Cancer Epidemiol. Author manuscript; available in PMC 2014 December 01.

Published in final edited form as:
Cancer Epidemiol. 2013 December ; 37(6): . doi:10.1016/j.canep.2013.09.004.

Skin Self-Examinations and Visual Identification of Atypical Nevi: Comparing Individual and Crowdsourcing Approaches

Andy J. King, Ph.D.[1], **Robert Gehl, Ph.D.**[2], **Douglas Grossman, M.D., Ph.D.**[3], and **Jakob D. Jensen, Ph.D.**[2,4]

[1]College of Media & Communication, Texas Tech University

[2]Department of Communication, University of Utah

[3]Department of Dermatology and Department of Oncological Sciences, University of Utah

[4]Department of Health Promotion & Education, University of Utah

Abstract

Purpose—Skin self-examination (SSE) is one method for identifying atypical nevi among members of the general public. Unfortunately, past research has shown that SSE has low sensitivity in detecting atypical nevi. The current study investigates whether crowdsourcing (collective effort) can improve SSE identification accuracy. Collective effort is potentially useful for improving people's visual identification of atypical nevi during SSE because, even when a single person has low reliability at a task, the pattern of the group can overcome the limitations of each individual.

Methods—Adults (N = 500) were recruited from a shopping mall in the Midwest. Participants viewed educational pamphlets about SSE and then completed a mole identification task. For the task, participants were asked to circle mole images that appeared atypical. Forty nevi images were provided; nine of the images were of nevi that were later diagnosed as melanoma.

Results—Consistent with past research, individual effort exhibited modest sensitivity (.58) for identifying atypical nevi in the mole identification task. As predicted, collective effort overcame the limitations of individual effort. Specifically, a 19% collective effort identification threshold exhibited superior sensitivity (.90).

Conclusions—The results of the current study suggest that limitations of SSE can be countered by collective effort, a finding that supports the pursuit of interventions promoting early melanoma detection that contain crowdsourcing visual identification components.

Corresponding Author: Andy J. King, Ph.D., Assistant Professor, College of Media & Communication, Texas Tech University, Box 43082, Lubbock, TX 79409-3082, andy.king@ttu.edu, 806-834-3535.
Co-Authors:
Robert Gehl, Ph.D., Assistant Professor, Department of Communication, University of Utah, 255 South Central Campus Drive, Salt Lake City, UT 84112, robert.gehl@utah.edu
Douglas Grossman, M.D., Ph.D., Associate Professor, Departments of Oncological Sciences & Dermatology, University of Utah/ Huntsman Cancer Institute, 2000 Circle of Hope, Salt Lake City, UT 84112, doug.grossman@hci.utah.edu
Jakob D. Jensen, Assistant Professor, Department of Communication, University of Utah, Departments of Communication & Health Promotion/Education, 255 South Central Campus Drive, Room, Salt Lake City, UT 84112, jakob.jensen@utah.edu

Keywords

crowdsourcing; prevention; detection; melanoma; skin self-examination; visual identification; screening

Detecting all types of skin cancer is an important public health goal. Melanoma, the most deadly type of skin cancer, is often the focus of public health interventions and screening efforts as it becomes more common in the general population [1]. The five-year survival rate for distant stage melanoma is only 15% with approximately one person dying of this illness every 61 minutes in the United States. Five-year survival rates for melanoma improve dramatically if the cancer is caught before it advances to a distant stage. The survival rate is 98% if the cancer has not spread to lymph nodes and 61% if at the regional stage [1].

Atypical nevi can be identified as melanoma early through routine clinical examinations by a dermatologist [2]. Yet, mass screening by dermatologists is neither cost-efficient nor feasible; therefore, routine clinical examination is only recommended for high-risk individuals or those with numerous nevi [3]. One of the primary non-clinical methods for initial identification of atypical nevi is skin self-examination (SSE), which is a patient-initiated behavior designed to identify atypical nevi on the skin, often used by patients between clinical visits [3]. Unfortunately, SSE has low sensitivity for detecting atypical moles [3–6], is only marginally improved by existing educational techniques [7], and is rarely practiced or effective at directing people to clinics [6,8]. As a result, SSE is not strongly recommended by most public health organizations [9], but using strategies that increase the accuracy of SSE and related behavior change could be beneficial to members of the public at risk. This paper focuses on trying to improve the accuracy of judgments resulting from people engaging in SSE.

