

REVIEWS

Crowdsourcing—Harnessing the Masses to Advance Health and Medicine, a Systematic Review

Benjamin L. Ranard, AB¹, Yoonhee P. Ha, MSc¹, Zachary F. Meisel, MD, MPH, MS^{3,4}, David A. Asch, MD, MBA^{2,3,5,6}, Shawndra S. Hill, PhD⁵, Lance B. Becker, MD⁴, Anne K. Seymour, MS⁷, and Raina M. Merchant, MD, MSHP^{2,3,4}

¹Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ²Penn Medicine Center for Innovation, University of Pennsylvania, Philadelphia, PA, USA; ³The Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA; ⁴Department of Emergency Medicine, University of Pennsylvania, Philadelphia, PA, USA; ⁵Philadelphia Veterans Affairs Medical Center, Philadelphia, PA, USA; ⁶The Wharton School, University of Pennsylvania, Philadelphia, PA, USA; ⁷University of Pennsylvania Libraries, University of Pennsylvania, Philadelphia, PA, USA.

OBJECTIVE: Crowdsourcing research allows investigators to engage thousands of people to provide either data or data analysis. However, prior work has not documented the use of crowdsourcing in health and medical research. We sought to systematically review the literature to describe the scope of crowdsourcing in health research and to create a taxonomy to characterize past uses of this methodology for health and medical research. **DATA SOURCES:** PubMed, Embase, and CINAHL through March 2013.

STUDY ELIGIBILITY CRITERIA: Primary peer-reviewed literature that used crowdsourcing for health research.

STUDY APPRAISAL AND SYNTHESIS METHODS: Two authors independently screened studies and abstracted data, including demographics of the crowd engaged and approaches to crowdsourcing.

RESULTS: Twenty-one health-related studies utilizing crowdsourcing met eligibility criteria. Four distinct types of crowdsourcing tasks were identified: problem solving, data processing, surveillance/monitoring, and surveying. These studies collectively engaged a crowd of >136,395 people, yet few studies reported demographics of the crowd. Only one (5 %) reported age, sex, and race statistics, and seven (33 %) reported at least one of these descriptors. Most reports included data on crowdsourcing logistics such as the length of crowdsourcing (n=18, 86 %) and time to complete crowdsourcing task (n=15, 71 %). All articles (n=21, 100 %) reported employing some method for validating or improving the quality of data reported from the crowd.

LIMITATIONS: Gray literature not searched and only a sample of online survey articles included.

CONCLUSIONS AND IMPLICATIONS OF KEY FINDINGS: Utilizing crowdsourcing can improve the quality, cost, and speed of a research project while engaging large segments of the public and creating novel science. Standardized guidelines are needed on crowdsourcing metrics that

should be collected and reported to provide clarity and comparability in methods.

KEY WORDS: crowdsourcing; crowd sourcing; citizen scientist; citizen science; human computing.

J Gen Intern Med 29(1):187–203

DOI: 10.1007/s11606-013-2536-8

© Society of General Internal Medicine 2013

INTRODUCTION

Crowdsourcing is an approach to accomplishing a task by opening up its completion to broad sections of the public. Innovation tournaments, prizes for solving an engineering problem, or paying online participants for categorizing images are examples of crowdsourcing. What ties these approaches together is that the task is outsourced with little restriction on who might participate. Despite the potential of crowdsourcing, little is known about the applications and feasibility of this approach for collecting or analyzing health and medical research data where the stakes are high for data quality and validity.

One of the most celebrated crowdsourcing tasks was the prize established in 1714 by Britain's Parliament in the Longitude Act, offered to anyone who could solve the problem of identifying a ship's longitudinal position.¹ The Audubon Society's Christmas Bird Count began in 1900 and continues to this day as a way for "citizen scientists" to provide data that can be used for studying bird population trends.² However, today the world has 2.3 billion Internet users and 6 billion mobile phone subscriptions,³ providing access that facilitates crowdsourcing to a much greater extent than was available to Britain's Parliament and the Longitude Act. The Galaxy Zoo project (galaxyzoo.org) successfully classified nearly 900,000 galaxies with the help of hundreds of thousands of online volunteers. The simple visual classification was easily performed by humans but not by computers.⁴ Other examples include Whale.fm (whale.fm), which has almost 16,000 whale

Received February 1, 2013

Revised May 6, 2013

Accepted May 20, 2013

Published online July 11, 2013

calls that volunteers are classifying⁵ in order to help process large data sets that have become unmanageable for researchers alone to classify.⁶ The online platform eBird (ebird.org) collected more than 48 million bird observations from well over 35,000 contributors.^{7,8}

