns about low quality and inconsistent information. These concerns led us to evaluate internet cancer risk calculators.

DESIGN: After a systematic search to find all cancer risk calculators on the internet, we reviewed the content of each site for information that users should seek to evaluate the quality of a website. We then examined the consistency of the breast cancer risk calculators by having 27 women complete 10 of the breast cancer risk calculators for themselves. We also completed the breast cancer risk calculators for a hypothetical high- and low-risk woman, and compared the output to Surveillance Epidemiology and End Results estimates for the average same-age and same-race woman.

RESULTS: Nineteen sites were found, 13 of which calculate breast cancer risk. Most sites do not provide the information users need to evaluate the legitimacy of a website. The breast cancer calculator sites vary in the risk factors they assess to calculate breast cancer risk, how they operationalize each risk factor and in the risk estimate they provide for the same individual.

CONCLUSIONS: Internet cancer risk calculators have the potential to provide a public health benefit by educating individuals about their risks and potentially encouraging preventive health behaviors. However, our evaluation of internet calculators revealed several problems that call into question the accuracy of the information that they provide. This may lead the users of these sites to make inappropriate medical decisions on the basis of misinformation.

Electronic supplementary material The online version of this article (doi: 10.1007/s11606-007-0484-x) contains supplementary material, which is available to authorized users.

The manuscript and data have not been published previously and are not under consideration for publication elsewhere. This research did not have financial support from any source.

Received December 15, 2006
Revised June 8, 2007
Accepted October 29, 2007
Published online January 11, 2008

KEY WORDS: internet; web risk calculators; cancer risk; breast cancer.
J Gen Intern Med 23(3):229–35
DOI: 10.1007/s11606-007-0484-x

INTRODUCTION

Increasingly, people turn to the internet for health information. In 2002, 63% of the population had ever gone online, and 51% of individuals who used the internet had used it to search for health information. Almost half of the general public who had searched for cancer information had chosen the internet as their first source for information.[1] In addition to providing general health information, the internet also enables people to obtain information tailored to their individual characteristics, including the prediction of an individual's disease risk by internet risk calculators.

Cancer risk prediction has become an increasingly important focus for the medical profession and the lay public. The prediction of individual cancer risk from risk factors has traditionally been based upon predictive models derived from epidemiologic studies of cancer incidence.[2] Although considerable effort has been devoted to identifying cancer risk factors, there are relatively few published cancer risk prediction models. The most well known is the Gail breast cancer prediction model.[3,4] Recent attempts to improve on the Gail model through the inclusion of additional risk factors have had limited success and have not yet been validated in multiple data sources.[5,6] Risk prediction models have recently been developed for lung and prostate cancer[7,8] but have not yet been developed in many other cancers.

Cancer risk calculators are a relatively new development on the internet, although they are already widely used (the Harvard Your Disease Risk website, for instance, receives 76 million visitors annually; Graham Colditz, August 15, 2005). These tools ask the site visitor for risk factor information and then calculate the individual's risk of the cancer. These websites have the potential to provide benefit by disseminating information about individual cancer risk that can improve knowledge, decision making, and preventive health behavior.

However, the internet is unregulated; anyone can post a health information website or a cancer risk assessment website. Perhaps, as a result, studies have found some health information on the internet to be of poor quality.[9–12] Inaccurate information may lead to inappropriate medical decisions,[13–15] such as whether or not to undergo recommended cancer screening or change a diet or exercise regimen. These concerns may be greater for internet risk calculators than general health

information websites because the latter provide aggregate information, whereas the former provide information that is specific to the individual, leaving the user less likely to question whether or how well the information applies to them.

However, little is currently known about the quality of the information provided by cancer risk calculators on the internet.[16,17] Thus, the purpose of this study was to evaluate cancer risk calculators on the internet with respect to the content of the sites (e.g., whether information is provided for a user to evaluate the credibility of the site). Because the majority of internet cancer risk calculators focus on breast cancer, we were able to conduct a more extensive evaluation of breast cancer risk calculator sites, examining the consistency of the risk factors assessed to calculate risk and variation in and consistency of the risk estimates provided. Although even the best risk calculators are known to be imperfect predictors,[18] information about the content of these sites could be of great use to the millions of people who visit these sites annually.