SSE is typically a solitary and inconsistently effective activity; however, evidence suggests that self-examination quality increases as family members assist with the task. Interventions aimed at improving patient self-efficacy for SSE found that people in relationships with motivated partners who perceived being in a high-quality relationship, showed greatest improvements in self-efficacy [10]. Improving self-efficacy is important for increasing quality of skin self-examination, but improving people's ability to identify atypical nevi specifically is a different challenge. Improving self-efficacy increases the likelihood that a behavior will be engaged, but it does not guarantee improvement in necessary skills (e.g., visual identification of a nevus as being atypical).

Issues in judging atypical nevi via SSE could be addressed by collective effort problem solving. Collective effort problem solving, also known as crowdsourcing, uses the intelligence of connected or distributed groups to make judgments; for example, citizen science projects ask motivated members of the public to assist in classification and categorization tasks that might otherwise be overwhelming for only formally trained individuals to conduct. Collective effort can be effective even when a single person has low reliability at a task, as the pattern of the group overcomes the limitations of each individual. In the case of skin cancer prevention and skin self-examination, motivation has been suggested as a key moderator in improving relevant outcomes [10], and crowdsourcing approaches to decision making take advantage of collective motivation, as participation is voluntary, to engage in classification tasks.

Crowdsourcing has been explored as a strategy in various public health and medical domains, although there are limited extant examples directly relevant to cancer prevention. For example, by using crowds of people, typically linked up via telecommunications networks and social media, disaster relief efforts can be better coordinated because affected people can report problems easily to centralized relief organizations [13]. Researchers can compare "real world" uses of prescription medication to clinical trials and further refine knowledge of treatment efficacy because patients can report their experiences with medicine directly to doctors and pharmaceutical companies [14]. Patients can engage in self-diagnosis and support one another in online forums [15]. Finally, the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) can be further refined by opening up medical documentation to large crowds who can propose connections between sub-concepts [16].

Collective effort strategies may not always be optimal, but these strategies may be viable when individual effort is underperforming and/or inefficient and collective effort improves accuracy and/or efficiency. For example, NASA successfully employed collective effort by recruiting thousands of lay-users to comb through millions of planetary images distinguishing craters from shadows [17,18]. Individual NASA employees could have analyzed each image, but it would have taken years to complete the task and individual evaluator fatigue or errors could have potentially undermined results. Given the inconsistency of SSE results in identifying atypical nevi, crowdsourcing the task of making judgments about moles people are suspicious about has the potential to improve one's quality of decision-making in moving from identification to action (e.g., talking to a dermatologist). By having a collective of individuals, trained through online modules similar to the crowdsourcing projects discussed, collective effort may better identify nevi that should be examined in a clinical setting.

To explore this potential of crowdsourcing as a component of a more comprehensive skin cancer prevention effort, the current study evaluates whether collective effort outperforms individual effort in the context of visual identification of atypical nevi. Specifically—is considering collective effort a more efficient way of trying to identify atypical nevi among mass audiences? Existing research has revealed that SSE has limited accuracy when carried out by an individual [3]. Currently unknown is whether the accuracy limitations of atypical nevi recognition during SSE can be countered by collective effort. That is, if a person could use "the crowd" to determine whether a mole was atypical, would the audience exhibit superior sensitivity and specificity?

Methods

Procedure

Data for the present study were collected as part of a larger project testing the effectiveness of using pamphlets to improve people's intention to practice SSE and their ability to identify atypical nevi. Participants over the age of 18 were recruited by the research team from a shopping mall located in a mid-sized Midwestern city. The mall was modest in size and targeted a traditional college town population, meaning there was diversity in store types to appeal to a variety of niche audiences on dimensions such as age and socioeconomic status. Large signs informed mall shoppers about the study, including the incentive ($15 gift card). In total, 500 individuals were recruited into the study. Based on observations of a research team member charged with the task of lining up participants and monitoring mall traffic, approximately 1 in 25 people stopped to participate in the study.