While this prior work illustrates the promise of crowdsourcing as a research tool, little is known about the types of questions crowdsourcing is best suited to answer and about the limitations of its use. Health research in particular requires high standards for data collection and processing, tasks traditionally conducted by professionals and not the public. Furthermore, human health research often requires protections for privacy and against physical harm. To better understand the potential of crowdsourcing methods in health research, we conducted a systematic literature review to identify primary peer-reviewed articles focused on health-related research that used crowdsourcing of the public. Our aim was to characterize the types of health research tasks crowdsourcing has been used to address and the approaches used in order to define future opportunities and challenges.

METHODOLOGY

Data Sources and Searches

Definitions. Crowdsourcing was defined as soliciting over the Internet from a group of unselected people, services and data that could not normally be provided solely by automated sensors or computation lacking human input. Crowdsourcing participants had to be actively engaged in the crowdsourcing task and not simply passively have their data mined without their knowledge.

Health research was defined as research that contributes to the World Health Organization's definition of health: "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity."⁹

Systematic Literature Search. A systematic literature search was performed on March 24, 2013, by searching PubMed, Embase, and CINAHL using the following Boolean search string: crowdsourc* OR "crowd source" OR "crowd sourcing" OR "crowd sourced" OR "citizen science" OR "citizen scientist" OR "citizen scientists." Articles underwent a multistage screening process whereby results were pooled and duplicates were removed. Two reviewers (BR and RM) screened abstracts for relevance and then screened full-text articles to confirm eligibility criteria. Two literature-informed database searches were performed to identify additional studies. Authors of published papers meeting eligibility criteria are likely to be experts in their field. As a result, PubMed was searched using the full names of first

authors of all papers meeting eligibility criteria. Two common crowdsourcing platforms used by articles meeting manuscript criteria were Amazon Mechanical Turk (AMT) (mturk.com) and Foldit (fold.it). Therefore, a literature-informed search of all three databases was performed by including the terms "mechanical turk" and "foldit." Reference lists of articles meeting eligibility criteria were reviewed to identify additional articles, as were relevant review articles that were returned by the database search. Project websites that were specifically cited by references were manually searched for relevant publications.

Crowdsourced Literature Search. A crowdsourced search for literature was performed by posting an open call for articles on two free websites: Yahoo! Answers (answers.yahoo.com) and Quora (quora.com). The title used in the question was: "Crowdsourcing: published literature on crowdsourcing in health/medicine?" and the body of the question was: "What scientific research articles in health/medicine have been published that use crowdsourcing in part or in whole to achieve their research objectives?" Results were collected after seven days. Responders consented to allow use of provided references.

Study Selection. Studies were included if they met the following criteria: (1) primary peer-reviewed journal article representing original health research; (2) methodology and results provided; and (3) citizen crowdsourcing used by scientists to obtain at least part of the results. Excluded were studies soliciting opinions only from other experts (i.e., experts collaborating with each other), abstracts, editorials, and wikis that existed simply to create content but not to answer a specific research question with original data. Also excluded were papers that used an Internet survey that did not contain any of the original crowdsourcing keywords from our Boolean search string in the title or abstract of the paper. Behavioral research has been conducted via the Internet through Internet surveys for over 15 years,¹⁰ and therefore this is not a particularly novel research method. We included the surveys returned by our database search to provide a few examples from the field.

Data Extraction and Quality Assessment. Two reviewers independently extracted data (BR and YH). The following data were extracted from articles (including participant data items recommended by the Cochrane Handbook checklist¹¹): study background information (title, author, publication year, research field, methodology type, study objective, study outcome), demographic and other characteristics of the crowd (size of the crowd, age, gender, racial/ethnic background, geographic location, occupation, education, relationship to the research problem, referral source, stated conflict of interest, motivation), and the

logistics of the crowdsourcing [length of time crowdsourcing was conducted, use of a web platform and/or a mobile platform, use of individuals compared to teams, intracrowd sharing techniques (such as team wiki or forum), data collected or processed, complexity of the task, time given to do the task, advertisement of project, skill set required, monetary incentives offered, and data validation techniques]. Additional extracted data included the viewer-to-participant ratio, reflecting how many people saw the task (or website with the task) to how many people completed the task.

Data Synthesis and Analysis. Summary statistics were used to describe the number of studies reviewed and to characterize the data extracted from these studies.

This study was approved by the University of Pennsylvania Institutional Review Board.