METHODS

We conducted a systematic search of the internet to identify all sites with a risk calculator for any cancer. We searched the terms "cancer risk" (without quotes) in each of the following search engines using Microsoft Internet Explorer: Google, MSN, Yahoo,

Table 1. Content of the Cancer Risk Calculators on the Internet

Primary affiliation	Cancers assessed	Source information	Has a disclaimer	Affiliation information	Information on when site was last updated	Identify algorithm or model used
1. About.Com, New York Times Company*	Colon	No	No	*Yes*	No	No
2. AstraZeneca*	Breast	*Yes*	*Yes*	*Yes*	No	*Yes (Gail model)*
3. Department of Surgery at Saint Vincent Charity Hospital*	Breast	No	No	*Yes*	No	No
4. GenneX Healthcare Technologies	Breast	No	*Yes*	*Yes*	No	No
5. Griffin Hospital	Colon	No	No	*Yes*	No	No
6. Harvard Center for Cancer Prevention	Bladder, breast, cervical, colon, kidney, lung, melanoma, ovarian, pancreatic, prostate, stomach, uterine	*Yes*	*Yes*	*Yes*	*Yes*	No
7. Mary Bird Perkins Cancer Center*	Bladder, breast, cervical, colon, kidney, lung, melanoma, ovarian, pancreatic, prostate, stomach, uterine	No	*Yes*	*Yes*	*Yes*	No
8. Medical Network Inc.	Bladder, colon, lung, ovarian, prostate, skin, uterine	*Yes*	*Yes*	*Yes*	*Yes*	No
9. Memorial Healthcare System	Breast	No	*Yes*	*Yes*	No	No
10. Memorial Sloan Kettering Cancer Center	Lung	*Yes*	*Yes*	*Yes*	Yes	No
11. MSNBC	Prostate	No	*No*	*Yes*	No	No
12. National Cancer Institute 42	Breast	*Yes*	*Yes*	*Yes*	No	No
13. National Surgical Breast and Bowel Project	Breast	No	*Yes*	*Yes*	No	No
14. Ohio State University Comprehensive Cancer Center	Cancer in general	*Yes*	*Yes*	*Yes*	No	No
15. RealAge	Breast, prostate	No	*Yes*	*Yes*	No	No
16. St. Luke's Episcopal Health System	Breast, skin	*Yes*	*Yes*	*Yes*	Yes	No
17. Steven Halls, M.D.	Breast	*Yes*	*Yes*	No	Yes	Yes (Gail model modified and Gail model)
18. Women's Cancer Network	Breast, ovarian, endometrial, cervical, vulvar, vaginal	No	*Yes*	*Yes*	No	No
19. Unidentified*	Breast	No	*Yes*	No	No	No
Totals		8	15	17	6	2

The italicized entries represent "yes": the presence of the information on the website. Website addresses: 1 http://coloncancer.about.com/od/gamesandquizzes/a/AreYouAtRisk4CC.htm, 2 http://www.bcancerrisk.com/consumer/test.asp?risk=1, 3 http://www.breastrisk.com/, 4 http://www.estronaut.com/a/breastInteractive2.htm, 5 http://www.griffinhealth.org/patientvisitor/patientservices/digestivecenter/screening.aspx, 6 http://www.yourdiseaserisk.harvard.edu/, 7 http://www.marybird.org/mycancerrisk/hccpquiz.pl?func=show&page=screening_test, 8 http://www.healthatoz.com/healthatoz/Atoz/tl/rq/quizbody.jsp, 9 http://www.memorial.org/, 10 http://www.mskcc.org/mskcc/html/12463.cfm, 11 http://www.msnbc.com/news/604780.asp?cp1=1, 12 http://bcra.nci.nih.gov/brc/, 13 http://www.breastcancerprevention.org/, 14 http://www.jamesline.com/patientsandvisitors/prevention/cancergenetics/#Start%20Session, 15 http://www.breastcancer.realage.com/content.aspx/topic/9, 16 http://www.sleh.com/sleh/index.cfm, 17 http://www.halls.md/breast/risk.htm, 18 http://www.wcn.org/interior.cfm?diseaseid=13&featureid=3, 19 http://www.breastcancerquiz.com/consumer/quiz/

* The site is no longer active.

Table 2. Risk Factors Assessed by the Breast Cancer Risk Calculators

Risk factor	Astra-Zeneca	Department of Surgery at St. Vincent Charity Hospital	GenneX Healthcare Technologies	Harvard Cancer for Cancer Prevention	Mary Bird Perkins Cancer Center	Memorial Health Care System
Age	No	*Yes*	*Yes*	*Yes*	*Yes*	No
Age at menarche	*Yes*	*Yes*	No	*Yes*	*Yes*	No
Age of first birth	*Yes*	*Yes*	No	*Yes*	*Yes*	No
Benign breast disease	No	*Yes*	No	*Yes*	*Yes*	No
Breastfeeding	No	*Yes*	No	*Yes*	*Yes*	No
Family history	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Hormone replacement therapy	No	*Yes*	No	*Yes*	*Yes*	No
Jewish ancestry	No	No	*Yes*	*Yes*	*Yes*	*Yes*
Menopause	No	*Yes*	No	*Yes*	*Yes*	No
Number of live births	No	*Yes*	No	*Yes*	*Yes*	No
Oral contraceptives	No	*Yes*	No	*Yes*	*Yes*	No

The italicized entries represent "yes": The risk factor is assessed on the respective website.