Participants first completed a pretest survey, then received printed materials to examine, and finally completed a posttest survey. The printed materials were pamphlets that explained standard nevi examination techniques used in SSE contexts (e.g., ABCDs, Ugly Duckling Sign). In the posttest, participants completed a mole identification task. Forty mole images were used in the task. Mole images were obtained from the company MoleMap (http://www.molemap.co.nz/), as well as Internet sources such as a National Cancer Institute

NIH-PA Author Manuscript
NIH-PA Author Manuscript

database (http://visualsonline.cancer.gov/). Of the forty images, 31 were of moles not diagnosed as melanoma and nine images were moles diagnosed as melanoma. Participants were asked to circle all mole images that appeared to be atypical nevi. The language of the task did not use the term atypical nevi, but rather asked participants to circle mole images they believed to be atypical. Following completion of the mole identification task, participants were thanked for their participation and provided compensation. A university institutional research board approved the research protocol, questions, and materials.

Participants

Average participant age was 36 years old (M = 36.14, SD = 14.22), with ages between 18 and 80. Participants were more likely to be female (57.2%, n = 286), white (73.8%, n = 369), and at least a high school graduate (92.8%, n = 461). Most participants were either single (38.4%, n = 192) or married (41.6%, n = 208). Skin cancer risk was measured using the brief skin cancer risk assessment test (BRAT) [19]. BRAT estimates classified just over half of participants as low risk (54.8%, n = 274), a third at moderate risk (34.8%, n = 174), and a small proportion at high risk (10.2%, n = 51). Three participants failed to complete the mole identification task as instructed, while other participants may have skipped other sections of the survey. There was no pattern in the missing data, but when data were missing listwise deletion was used to eliminate individuals with incomplete information. Information about the sample size for various analyses is reported in text and tables.

Statistical Methods

Participants completed the mole identification task by circling the mole images they believed to be atypical. Melanoma images classified as atypical were true positives (TP). Non-melanoma mole images classified as typical were true negatives (TN). Non-melanoma moles classified as atypical were considered false positives (FP), and false negatives (FN) were mole images of melanomas that were not identified as atypical.

After classifying individual participant effort into units of true/false positives/negatives, four proportional scores were calculated: sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) [20–22]. These four proportions represent individual effort scores for the identification of mole images classified correctly as atypical/typical. Individual effort refers to the mean performance of individuals in the sample.

Following the calculation of individual effort scores, collective effort scores were calculated. Collective effort considers the pattern across all users (e.g., 35% of participants think a particular mole is atypical). For instance, imagine that 100 people were asked to examine five mole images (one of which was a clinically diagnosed melanoma). In this hypothetical situation, most people (65%) incorrectly classify the melanoma as typical (i.e., a false negative), a response that yields a low average individual effort score. Collective effort ignores the limitations of individual effort and considers the pattern of the group. In this case, the pattern of the group is revealing as 35% of people did score the melanoma as atypical (i.e., a true positive) which is a relatively high number compared to the rest of the moles. Thus, this hypothetical scenario illustrates the potential of collective effort to successfully overcome the limitations of individual ability in SSE.

The pattern across all users is useful information, but research can aid users of collective effort data by identifying meaningful thresholds or cut-off points. In the current data (see Table 1), the majority of people viewed most moles as typical, non-melanoma images (Median % scored atypical = 16.25%). In fact, for 60% of the non-melanoma images fewer than 19% of raters were concerned. Given that and general statistical advice for analyzing classification tasks [38], we examined 19% as a threshold for identifying melanoma images.

For comparative purposes, we also examined 65% as a collective effort threshold. That threshold was selected as the majority of the moles (85%) were identified as atypical by fewer than 65% of people.