RESULTS

Systematic Literature Search

There were 231 unique articles identified from the initial database and crowdsourced search and 209 articles identified in the literature-informed database search. Of the 440 articles, 76 articles underwent a full-text review. Sixteen of these articles met eligibility criteria. Five additional articles were identified from reference lists of eligible articles (Fig. 1). The final article cohort consisted of 21 unique health research-related primary peer-reviewed publications that used crowdsourcing as a methodology¹²⁻³² (Table 1).

Study Characteristics

We identified four types of research tasks that articles in our cohort employed crowdsourcing to accomplish: problem solving, data processing, surveillance/monitoring, and surveying (Appendix Fig. 2 and Appendix Table 4).

Problem Solving

Seven of 21 articles (33 %) from the final cohort employed crowdsourcing for problem solving. Six of these used Foldit, an online game that allows users to manipulate the three-dimensional structures of proteins in order to find the most likely tertiary structure.^{12,13,17-19,21} Also described was the online game Phylo, where users moved colored blocks representing different nucleotides of a gene promoter sequence around on screen in order to make the most parsimonious phylogenetic tree.²³

Data Processing

Crowdsourcing was used to provide data processing in 7 of 21 articles (33 %). Three papers used AMT, a service from Amazon.com that allows individuals to create accounts and sign up to do an online task in exchange for payment. AMT Knowledge Workers (KWs) classified polyps in computer tomography (CT) colonography images²⁸ and then were asked questions to gauge how to optimize presentation of the polyps.²⁷ AMT was also used to annotate public webcam images to determine how the addition of a bike lane changed the mode of transportation observed in the images.³¹ Three manuscripts reported on two independent games that used crowdsourcing to either identify red blood cells (RBCs) infected with^{25,26} or thick blood smears containing²⁴ malaria parasites (*Plasmodium falciparum*). The final paper in the data processing category attempted to use the crowd to update the literature and evidence covered by a systematic review.¹⁶

Surveillance/Monitoring

Surveillance/monitoring was employed in 3 of 21 (14 %) studies. One of the papers used AMT to ask users about their malaria symptoms in order to assess malaria prevalence in India.²⁰ Another used a mobile phone application that allowed users to report potential flu-like symptoms along with GPS coordinates and other details, which enabled researchers to chart incidence of flu symptoms that matched relatively well with Centers for Disease Control and Prevention data.¹⁴ The last paper created a map of automated external defibrillators (AEDs) by having users of a mobile phone application locate and take pictures of AEDs.³²

Surveying

Crowdsourcing to conduct surveys was reported in 4 of 21 (19 %) papers. All four used AMT to administer surveys. One study allowed surveyors to include a more diverse population than the typical university research subject pool while maintaining reliability (as measured by Cronbach's coefficient alpha).¹⁵ A second study used AMT to study the human decision-making process and showed that AMT replicated previous laboratory findings, indicating that it may be a good platform for future decision-making studies.²² A third study used AMT to administer surveys about health promotional materials and solicit feedback.²⁹ The last study used KWs as subjects for cognitive behavioral tests administered through AMT.³⁰