Lycos, and Excite. The first 1,000 hits from each search were visited to determine whether they included a cancer risk calculator. After excluding duplicates (sites with links to the same risk calculator) and non-English sites, a list of 19 unique cancer risk calculators was generated. Additional searches were conducted using the terms *cancer risk calculator*, *personal cancer risk*, and *individual cancer risk* but yielded no new calculators.

We developed a site-abstraction form to collect data on whether each of the following was present on each cancer risk calculator website: (1) source information (e.g., a medical reference), (2) information on when the site was last updated, (3) the algorithm or model used to calculate the risk estimate, (4) a disclaimer of liability, and (5) information on the website's affiliations. We also determined the format of the output of each risk calculator and which cancers each site addressed. For the breast cancer sites, we examined the risk factors they assessed and how they operationalize each risk factor.

With Institutional Review Board approval from the University of Pennsylvania, we then recruited a convenience sample of women from the general population (N=27) to complete ten of the breast cancer risk calculators for themselves and provide us with print-outs of their output from each site. Subjects had a mean age of 42 (range, 21–61), all were Caucasian and 43% were Ashkenazi Jewish. Their average lifetime breast cancer risk, according to the Gail model, was 14.12%. We compared the output the subjects provided using simple descriptive analyses, as well as calculations of the correlations of the estimates provided by each site for a given subject. To calculate these correlations, where necessary, we converted all nonnumeric scales used by sites (e.g., "very much below average," to "very much above average") to ordinal scales (e.g., a 7-point scale where 1="very much below average" and 7="very much above average").

We also compared the output provided by each site for two hypothetical patients, one created with characteristics leading to a low risk of breast cancer and the other with characteristics indicative of a high risk of breast cancer (Appendix 1). These hypothetical patients' risk factors were based on well-established breast cancer risk factors (Tables 2 and 3).[19–32] If an internet risk calculator asked questions about risk factors that were not in the description created (e.g., education), *unknown/not sure* was entered when that option was provided or were answered with the midpoint or average response and recorded. The latter was required only 9 times.

Because there is no gold standard for breast cancer risk prediction and the currently available risk prediction models have significant limitations, we provided a framework for assessing the risk estimates by comparing the outputs for high- and low-risk women to population-based incidence rates of breast cancer provided by Surveillance Epidemiology and End Results (SEER).[33] According to SEER data, the average 55-year-old white woman has a 2.5% 10-year breast cancer risk, 9.35% 30-year risk, and 11.5% lifetime risk. The average 55-year-old black woman has a 2.1% 10-year risk, 6.63% 30-year risk, and 7.75% lifetime risk. We predicted that the output

Table 3. Risk Factors Assessed by the Breast Cancer Risk Calculators

Risk factor	National Cancer Institute	National Surgical Breast and Bowel Project	RealAge	St. Luke's Episcopal Health System	Steven Halls, M.D.	Women's Cancer Network	Unidentified
Age	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	No
Age at menarche	*Yes*	*Yes*	*Yes*	No	*Yes*	*Yes*	*Yes*
Age of first birth	*Yes*	*Yes*	*Yes*	No	*Yes*	*Yes*	*Yes*
Benign breast disease	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	No
Breastfeeding	No	No	No	No	No	Yes	No
Family history	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Hormone replacement therapy	No	No	*Yes*	*Yes*	*Yes*	*Yes*	No
Jewish ancestry	*Yes*	*Yes*	No	No	*Yes*	*Yes*	No
Menopause	No	*Yes*	*Yes*	No	No	*Yes*	No
Number of live births	No	*Yes*	*Yes*	No	No	*Yes*	*Yes*
Oral contraceptives	No	No	*Yes*	No	*Yes*	*Yes*	*Yes*

The italicized entries represent "yes": The risk factor is assessed on the respective website.

for the high-risk women should be higher than the average SEER risk, whereas the output for the low-risk women should be lower than the average SEER risk. In addition, we examined the variation in the range between the estimates for the high- and low-risk women across the calculators.

RESULTS

Table 1 shows the internet risk calculators found in our search. There are 13 calculators for breast cancer risk, 5 for colon, 4 for prostate, 4 for lung, 3 for bladder, 3 for ovarian, 3 for melanoma, 2 for uterine, 2 for kidney, 2 for pancreatic, 2 for cervical, 1 for stomach, and 1 for general cancer risk.