Results

The goal of the current study was to determine if collective effort was more effective than individual effort at distinguishing between nevi types (see Table 2). In the current study, individual effort correctly categorized 58% of melanoma images as being “atypical” (sensitivity) and correctly categorized 81% of non-melanoma images as being “typical” (specificity). For mole images categorized by participants as “atypical,” 49% were images of clinically diagnosed melanoma cases (PPV). For mole images categorized as typical, 87% were non-melanoma images (NPV). Collective effort scores were calculated using the methods and thresholds specified in the statistical methods section. Using the 19% threshold, collective effort correctly classified 90% of melanoma images and 72% of non-melanoma images. For the mole images categorized by participants as being atypical, 50% were melanoma images. For mole images identified as “typical,” 96% were non-melanoma images.

Using the 65% threshold, collective effort correctly categorized 67% of melanoma images and 100% of non-melanoma images. For mole images identified as being “atypical,” 100% of those were melanoma images. For mole images categorized as “typical,” 91% were non-melanoma images.

Thus, the 19% collective effort threshold appeared to yield optimal sensitivity compared to other strategies. To determine if the observed differences between individual and collective effort were statistically significant, z-tests comparing proportions were calculated. There were substantive and statistically significant differences between individual effort sensitivity and collective effort sensitivity at the 19% threshold, $z = 12.34$, $p < .001$, as well as differences between a 19% threshold and 65% threshold, $z = 9.18$, $p < .001$, and differences between individual estimate sensitivity and a 65% collective effort threshold, $z = 2.94$, $p = .002$. Proportions for all other dimensions—specificity, PPV, and NPV—were all significantly different as well with one exception. There was no difference between the PPV for individual effort and the 19% collective threshold.

In any crowdsourcing project, it can be difficult to know the demographics of online participants, and thus it can be difficult to break down task performance across subgroups. However, in the case of the present study, it is possible to examine how segments of sample subgroups perform on the task. Considering three variables—skin cancer risk, sex, and education—subgroup analyses were performed. These results appear in Tables 3 to 5. Taken together, the subgroups showed few deviations from the total sample results with one notable exception. Collective effort, using a 19% agreement threshold, resulted in a higher sensitivity score for people with a high school degree or less (.89) compared to those with more than a high school education (.67; $z = 6.07$, $p < .001$).

Discussion

In summary, though specificity was lower (i.e., a higher false-positive rate), all other benchmarks found crowdsourcing to be more effective in identifying atypical nevi than individual effort. Results for individual effort in this study are consistent with those found in past evaluations of SSE accuracy [9].

Notably, a 19% collective effort threshold was considerably more sensitive than individual effort at detecting melanoma images—for the entire sample as well as the lower education subgroup. For visual identification of atypical nevi, sensitivity might be considered a more relevant component as a false negative would result in missed melanoma diagnosis, whereas a false positive presumably leads the individual to schedule a consultation with a dermatologist for a clinical examination. Given the goal of skin self-examination is to begin a dialogue and potential treatment with a physician, this false positive does not necessarily mean invasive clinical treatment of any sort.

This approach to identifying atypical nevi will likely only meaningfully improve early identification of melanoma if the identification results in a person deciding to speak with their physician or dermatologist about the nevus. There is always the possibility that the "crowd" could incorrectly classify an image of a mole. As noted earlier, that is why a collective effort approach—based on results of this study—could be incorporated into a larger intervention effort intended to improve behavioral outcomes of skin self-examination, such as increasing intentions to visit doctor about atypical nevi and to talk with doctors about clinical examinations at annual health screenings.

From a broader perspective, the results of the current study suggest that some limitations of individual visual identification of atypical nevi during SSE can be countered by a crowdsourcing approach. Thus, the next step in this research is examination of the potential implementation and evaluation of collective effort interventions. For example, collective effort could be implemented via a web-based interface that allowed individuals to post images of their moles to receive crowdsourced feedback. Individuals could post their own images or post images with the assistance of a portable camera system. The latter could be introduced to underserved populations via portable healthcare units, such as those utilized by public health nurses [23]. For example, rural populations are less likely to have access to dermatologists or healthcare professionals trained in dermoscopy, a service gap that increases melanoma mortality rates in this population [24]. Yet, rural populations increasingly have access to portable healthcare units. Such a system could provide users with a more reliable means for managing their own care, encourage innovate telemedicine efforts, and nudge users toward action [25]. The potential implications of false negatives, however, cannot be overlooked—what happens if the "crowd" fails an individual user? Shortcomings and successes of other crowdsourced projects provide some potential guidance in dealing with such issues.