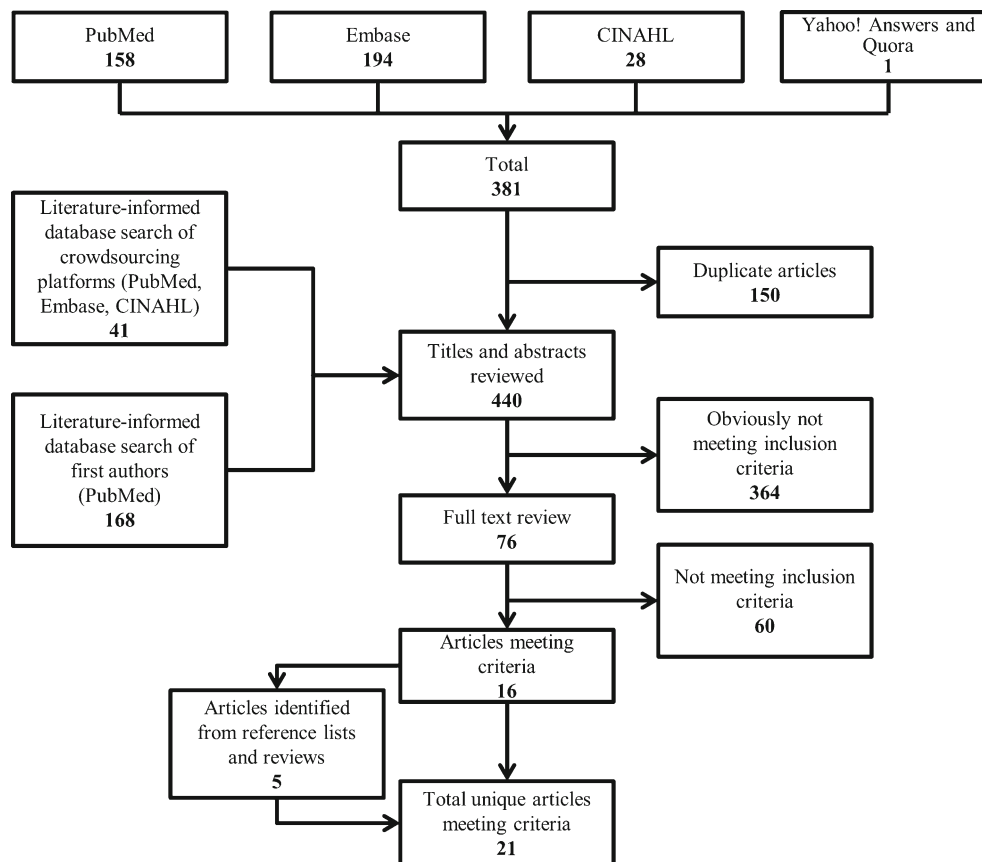


Figure 1. Results of the systematic literature search for health-related crowdsourcing studies. This figure shows the results of the systematic literature search for primary peer-reviewed articles that used crowdsourcing for health research.

Crowdsourcing Logistics

The length of the crowdsourcing study was mentioned in 18 of 21 (86 %) studies and varied from <2 hours to 10 months. Teams were used in 6 of 21 (29 %) studies, and intracrowd sharing was allowed in 7 of 21 (33 %) of studies. Monetary incentives were offered in 9 of 21 (43 %) studies and ranged from \$0.01 USD to \$2.50 USD per task, with many studies offering bonuses for either good completion or as a raffle/prize. The reported size of the crowd engaged in studies ranged from 5 to >110,000 people, with >136,395 people collectively engaged. Eleven of 21 studies (52 %) reported what advertising was used to attract participants.

All articles (100 %) reported employing some method for validating or improving the quality of data reported from the crowd. The types of validation techniques varied from inserting random questions with known answers into the task to screen for users who were incorrectly marking answers to comparing responses among multiple users and discarding outliers (Table 2).

Demographics of the Crowd

Reporting of the demographics of the crowd varied widely, with studies reporting crowd size (16/21, 76 %), age (7/21, 33 %), gender (5/21, 24 %), race (1/21, 5 %), geographic location (10/21, 48 %), occupation (4/21, 19 %), education (3/21, 14 %), relationship to the research question (4/21, 19 %), referral source (2/21, 10 %), conflict of interest (2/21, 10 %), reported motivation (3/21, 14 %), and viewer-to-participant ratio (4/21, 19 %). One (5 %) study reported age, sex, and race, and seven (33 %) studies reported at least one of these three descriptors (Table 3).

DISCUSSION

This is the first study to identify the types of crowdsourcing tasks used in primary peer-reviewed health research. This study has three main findings. First, we identified only 21 articles reflecting the use of crowdsourcing for health-