Content of Websites

The websites varied in the information they provided to allow a user to assess the credibility of the site (Table 1). The minority of the sites included source information (7/19), update information (5/19), or identified the model or algorithm they used to calculate the risk estimate (2/19, for both, it was the Gail model or the modified Gail model), but the majority included a disclaimer (14/19) and affiliations (17/19). A sample disclaimer from the Medical Network Inc. website read: "The text presented on these pages is for your information only. It is not a substitute for professional medical advice. It may not represent your true individual medical situation. Do not use this information to diagnose or treat a health problem or disease without consulting a qualified healthcare provider. Please consult your healthcare provider if you have any questions or concerns."

The sites also differed in the format of the output they provided. Sites used numeric test scores with score explanations, graphs, comparative risk descriptions, verbal risk descriptions, and numeric risk estimates with varying time frames.

Risk Factors Measured

The breast cancer risk calculators also varied in the risk factors collected for the calculation. As shown in Tables 2 and 3, only some of the breast cancer sites asked users about each of several well-established breast cancer risk factors.[19–32] For instance, of the 13 breast cancer calculators, all assessed family history, 10 assessed age of menarche, 8 assessed Jewish ancestry, 7 assessed hormone therapy use, and 4 assessed breastfeeding.

The websites also varied in how they operationalized each risk factor. For example, across the breast cancer websites, family history was assessed each of the following ways: (1) Yes/no first degree relative had cancer, and if yes, cancer type and age at diagnosis for each. (2) Yes/no mother had breast cancer. (3) Yes/no mother or sister had breast cancer (followed by age of diagnosis question for each on some sites). (4) How many of your mother, sisters, and daughters had breast cancer? (5) Who in your family had breast cancer? Choices: (a) mother, (b) sister, (c)

Table 4. Comparison of Breast Cancer Risk Calculators' Output for 4 Sample Subjects

	Subject 16 (21 y old)	Subject 1 (26 y old)	Subject 37 (42 y old)	Subject 31 (48 y old)
5-year risk				
NCI	*	0.5%	0.8%	1.3%
Halls, MD	<0.2%	<0.2%	0.7%	1.7%
National Surgical Breast and Bowel Project	Website only for those >35 y	Website only for those >35 yrs	0.74%	3.76%
10 year risk				
GenneX	0.42%	2.86%	2.53%	3.54%
Halls, MD	1.4%	1.6%	1.9%	3.6%
30 year risk				
GenneX	4%	30%	9%	13%
Halls, MD	6.6%	7.7%	6.8%	11.1%
RealAge	*	3.3%	7.5%	*
Lifetime risk				
NCI	20.1%	17.7%	11.9%	12.1%
Halls, MD	16.9%	18.4%	11.1%	15.6%
Comparative risk				
Harvard Center for Cancer Prevention	Average	Above average	Below average	Much above average
Women's Cancer Network	Average risk	Average risk	Average risk	Average risk
Mary Bird Perkins	Average	Above average	Below average	Above average
Verbal risk				
AstraZeneca	You scored a total of 1. If your answers added up to 2 or more, you *may* be at high risk for developing breast cancer	You scored a total of 3. If your answers added up to 2 or more, you *may* be at high risk for developing breast cancer	You scored a total of 2. If your answers added up to 2 or more, you *may* be at high risk for developing breast cancer	You scored a total of 2. If your answers added up to 2 or more, you *may* be at high risk for developing breast cancer
Breastcancerquiz.com	Website only for those >35 y	You scored a total of 3. If your answers added up to 2 or more, you *may* be at high risk for developing breast cancer	You scored a total of 2. If your answers added up to 2 or more, you *may* be at high risk for developing breast cancer	You scored a total of 2. If your answers added up to 2 or more, you *may* be at high risk for developing breast cancer

*Subject did not provide her output.

two or more sisters or mother and sister, (d) maternal grandmother or aunt, (e) paternal grandmother or aunt, (f) two or more grandmothers or aunt. (6) Check box for who had any cancer before age 50: you, mother, sister, or daughter.

Calculator Output

The websites also differed in the output they provided for the same person. In fact, each website's output for each subject was not consistently correlated with the other websites' output. The average correlation coefficient for the estimates provided by the 10 breast cancer websites for the study participants was 0.42 with a range from −0.74 to +1.00. Eighteen (40%) of the 45 pairs reached statistical significance. However, it is important to note that some internet sites were completed by only a subset of the subjects (an omission by the subject), giving those sites lower power to reach statistical significance because of a smaller sample size.