The initial suggestions that crowdsourcing and collective effort would always be successful are proving untrue [26]. For example, Lichtenthaler and Ernst found that highly regarded collective effort projects have not lived up to the high expectations of much of the early research and projections [27]. Chanal and Caron-Fasan produced an evaluation of a failed collective effort site, CrowdSpirit, a project created to allow crowds to design, test, and produce products; they found that early conceptual problems undermined this project [28].

Past failures have identified several obstacles that can undermine the success of collective effort interventions that should be considered in the context of skin cancer. For example, collective effort often fails to achieve meaningful results because of problems related to scale. Scale refers to several factors including the size of the sample material the crowd will work with (e.g., the number of mole images), the size of the crowd, and the amount of material produced by the crowd (e.g., quantity of feedback). Raymond argued that a successful collective effort project needs a seedbed of material for users to work with [29]. The Mars Clickworkers project [30], ReCaptcha [31], and The GoldCorp Challenge [11] all started with a large database of objects for users to analyze. This suggestion indicates a need for any intervention related to skin cancer to develop a large database of potential moles to identify, as well as determine what type of feedback system is most useful for end users.

Scale also refers to the size of the crowd. Scholars of collective effort projects have frequently noted that a bigger user base translates to greater success [33], which is conceptually in line with the basic presuppositions of collective effort. With more participation by diverse and motivated people, better solutions, ideas, and products emerge. The effects of crowd size can be seen within existing and well-regarded crowdsourced projects: for example, Wikipedia entries with many participants tend to suffer vandalism for far shorter periods than less-trafficked articles [34].

Finally, even with a large initial seedbed and an adequate crowd, there is the issue of vetting results. Scale can become a challenge after the crowd weighs in: what is to be done with the mass of data and ideas the crowd produces [33]? Often this judgment falls on the sponsor of the project. For example, in disaster relief, a collective effort strategy struggled because verification of geographic data and the elimination of fraud required too much work on the part of the relief organizations [13]. In design and idea production contests, someone has to act as judge, and with more entries, this work is harder [12].

Future work needs to address challenges identified by past SSE research. For example, researchers have noted that SSE is undermined by a failure to completely scan all parts of the body [35–36]. Users might be willing to upload a photo of a single mole, but that could ultimately prove suboptimal if they are failing to monitor other parts of their body. As such, any intervention still requires a multilevel approach that addresses education, behavior change, and actual behavioral ability/accuracy. Grossman and colleagues have spent the last decade studying the relationship between various forms of mole imagery and melanoma screening accuracy [37]. Their research to date suggests that accuracy is improved by the addition of regional photographs so that individual lesions can be viewed in the context of other lesions. For example, if a patient has a suspicious lesion on his/her arm, then it is useful to have a photograph of the lesion (up close) and the arm (i.e., the region). Regional photographs provide context and allow for the possible identification of melanoma arising *de novo* (i.e., from normal-appearing skin).

Despite these challenges, the potential of collective effort is alluring, especially for researchers and practitioners interested in improving early melanoma detection. SSE is a behavioral and visual identification task, and some of the most effective collective effort interventions have focused on the processing of visual information [17]. In fact, there is evidence that individuals are already utilizing quasi-collective effort approaches to identify atypical moles. A search of the Internet reveals numerous people informally posting mole images to health forums and asking for feedback. The continued academic study of approaches such as crowdsourcing could assist in creating formal, credible opportunities for the public to engage in melanoma detection and, as part of a more comprehensive intervention, could improve the effectiveness of SSE for targeted populations.