Table 1. Study Background Information

Reference no.	Research field	Methodology type	Objective	Outcome
1	Molecular biology	Observational	To evaluate whether players of the online protein folding game Foldit could find solutions to protein structure prediction problems	Several protein structures submitted to the CASP8 protein structure prediction competition ranked in the top 3 of all submitted entries, indicating that “the game has been designed in such a way that players can use it to solve scientific problems”
2	Molecular biology	Observational	To determine whether Foldit players could solve complex protein structure prediction problems	Foldit players were able to solve “challenging [protein] structure refinement problems” and outperform the Rosetta computer algorithm on some problems
3	Epidemiology	Observational	To track an influenza outbreak by using data submitted by users of a mobile phone application	Crowdsourced aggregated metric available in near real time “correlated highly” with CDC influenza metrics available after a minimum 1 week lag following data collection
4	Psychology	Survey	To examine the feasibility of using AMT respondents instead of the traditional undergraduate participant pool for survey research	“Crowdsourcing respondents were older, were more ethnically diverse, and had more work experience” and “reliability of the data... was as good as or better than the corresponding university sample”
5	Human behavior	Observational	To determine whether a scoping systematic review on asynchronous telehealth could be kept up to date through crowdsourcing	A total of 6 contributions were made to the wiki, but “none of the contributions enhanced the evidence base of the scoping review”
6	Molecular biology	Observational	To evaluate the “potential of automation in a social context for propagating the expert skills of top Foldit players and increasing the overall collective problem solving skills”	“Players take advantage of social mechanisms in the game to share, run, and modify recipes” indicating “potential for using automation tools to disseminate expert knowledge”
7	Molecular biology	Observational	To determine whether Foldit players could solve the structure of a retroviral protease	Players of the online game Foldit were able to solve a retroviral protease structure that had remained unsolved by automated methods
8	Molecular biology	Observational	To determine whether structured collaboration among individual Foldit players could produce better folding algorithms	5,400 different folding strategies were made by the Foldit community. By sharing and recombining successful “recipes” two dominant recipes emerged that outperformed previously published methods and bore “striking similarity to an unpublished algorithm developed by scientists over the same period”
9	Epidemiology	Survey	To use AMT to determine whether micro-monetary incentives and online reporting could be harnessed for public health surveillance of malaria	“This methodology provides a cost-effective way of executing a field study that can act as a complement to traditional public health surveillance methods”
10	Molecular biology	Observational	To determine whether crowdsourcing could be used to remodel a computationally designed enzyme	The Foldit community was able to “successfully guide large-scale protein design problems” and produced an “insertion, that increased enzyme activity >18-fold”
11	Human behavior	Survey	To use AMT to study the relationship between delay and probability discounting in human decision-making processes	AMT “findings were validated through the replication of a number of previously established relations” and the present study showed that “delay and probability discounting may be related, but are not manifestations of a single construct (e.g., impulsivity)”
12	Comparative genomics	Observational	To determine whether players of the online game Phylo could improve the multiple sequence alignment of the promoters of disease-related genes	A complex scientific problem was successfully embedded into the “casual” online game Phylo (much of the science was “hidden” leaving just a puzzle for users to work on). Players were able to improve multiple sequence alignment accuracy in “up to 70 % of the alignment blocks considered”
13	Pathology/hematology	Observational	To determine whether players of the online game MalariaSpot could accurately identify malaria parasites in digitized thick blood smears	By combining player inputs, “nonexpert players achieved a parasite counting accuracy higher than 99 %”
14	Pathology/hematology	Observational	To determine whether players of an online game could accurately identify RBCs infected with malaria parasites in digitized thin blood smears	Crowdsourcing resulted in a “diagnosis of malaria infected red blood cells with an accuracy that is within 1.25 % of the diagnostics decisions made by a trained medical professional”
15	Pathology/hematology	Observational	To expand on a prior small scale experiment by determining whether players of a large-scale online game could identify RBCs infected with malaria parasites	989 previously untrained gamers from 63 countries made more than 1 million cell diagnoses with a “diagnostic accuracy level that is comparable to those of expert medical professionals”

Table 1. (Continued)

Reference no.	Research field	Methodology type	Objective	Outcome
16	Radiology	Observational	To determine whether AMT could be used to conduct observer performance studies to optimize systems for medical imaging applications and to find out if the results were applicable to medical professionals	“Numerous parallels between the expert radiologist and the [knowledge workers]” were observed, pointing to the potential to cheaply test and optimize systems on the crowd before use with physicians
17	Radiology	Observational	To determine whether AMT KWs with minimal training could correctly classify polyps on CT colonography images	“The performance of distributed human intelligence is not significantly different from that of [computer-aided detection] for colonic polyp classification”
18	Public health	Survey	To determine whether crowdsourcing could be used to solicit feedback on oral health promotional materials	AMT provided an “effective method for recruiting and gaining feedback from English-speaking and Spanish-speaking people” on health promotional materials
19	Psychology	Survey	To determine whether classic behavioral cognitive experiments that require complex multi-trial designs could be replicated using AMT	“Data collected online using AMT closely resemble data collected in the lab under more controlled situations” for many of the experiments
20	Public health	Observational	To determine whether crowdsourcing could be used to annotate webcam images for analysis of behavioral modifications in active transportation following built-environment change	“Publicly available web data feeds and crowdsourcing have great potential for capturing behavioral change associated with built environments”
21	Public health	Observational	To determine whether crowdsourcing could be used to create a map of AED locations	Crowdsourcing generated “the most comprehensive AED map within a large metropolitan US region reported in the peer-reviewed literature” with 1,429 AEDs identified