Table 4 shows the output from the breast cancer sites for four sample subjects whose risk factors are presented in Appendix 2 (the anonymous output from each site for all subjects is available on the JGIM website). Although the variation in the format of the output from each website makes comparison of the nonstandardized outputs difficult, some comparisons can be made. For instance, the three sites that provided a comparative statement of risk provided different results from each other for three of the four sample subjects. Comparisons of the sites with numeric estimates to each other and to the comparative estimates (assuming an average woman's risk of 12.5%)[34] revealed some small and some large differences. For example, for subject 1, the Halls MD and NCI sites, which presented lifetime risk estimates did not differ much from each other (18.4% vs 17.7%), but they did deviate from the comparative risk estimate of "average risk" from the Women's Cancer Network site. In addition, the three websites presenting 30-year risk estimates (GenneX, Halls MD, and RealAge) differed greatly from each other (30% vs 7.7% vs 3.3% for subject 1; Table 4).

As shown in Table 5, the output provided by the websites for the hypothetical high- and low-risk women also varied. For instance, for the high-risk woman, the 30-year estimates for Halls MD and the National Surgical Breast and Bowel Project did not differ much (38.7% vs 35%) but were very different from the GenneX estimate (64%). The lifetime estimates for the high-risk woman were similar for NCI and Halls MD (40.5% vs 42.3%) but differed greatly from the RealAge estimate (15.2%).

Furthermore, when comparing numeric estimates from each site for the hypothetical high- and low-risk women to the appropriate SEER estimate (matched by gender, age, and time frame of the estimate), there was a fivefold difference across internet calculators' estimates for the high-risk woman and a twofold difference for the low-risk woman. For the high-risk woman, the GenneX site provided a 30-year estimate that was 6.8 times the SEER 30-year estimate for the average same-age, Caucasian woman (64% from GenneX vs 9.35% from SEER), whereas the RealAge site provided a lifetime risk estimate that was 1.3 times the SEER lifetime estimate for the average same-age white woman (15.2% vs 11.5%). For the low-risk woman, the GenneX site provided a 30-year estimate that was nearly identical to the SEER estimate for the average same-age black woman (6% vs 7.75%), whereas the National Surgical Breast and Bowel Project site provided an estimate that was less than half of the SEER estimate (3% vs 6.6%).

Table 5 Breast Cancer Risk Calculators' Output for Hypothetical Low- and High-Risk Woman

	Output for hypothetical low risk woman	Output for hypothetical high risk woman
5 year risk		
NCI	0.6%	7.2%
National Surgical Breast and Bowel Project	0.57%	7.17%
10 year risk		
GenneX	2.7%	24.5%
Halls, MD	1.4%	16.1%
RealAge	0.8%	4.6%
30 year risk		
GenneX	6%	64%
Halls, MD	4.2%	38.7%
National Surgical Breast and Bowel Project	3%	35%
Lifetime risk		
NCI	3.7%	40.5%
Halls, MD	5%	42.3%
RealAge	2.5%	15.2%
Comparative risk		
Harvard Center for Cancer Prevention	Below average	Very much above average
Women's Cancer Network	Low risk	Moderate risk
Mary Bird Perkins	Below average	Very much above average
Verbal risk		
AstraZeneca	You scored a total of 0. If your answers added up to 2 or more, you may be at high risk for developing breast cancer	You scored a total of 8. If your answers added up to 2 or more, you may be at high risk for developing breast cancer
Breastcancerquiz.com	You scored a total of 0. If your answers added up to 2 or more, you may be at high risk for developing breast cancer	You scored a total of 8. If your answers added up to 2 or more, you may be at high risk for developing breast cancer
Department of Surgery at St. Vincent Charity Hospital	Your risk score is 1. 0–3, average risk; 4–7, moderate risk; 8 or higher, high risk	Your risk score is 8. 0–3, average risk; 4–7, moderate risk; 8 or higher, high risk
Memorial Healthcare System	You appear to not be at higher risk for breast or ovarian cancer	Please call our Cancer Risk Counseling Service at xxx-xxx-xxxx to evaluate your risk…

SEER estimate for same-age and same-race woman. Low risk: 2.1%, 10-year risk; 6.6%, 30-year risk; 7.75%, lifetime risk. High risk: 2.5%, 10-year risk; 9.4%, 30-year risk; 11.5%, lifetime risk. The 13th site, St. Luke's Episcopal Health System, is omitted from this table because it does not provide a risk estimate; it only provides behavioral recommendations for the user.