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA: A Cancer Journal for Clinicians. 2012; 62:10–29. [PubMed: 22237781]
2. Goulart JM, Malvehy J, Puig S, Martin G, Marghoob AA. Dermoscopy in SSE: A useful tool for patients. Archives of Dermatology. 2011; 147:53–58. [PubMed: 21242394]
3. Goodson AG, Grossman D. Strategies for early melanoma detection: Approaches to the patient with nevi. Journal of the American Academy of Dermatology. 2009; 60:719–735. [PubMed: 19389517]

4. Buettner PG, Garbe C. Agreement between self-assessment of melanocytic nevi by patients and dermatologic examination. American Journal of Epidemiology. 2000; 151:72–77. [PubMed: 10625176]
5. Carli P, De Giorgi V, Nardini P, et al. Melanoma detection rate and concordance between self-skin examination and clinical evaluation in patients attending a pigmented lesion clinic in Italy. British Journal of Dermatology. 2002; 146:261–266. [PubMed: 11903237]
6. Hamidi R, Peng D, Cockburn M. Efficacy of SSE for the early detection of melanoma. International Journal of Dermatology. 2010; 49:126–134. [PubMed: 20465635]
7. Grob JJ, Bonerandi JJ. The 'ugly duckling' sign: identification of the common characteristics of nevi in an individual as a basis for melanoma screening. Archives of Dermatology. 1998; 134:103–104. [PubMed: 9449921]
8. Miller DR, Geller AC, Wyatt SW, Halpern A, Howell JB, Cockerell C, et al. Melanoma awareness and self-examination practices: Results of a United States Survey. Journal of the American Academy of Dermatology. 1996; 34:962–970. [PubMed: 8647989]
9. U S. Preventive Services Task Force. Screening for skin cancer: U.S. Preventive Services Task Force Recommendation Statement. Annals of Internal Medicine. 2009; 150:188–193. [PubMed: 19189908]
10. Robinson JK, Stapleton J, Turrisi R. Relationship and partner variables increase self-efficacy of performing skin self-examination. Journal of the American Academy of Dermatology. 2008; 58:755–762. [PubMed: 18258332]
13. Gao H, Barbier G, Goolsby R. Harnessing the crowdsourcing power of social media for disaster relief. IEEE Intelligent Systems. 2011; 26(3):10–14.
14. Armstrong AW, Harskamp CT, Cheeney S, Schupp CW. Crowdsourcing for research data collection in rosacea. Dermatology Online Journal. 2012; 18(3):15. http://dermatology.cdlib.org/1803/08_ltr/15_12-00035/article.html. [PubMed: 22483526]
15. Leslie S. Crowdsourcing Your Diagnosis. MLA News. 2012; 52(2):8.
16. Simpson CR, Anandan C, Fischbacher C, Lefevre K, Sheikh A. Will systematized nomenclature of medicine-clinical terms improve our understanding of the disease burden posed by allergic disorders? Clinical & Experimental Allergy. 2007; 37:1586–1593. [PubMed: 17883425]
17. Gehl RW. The archive and the processor: the internal logic of web 2.0. New Media & Society. 2011; 13(8):1228–1244.
18. NASA Open Government Plan. Public participation in aeronautics research and space exploration. 2010. Retrieved September 5, 2012 from http://www.nasa.gov/open/plan/peo.html
19. Glanz K, Schoenfeld E, Weinstock MA, Layi G, Kidd J, Shigaki DM. Development and reliability of a brief skin cancer risk assessment tool. Cancer Detection & Prevention. 2003; 27:311–315. [PubMed: 12893080]
20. McNeil BJ, Keller E, Adelstein SJ. Primer on certain elements of medical decision making. N Engl J Med. 1975; 293:211–5. [PubMed: 806804]
21. Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures. Principles and applications. Ann Intern Med. 1981; 94(4 Pt 2):557–92. [PubMed: 6452080]
22. Smith JE, Winkler RL, Fryback DG. The first positive: Compute positive predictive value at the extremes. Ann Intern Med. 2000; 132:804–809. [PubMed: 10819704]
23. Hutchison, L.; Hawes, C.; Williams, L. Access to quality health services in rural areas--long-term care: A literature review. In: Gamm, L.; Hutchison, L., editors. Rural Healthy People 2010: A companion document to Healthy People 2010. Vol. 3. College Station, TX: The Texas A&M University System Health Science Center, School of Rural Public Health, Southwest Rural Health Research Center; 2005. p. 1-28.
24. Aneja S, Aneja S, Bordeaux JS. Association of increased dermatologist density with lower melanoma mortality. Archives of Dermatology. 2012; 148:174–178. [PubMed: 22351816]
25. Hesse BW, Ahern DK, Woods SS. Nudging best practice: the HITECH act and behavioral medicine. Translational Behavioral Medicine. 2011; 1:175–181. [PubMed: 23662166]
26. Marjanovic S, Fry C, Chataway J. Crowdsourcing based business models: in search of evidence for innovation 2.0. Science and Public Policy. 2012; 39(3):318–332.