Reference number key

- 1- (Cooper et al., 2010)¹²
 2- (Cooper et al., 2010)¹³
 3- (Freifeld et al., 2010)¹⁴
 4- (Behrend et al., 2011)¹⁵
 5- (Bender et al., 2011)¹⁶
 6- (Cooper et al., 2011)¹⁷
 7- (Khatib et al., 2011)¹⁸
 8- (Khatib et al., 2011)¹⁹
 9- (Chunara et al., 2012)²⁰
 10- (Eiben et al., 2012)²¹
 11- (Jarmolowicz et al., 2012)²²
 12- (Kawrykow et al., 2012)²³
 13- (Luengo-Oroz et al., 2012)²⁴
 14- (Mavandadi et al., 2012)²⁵
 15- (Mavandadi et al., 2012)²⁶
 16- (McKenna et al., 2012)²⁷
 17- (Nguyen et al., 2012)²⁸
 18- (Turner et al., 2012)²⁹
 19- (Crump et al., 2013)³⁰
 20- (Hipp et al., 2013)³¹
 21- (Merchant et al., 2013)³²

related research. Second, we found that these studies used crowdsourcing for four different principal objectives and that there are several advantages to utilizing crowdsourcing. Third, we found considerable variability in how the methods of crowdsourcing were reported.

While citizen science has been in existence for more than a century and crowdsourcing has been used in science for at least a decade, crowdsourcing has been utilized primarily by non-medical fields, and little is known about its potential in health research. Every health field from studying chronic diseases to global health has a potential need for human computing power that crowdsourcing could fill in order to accelerate research. Prior work has heralded crowdsourcing as a feasible method for data collection, but a clear roadmap for

the types of questions crowdsourcing could answer and the ways it could be applied has been lacking. Understanding how crowdsourcing has been used successfully in health research is crucial to understanding where crowdsourcing fits in the health care space, especially when there may be higher standards or tighter regulations for data quality and validity compared to the science fields that were early adopters.

The limited number of articles using crowdsourcing is surprising given the potential benefits of this approach. Although we identified 21 articles, most of which used crowdsourcing successfully, crowdsourcing clearly is not used pervasively in health research, and it is important to understand the quality of data it provides. Even though the use of crowdsourcing in health research is in its infancy, the

Table 2. Logistics of the Crowdsourcing

Reference no.	Length of crowd sourcing	Platform	Team	Intra-crowd sharing	Data collected or processed	Complexity of the task	Task time	Advertisement of the project	Skill set required	Monetary incentives (USD)	Data validation techniques
1	No	Custom game	Yes	Yes	Game play data and puzzle results	Complex - solving protein structure puzzles	No	No - but Foldit community was used	Three-dimensional problem solving and visualization	Not monetary	Final submissions were chosen by experts who looked at aggregates of returned solutions from Foldit players Compared response to known answers
2	No	Custom game	Yes	Yes - social aspect of game allows for sharing and collaboration	Game play data and puzzle results, optional demographic survey and motivation survey on website, survey sent to top players, website statistics	Complex - solving protein structure prediction puzzles	Mean 155 to 216 minutes to complete, depending on puzzle type, no time limit	Rosetta@home forum announcements, publicity from press, talks, and word of mouth	Three-dimensional problem solving and visualization	Not monetary - paper described several features to make game more engaging including scoring, achievement system, player status and rank, social aspect (teams, chats, forums, wiki, community), and connection to scientific outcomes	Compared response to known answers
3	N/A - continuous monitoring that had run for about a year at time of publication	Mobile phone application	None	None	User observational report collected along with optional GPS and/or photograph of person involved	Moderate - report observation	N/A - unlimited time for duration of study	No	Use of smartphone	Not monetary	Reports filtered for spam, duplicates, and mistakenly submitted reports, response data compared to CDC data
4	2 days	AMT	None	None	Survey responses including some demographics	Easy - answer survey questions	Mean 26.45 minutes, unknown time limit	AMT	None	\$0.80 for survey	Surveys were screened for excessive response consistency (Long String Index) and random responders (screened for by having a pair of Likert scale items that should have opposite responses and also having a pair of Likert scale items that should have similar responses)

Table 2. (Continued)