DISCUSSION

Cancer risk calculators on the internet disseminate individualized cancer risk estimates and recommendations to site visitors. Thus, cancer risk calculators on the internet can be a valuable tool for improving people's understanding of their cancer risks and encouraging preventive health measures, which can reduce cancer risk.[28,29,35,36] However, the lack of regulation or oversight of information on the internet raises concerns

about the quality and accuracy of the risk information that these calculators provide. Our study provides evidence to support these concerns on several levels.

First, a prudent internet user should evaluate the website before accepting the information it provides, but we found that many of the sites do not provide the information (e.g., source information) that a user would need to conduct this evaluation. Others have raised this same concern about cardiovascular risk calculators on the internet[37] and health information websites more generally.[11] The absence of this information on some websites raises the question of whether a reliable source exists.

Second, the breast cancer risk calculators vary in the risk factors that they use to assess a visitor's risk of breast cancer and how they operationalize these risk factors. Some calculators do not include well-established risk factors in their calculation. Third, the breast cancer risk calculators vary in the estimates they provide for the same woman's risk of breast cancer, and their estimates are not consistently correlated with each other. This variation strongly suggests that the risk estimates from some of the calculators are inaccurate. Because so few sites calculate risks of cancers other than breast cancer, we could not examine whether these two issues apply to calculators for other cancers as well.

These variations and potential inaccuracies in internet risk calculator output may generate inaccurate perceptions of cancer risks in the general public, which may lead to inappropriate decisions and behaviors. For instance, a visitor who is told that her risk of breast cancer is lower than it actually is might decide not to get routine breast screening, placing her at increased risk of death from breast cancer. Conversely, a visitor who is told that her risk of breast cancer is twice as high as it actually is may become distressed or obtain unnecessary screening tests or even invasive diagnostic procedures.

These scenarios are real possibilities. Studies have shown that the general public views information on the internet as trustworthy,[38,39] and 71% of online health information-seekers said that information they found on the internet influenced their decisions regarding their own healthcare or that of a loved one.[40] Thus, the dissemination of inaccurate risk information may have significant consequences for people's health decisions and health behavior.

The results of this study must be considered within its limitations. First, the subjects in this study were from a convenience sample and were not representative of any particular patient population or the US population. However, because this study compared internet risk calculator output either within-subject or between a subject and corresponding SEER estimates, the nature of the sample should have no impact on the findings. Second, the absence of a gold standard (i.e., the true risk estimate for a given person) and the variation in output format hinders our ability to evaluate the risk calculator output provided by each site for each subject. However, the SEER estimates provided information on the average same-sex and same-race patient as the hypothetical high- and low-risk women, which facilitated an evaluation of the sites' output. Furthermore, even without knowing the true risk for a given person, the variation in the output across the sites makes it clear that at least some sites are providing inaccurate estimates; the different estimates cannot all be correct. Third, it is also difficult to evaluate these sites without considering the algorithm they used (and how it was validated) to calculate risk estimates. However, upon requesting information on the algorithm used by the sites that did not document this information on their webpage, we either received no response or a denial of our request. Finally, it is important to note that the World Wide Web is an ever-changing environment, and as a result, the conclusions from this study are time sensitive. In fact, during the review of this manuscript, 5 of the sites became inactive.

Nevertheless, the results of our evaluation of cancer risk calculators on the internet suggest several points for further consideration. First, the results highlight the need for oversight of these websites and correction of existing inaccuracies. The provision of inaccurate risk information by these sites may have significant consequences for patients' health outcomes. Second, these results can provide guidance to assist healthcare providers in determining which sites they ought to recommend to their patients and to assist patients in selecting the site that is most useful to them. Finally, these results also highlight the difficulties in developing a "gold-standard" risk prediction model. The multiple risk factors for any given cancer are generally not available from a single study requiring statistical methods to aggregate their effects into a summative estimate. These methods can be misleading if they fail to take into account lack of independence among the risk factors or the variation in their effects across different population subgroups. These difficulties contribute to the lack of a dominant approach to risk assessment and the variation across the websites.

There is little doubt that making individualized health information available on the internet for easy access by the general public is a worthwhile goal. However, if this information is not of sufficient quality, it will have the reverse of the intended effect and may increase patient and health system burden rather than reducing it.