27. Lichtenthaler U, Ernst H. Innovation intermediaries: why internet marketplaces for technology have not yet met the expectations. Creativity and Innovation Management. 2008; 17(1):14–25.
28. Chanal V, Caron-Fasan ML. The difficulties involved in developing business models open to innovation communities: the case of a crowdsourcing platform. M@n@gement. 2010; 4(13):318–340.
29. Raymond, ES. [Accessed November 20, 2012] The cathedral and the bazaar. First Monday. 2005. http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1472/1387
30. Benkler, Y. The wealth of networks: how social production transforms markets and freedom. New Haven, Conn: Yale University Press; 2006.
31. von Ahn, L. 46th ACM/IEEE Design Automation Conference, 2009. 2009. Human Computation; p. 418-419.
32. Malin MC, Edgett KC. Mars global surveyor mars orbiter camera: interplanetary cruise through primary mission. Journal of Geophysical Research. 2001; 106(E10):23429–23570.
33. Poetz MK, Schreier M. The value of crowdsourcing: can users really compete with professionals in generating new product ideas? Journal of Product Innovation Management. 2012; 29(2):245–256.
34. Lih, A. Wikipedia revolution, the: how a bunch of nobodies created the world's greatest encyclopedia. New York, New York: Hyperion; 2009.
35. Arnold MR, DeJong W. SSE practices in a convenience sample of U.S. university students. Preventive Medicine. 2005; 40:268–273. [PubMed: 15533538]
36. Jensen JD, Moriarty CM. Psychosocial correlates of SSEs. Journal of American College Health. 2008; 56:701–705. [PubMed: 18477528]
37. Goodson AG, Florell SR, Hyde M, Bowen GM, Grossman D. Comparative analysis of total body and dermatoscopic photographic monitoring of nevi in similar patient populations at risk for cutaneous melanoma. Dermatologic Surgery. 2010; 36:1087–1098. [PubMed: 20653722]
38. Pepe, MS. The statistical evaluation of medical tests for classification and prediction. Oxford; New York: Oxford University Press; 2004.

Table 1

Collective Effort Data for Mole Identification Task

Mole	% scored atypical
Mole #1 (Non-melanoma)	1.40%
Mole #2 (Non-melanoma)	2.90%
Mole #3 (Non-melanoma)	3.50%
Mole #4 (Non-melanoma)	4.80%
Mole #5 (Non-melanoma)	5.20%
Mole #6 (Non-melanoma)	5.20%
Mole #7 (Non-melanoma)	6.10%
Mole #8 (Non-melanoma)	6.40%
Mole #9 (Non-melanoma)	6.60%
Mole #10 (Non-melanoma)	7.40%
Mole #11 (Non-melanoma)	7.60%
Mole #12 (Non-melanoma)	9.50%
Mole #13 (Non-melanoma)	10.30%
Mole #14 (Non-melanoma)	11.20%
Mole #15 (Non-melanoma)	12.20%
Mole #16 (Non-melanoma)	13.00%
Mole #17 (Non-melanoma)	13.70%
Mole #18 (Non-melanoma)	15.50%
Mole #19 (Non-melanoma)	16.10%
Mole #20 (Non-melanoma)	16.40%
Mole #21 (Non-melanoma)	17.80%
Mole #22 (Non-melanoma)	18.30%
Mole #23 (Non-melanoma)	23.10%
Mole #24 (Non-melanoma)	25.20%
Mole #25 (Non-melanoma)	25.40%
Mole #26 (Non-melanoma)	28.50%
Mole #27 (Non-melanoma)	38.80%
Mole #28 (Non-melanoma)	60.70%
Mole #29 (Non-melanoma)	60.90%
Mole #30 (Non-melanoma)	61.00%
Mole #31 (Non-melanoma)	64.10%
Mole #32 (Melanoma)	11.00%
Mole #33 (Melanoma)	19.00%
Mole #34 (Melanoma)	20.50%
Mole #35 (Melanoma)	68.60%
Mole #36 (Melanoma)	70.00%
Mole #37 (Melanoma)	74.60%
Mole #38 (Melanoma)	77.00%
Mole #39 (Melanoma)	85.00%