Reference no.	Length of crowd sourcing	Platform	Team	Intra-crowd sharing	Data collected or processed	Complexity of the task	Task time	Advertisement of the project	Skill set required	Monetary incentives (USD)	Data validation techniques
5	10 months	Wiki	None	Yes - wiki was viewable by all	User changes to wiki, brief survey, web analytics of user activity	Complex - edit evidence for systematic review	N/A - unlimited duration of study	2 eBulletins and a Rapid Response article in <i>Open Medicine</i> , e-mail list, and eBulletin by Canadian Telehealth Forum	Knowledge to edit systematic review and use a wiki	Not monetary	Spam filters, user registration required to contribute content, site "rollbacks" by administrators when spam was posted
6	~3 months	Custom game	Yes	Yes	Game play and puzzle results	Complex - solving protein structure prediction puzzles	No	No - but already active Foldit community was used	Three-dimensional problem solving and visualization	Not monetary	Players could rate recipes they used
7	3 weeks	Custom game	Yes	Yes	Game play and puzzle results	Complex - solving protein structure prediction puzzles and creating solution algorithms	N/A - unlimited duration of study	No - but already active Foldit community was used	Three-dimensional problem solving and visualization	Not monetary	In game scoring of structures, top Foldit models screened for quality by Phaser program, crystal structure was validated
8	3.5 months	Custom game	Yes	Yes - within the game (group play, game chat) and outside of the game (wiki)	Game play and puzzle results	Complex - solving protein structure prediction puzzles and creating solution algorithms	N/A - unlimited duration of study	No - but already active Foldit community was used	Three-dimensional problem solving and visualization	Not monetary	Successful recipes were more likely to be shared and used by other players
9	42 days	AMT	None	None	Survey responses	Easy - answer short survey	No	AMT	None	\$0.02 for survey	Surveys were removed if repeat users, if answers had contradictory information, or if users were located outside of solicited area
10	1 week per puzzle	Custom game	None	None	Game play data and puzzle results	Complex - solving protein structure prediction puzzles	N/A - unlimited tries as long as puzzle was posted	No - but already active Foldit community was used	Three-dimensional problem solving and visualization	Not monetary	Predicted enzyme structure and activity verified experimentally. Advanced Foldit player served as intermediary between the Foldit community and experimental laboratory

Table 2. (Continued)

Reference no.	Length of crowd sourcing	Platform	Team	Intra-crowd sharing	Data collected or processed	Complexity of the task	Task time	Advertisement of the project	Skill set required	Monetary incentives (USD)	Data validation techniques
11	~24 hours	AMT	None	None	Survey results	Easy - answer survey	Mean 1,472 seconds, unknown time limit	AMT	None	\$2.50 for survey+\$2.50 bonus for good completion	KWs needed to have high prioritating, had to show they understood instructions, had to complete 80 % of survey, couldn't complete survey too quickly, couldn't show non-varying responses
12	7 months	Custom game	None	None	User puzzle solutions, player activity statistics	Moderate - solving puzzles which represent sequence alignment problems	Timed to give more of a game feel	No - but mentioned publicity from media	Puzzle solving	Not monetary - paper described several features to make the puzzles more entertaining including levels, timing, and high score boards	Answers were compared to current best sequence alignment, best scoring sequence was kept
13	1 month	Custom game	None	None	Game play data and game results	Moderate - identify malaria parasites in blood samples	1 minute per image time limit	No	Image analysis	Not monetary - paper described high score table	Gold standard was used and quorum algorithm combined user inputs to achieve higher accuracy
14	No	Custom game	None	None	Game play data and game results	Moderate - identify RBCs infected with malaria parasites	No	No	Image analysis	Not monetary	Multiple gamers reviewed each RBC image and player input included based on player's current score on embedded control images
15	3 months	Custom game	None	None	Game play data and game results	Moderate - identify RBCs infected with malaria parasites	No	No	Image analysis	Not monetary	Subset of pictures were known controls (ranked by medical experts), used to determine each player's performance and weight his or her answers accordingly. Only included gamers who rated at least 100 images. Multiple gamer reviews per image

Table 2. (Continued)