Acknowledgment: *The authors are grateful to Sanghamitra Savadatti, Nicole Maddox, and the participants in the study.*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Andrea Gurmankin Levy, PhD, MBe, Center for Community-Based Research, Dana-Farber Cancer Institute, 44 Binney St, LW 618, Boston, MA 02115, USA (e-mail: Adg11@cornell.edu).*

REFERENCES

1. **Viswanath K**. Science and society: the communications revolution and cancer control. Nat Rev Cancer. 2005;5(10):828–35.
2. **Armstrong K, Eisen A, Weber B.** Assessing the risk of breast cancer. N Engl J Med. 2000;342(8):564–71.
3. **Gail MH, Brinton LA, Byar DP, et al.** Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst. 1989;81:1879–86.
4. **Costantino JP, Gail MH, Pee D, et al.** Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst. 1999;91(18):1541–8.
5. **Bondy ML, Newman LA.** Assessing breast cancer risk: evolution of the Gail Model. J Natl Cancer Inst. 2006;98(17):1172–3.
6. **Barlow WE, White E, Ballard-Barbash R, et al.** Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006;98(17):1204–14.
7. **Bach PB, Kattan MW, Thornquist MD, et al.** Variations in lung cancer risk among smokers. J Natl Cancer Inst. 2003;95(6):470–8.
8. **Thompson IM, Ankerst DP, Chi C, et al.** Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst. 2006;98(8):529–34.
9. **Impicciatore P, Pandolfini C, Casella N, et al.** Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home. BMJ. 1997;314(7098):1875–9.
10. **Allen JW, Finch RJ, Coleman MG, et al.** The poor quality of information about laparoscopy on the World Wide Web as indexed by popular search engines. Surg Endosc. 2002;16(1):170–2.

11. **Eysenbach G, Powell J, Kuss O, et al.** Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA. 2002;287(20):2691–700.
12. **McLeod S**. The quality of medical information on the internet: a new public health concern. Arch Ophthalmol. 1998;116:1663–5.
13. **Benigeri M, Pluye P.** Shortcomings of health information on the internet. Health Promot Int. 2003;18(4):381–6.
14. **Theodosiou L, Green J.** Emerging challenges in using health information from the internet. Adv Psychiatr Treat. 2003;9:387–96.
15. **Robinson T, Patrick K, Eng T, et al.** An evidence-based approach to enteractive health communication. J Am Med Assoc. 1998;280(14): 1264–9.
16. **Ademiluyi G, Rees CE, Sheard CE.** Evaluating the reliability and validity of three tools to assess the quality of health information on the Internet. Patient Educ Couns. 2003;50(2):151–5.
17. **Ekman A, Hall P, Litton JE.** Can we trust cancer information on the Internet?—a comparison of interactive cancer risk sites. Cancer Causes Control. 2005;16(6):765–72.
18. **Freedman AN, Seminara D, Gail MH, et al.** Cancer risk prediction models: a workshop on development, evaluation, and application. J Natl Cancer Inst. 2005;97(10):715–23.
19. **Berkey CS, Frazier AL, Gardner JD, et al.** Adolescence and breast carcinoma risk. Cancer. 1999;85(11):2400–9.
20. **Sweeney C, Blair CK, Anderson KE, et al.** Risk factors for breast cancer in elderly women. Am J Epidemiol. 2004;160(9):868–75.
21. **Lagiou A, Lagiou P, Vassilarou DS, et al.** Comparison of age at first full-term pregnancy between women with breast cancer and women with benign breast diseases. Int J Cancer. 2003;107(5):817–21.
22. **Helewa M, Levesque P, Provencher D, et al.** Breast cancer, pregnancy, and breastfeeding. J Obstet Gynaecol Canada. 2002;24(2):164–80.
23. **Newcomb PA, Egan KM, Titus-Ernstoff L, et al.** Lactation in relation to postmenopausal breast cancer. Am J Epidemiol. 1999;150(2):174–82.
24. **Silvera SA, Miller AB, Rohan TE.** Oral contraceptive use and risk of breast cancer among women with a family history of breast cancer: a prospective cohort study. Cancer Causes Control. 2005;16(9):1059–63.
25. **Grabrick DM, Hartmann LC, Cerhan J, et al.** Risk of breast cancer with oral contraceptive use in women with a family history of breast cancer. JAMA. 2000;284(14):1791–8.
26. **Pesch B, Ko Y, Brauch H, et al.** Factors modifying the association between hormone-replacement therapy and breast cancer risk. Eur J Epidemiol. 2005;20(8):699–711.
27. **Hartmann LC, Sellers TA, Frost MH, et al.** Benign breast disease and the risk of breast cancer. N Engl J Med. 2005;353(3):229–37.
28. **Malin AS, Qi D, Shu XO, et al.** Intake of fruits, vegetables and selected micronutrients in relation to the risk of breast cancer. Int J Cancer. 2003;105(3):413–8.
29. **Gandini S, Merzenich H, Robertson C, et al.** Meta-analysis of studies on breast cancer risk and diet: the role of fruit and vegetable consumption and the intake of associated micronutrients. Eur J Cancer. 2000;36(5):636–46.
30. **Rubinstein WS.** Hereditary breast cancer in Jews. Familial Cancer. 2004;3(3–4):249–57.
31. **Terry KL, Willett WC, Rich-Edwards JW, et al.** Menstrual cycle characteristics and incidence of premenopausal breast cancer. Cancer Epidemiol Biomark Prev. 2005;14(6):1509–13.
32. **Terry MB, Zhang FF, Kabat G, et al.** Lifetime alcohol intake and breast cancer risk. Ann Epidemiol. 2006;16(3):230–40.
33. National Cancer Institute. SEER cancer statistics review: 1973–1992. Maryland: National Cancer Institute; 1995.
34. American Cancer Society. Statistics for 2007. http://www.cancer.org/docroot/STT/stt_0.asp. Cited May 29, 2007.
35. **Calle EE, Rodriguez C, Walker-Thurmond K, et al.** Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med. 2003;348(17):1625–38.
36. **Ross JA, Parker E, Blair CK, et al.** Body mass index and risk of leukemia in older women. Cancer Epidemiol Biomark Prev. 2004;13(11 Pt 1):1810–3.
37. **Gillois P, Colombet I, Dreau H, et al.** A critical appraisal of the use of Internet for calculating cardiovascular risk. Proc AMIA Symp. 1999; 775–9.
38. **Chen X, Siu LL.** Impact of the media and the internet on oncology: survey of cancer patients and oncologists in Canada. J Clin Oncol. 2001;19(23):4291–7.
39. HarrisInteractive. 4-County survey finds most cyberchondriacs believe online health care information is trustworthy, easy to find and understand. http://www.harrisinteractive.com/news/allnewsbydate.asp? NewsID=464; 2002.
40. **Fox S, Rainie L.** Vital decisions. How Internet users decide what information to trust when they or their loved ones are sick. Pew Internet and American Life Project; May 22, 2002.
41. NCI. Breast cancer risk assessment tool. Bethesda, MD: National Cancer Institute Office of Communications; 1998.