Mole	% scored atypical
Mole #40 (Melanoma)	92.60%

Note. Actual data from the mole identification task study. Moles identified as melanoma were clinically diagnosed as such. The percent of participants that scored a mole as atypical is listed under % scored atypical. Non-melanoma/melanoma mole images are presented here in numerical order based on % scored atypical. Actual mole images were presented to participants in a more random order.

Table 2

Comparing Individual Visual Identification Performance to Collective Effort Performance

	Individual Effort (average)	Collective Effort – 19% Threshold	Collective Effort – 65% Threshold
Sensitivity	.58	.90	.67
Specificity	.81	.72	1.00
PPV	.49	.50	1.00
NPV	.87	.96	.91

Note. *N* = 471. Individual effort is the average ability of a single user to detect a melanoma mole image. A 19% threshold means that moles are only considered atypical if at least 19% of the group deems them to be.

PPV = Positive Predictive Value NPV = Negative Predictive Value

Table 3

Comparing Individual Identification to Collective Identification (Segmented by Risk Level)

	Individual Effort (average)	Collective Effort – 19% Threshold	Collective Effort – 65% Threshold
	Low Risk/Mod-Hi Risk	**Low Risk/Mod-Hi Risk**	**Low Risk/Mod-Hi Risk**
Sensitivity	.58/.58	.89/.89	.56/1.00
Specificity	.80/.81	.68/.71	1.00/1.00
PPV	.48/.50	.44/.47	1.00/1.00
NPV	.87/.87	.96/.96	.89/.91

Note. N = 470 (Low risk n = 260, mod-hi risk n = 210). Individual effort is the average ability of a single user to detect a melanoma mole image. A 19% threshold means that moles are only considered atypical if at least 19% of the group deems them to be.

Mod-Hi Risk = Moderate to High Risk Scores on the BRAT measure

PPV = Positive Predictive Value

NPV = Negative Predictive Value

Table 4

Comparing Individual Identification to Collective Identification (Segmented by Sex)

	Individual Effort (average)	Collective Effort – 19% Threshold	Collective Effort – 65% Threshold
	Male/Female	**Male/Female**	**Male/Female**
Sensitivity	.57/.58	.78/.78	.56/.67
Specificity	.80/.81	.68/.71	.97/1.00
PPV	.48/.50	.41/.44	.83/1.00
NPV	.87/.87	.91/.92	.88/.91

Note. N = 467 (Male n = 198; female n = 269). Individual effort is the average ability of a single user to detect a melanoma mole image. A 19% threshold means that moles are only considered atypical if at least 19% of the group deems them to be.

PPV = Positive Predictive Value

NPV = Negative Predictive Value

Table 5

Comparing Individual Identification to Collective Identification (Segmented by Education)

	Individual Effort (average)	Collective Effort – 19% Threshold	Collective Effort – 65% Threshold
	HS or less/More than HS	**HS or less/More than HS**	**HS or less/More than HS**
Sensitivity	.55/.60	.89/.67	.44/.67
Specificity	.80/.81	.68/.74	1.00/.97
PPV	.47/.51	.44/.43	1.00/.86
NPV	.86/.88	.96/.89	.86/.91

Note. N = 469 (HS or less n = 191, More than HS n = 278). Individual effort is the average ability of a single user to detect a melanoma mole image. A 19% threshold means that moles are only considered atypical if at least 19% of the group deems them to be.

HS or less = High school degree or less

More than HS = Education beyond a high school degree

PPV = Positive Predictive Value

NPV = Negative Predictive Value