APPENDIX 1: RISK FACTORS OF THE HYPOTHETICAL HIGH AND LOW BREAST CANCER RISK WOMEN

The high-risk patient is a 55-year-old, 5′2″, 150-lb Jewish woman who smokes. Her mother had breast and ovarian cancer, and two sisters had breast cancer, one of whom also had ovarian cancer. She has no personal history of cancer and no other family history of cancer. She has no significant medical history except for benign breast disease and two negative breast biopsies. She had no pregnancies, no live births, began menstruating at age 11, reached menopause at age 53, and did not have a hysterectomy. She took oral contraceptives for 10 years and stopped 10 years ago. She did not take hormone replacement therapy. She does not have annual mammograms or conduct regular self breast exams. She has poor nutrition and exercise habits (the worst option in response to questions assessing these two risk factors was entered on each site).

The low-risk patient is a 55-year-old, 5′5″, 115-lb African-American woman who does not smoke. She has no family history of any cancer and no personal history of cancer. She has had no biopsies and has not been diagnosed with benign breast disease and no other significant medical history. She has had 3 pregnancies, which all resulted in live births, breast fed for a total of 2 years, and had her first child at age 24. Menstruation began at age 14; she reached menopause at age 47 and has not had a hysterectomy. She never took oral contraceptives or hormone replacement therapy. She has annual mammograms and conducts regular self breast exams. She has good nutrition and exercise habits (the best option in response to questions assessing these two risk factors was entered on each site).

Other risk factor questions were either answered as "unknown/not sure" or were answered randomly and recorded.

APPENDIX 2: RISK FACTORS FOR SAMPLE SUBJECTS PRESENTED IN TABLE 4

Risk factor	Subject 16	Subject 1	Subject 37	Subject 31
First period before age 12?	No*	No*	Yes†	No*
First child after age 30?	No*	No*	No*	No*
Childless?	Yes*	Yes*	Yes*	No†
Mother had/has breast cancer?	No*	Yes†	No*	No*
Sisters had/has breast cancer?	No†	No†	No†	No†
Daughters had/has breast cancer?	No†	No†	No†	No†
Ever had breast biopsy?	No†	No†	No†	Yes*
Doctor ever tell you biopsy showed premalignant or precancerous condition?	No†	No†	No†	No†
Doctor ever tell you that one of your biopsies showed early cancer that had not spread yet?	No†	No†	No†	No†
Have you ever used birth control pills?	No*	No*	No*	No*
Ashkenazi Jewish?	No†	No†	No†	No†

**Status of this risk factor is risk increasing (e.g., is Ashkenazi Jewish).* †*Status of this risk factor is risk decreasing (e.g., is not Ashkenazi Jewish).*