Reference no.	Length of crowd sourcing	Platform	Team	Intra-crowd sharing	Data collected or processed	Complexity of the task	Task time	Advertisement of the project	Skill set required	Monetary incentives (USD)	Data validation techniques
16	<3 months	AMT	None	None	Results from KW's and usage statistics from AMT	Moderate - classify polyps in CT colonography images and videos	Mean 17.61 to 20.36 seconds, time limit of 20 minutes	AMT	Visual picture/pattern recognition	\$0.01 per task+\$5 bonus for top 10 KW's	KW's needed to have high prior rating and complete qualification test, multiple KW's read each image, results compared to expert radiologist
17	3-3.5 days	AMT	None	None	Results from KW's and usage statistics from AMT	Moderate - classify polyps on CT colonography images	Mean 24.1 to 35.5 seconds, time limit of 20 minutes	AMT	Visual picture/pattern recognition	\$0.01 per task	KW's needed to have high prior rating, multiple KW's read each image, unreliable KW outliers were excluded (however, this did not change outcome)
18	<12 days	AMT	None	None	Survey responses including some demographics	Easy - answer short survey	Mean 4.7 minutes, time limit of 15 minutes	AMT	None	\$0.25 for survey	KW's needed to have high prior rating, surveys included questions to test comprehension of passages that were read
19	Ranged from ~2 hours to 4 days Not reported for some experiments	AMT	None	None	Results from KW's and usage statistics from AMT	Ranged from easy (e.g., response times) to moderate (e.g., learning)	5 to 30 minutes	AMT	None	Varied across experiments but included \$0.10, \$0.50, \$0.75, \$1 with \$10 bonus for 1 in 10 raffle winners, or \$2 with up to \$2.50 bonus based on performance	Varied across experiments but included exclusion of outlier responses, minimum user ratings, initial performance and completion requirements, verification user was unique, demonstration by user that instructions were understood
20	<8 hours	AMT	None	None	Results from KW's	Easy - annotate images	No	AMT	None	\$0.01 per annotation	5 unique reviews per image

Table 2. (Continued)

Reference no.	Length of crowd sourcing	Platform	Team sharing	Data collected or processed	Complexity of the task	Task time	Advertisement of the project	Skill set required	Monetary incentives (USD)	Data validation techniques
21	8 weeks	Custom mobile phone application and website	Yes	Data from mobile application, website, and participant survey	Moderate - locate and photograph an AED, strategize to collect most AEDs for grand prize	N/A - no time limit per task	Radio, TV, print, social, and mobile media	Use of smartphone	\$50 to first individual to identify each of 200 pre-identified, unmarked AEDs, \$10,000 to the individual or team that submitted the most eligible AEDs	Compared GPS coordinates of AED submissions to GPS coordinates of building locations, to previously known location from AED manufacturer, or to location that was checked in person, teams with >0.5 % false or inaccurate entries disqualified from grand prize competition

Reference number key

- 1- (Cooper et al., 2010)¹²
- 2- (Cooper et al., 2010)¹³
- 3- (Freifeld et al., 2010)¹⁴
- 4- (Behrend et al., 2011)¹⁵
- 5- (Bender et al., 2011)¹⁶
- 6- (Cooper et al., 2011)¹⁷
- 7- (Khatib et al., 2011)¹⁸
- 8- (Khatib et al., 2011)¹⁹
- 9- (Chunara et al., 2012)²⁰
- 10- (Eiben et al., 2012)²¹
- 11- (Jarmolowicz et al., 2012)²²
- 12- (Kawrykow et al., 2012)²³
- 13- (Luengo-Oroz et al., 2012)²⁴
- 14- (Mavandadi et al., 2012)²⁵
- 15- (Mavandadi et al., 2012)²⁶
- 16- (McKenna et al., 2012)²⁷
- 17- (Nguyen et al., 2012)²⁸
- 18- (Turner et al., 2012)²⁹
- 19- (Crump et al., 2013)³⁰
- 20- (Hipp et al., 2013)³¹
- 21- (Merchant et al., 2013)³²

*"No" indicates that data were not explicitly reported in the paper; "NP" indicates that data were collected but not published

Table 3. Demographics of the Crowd

Reference no.	Size	Age (years old)	Gender	Racial/ethnic background	Geographic location	Occupation	Education	Relationship to the research question	Referral source	Conflicting interests	Motivation	Viewer to participant ratio
1	No	No	No	No	No	No	No	No	No	No	No	No
2	No	Yes - ~67% ≤35	Yes - >75% male, <25% female	No	Yes	Yes	Yes - >25% chose "some masters program"	No	Yes - initially 59% from referring sites, 12% from search engines, 29% was direct traffic. After 1 month, 41% direct traffic	No	Yes - purpose, achievement, social or immersion	No
3	>110,000 downloads, >2,400 submissions	No	No	No	NP	No	No	Yes - outbreak reporting relevant to community	No	No	No	>110,000 application downloads : >2,400 submissions (2.18%)
4	270	Yes - 33% 18 to 25, 14% 46+	Yes - 36% male, 63% female	Yes	Yes	Yes	Yes - 48% 4-year degree or greater	No	No	No	Yes - compensation ranked highly	No
5	5	Yes - all 30 to 60	Yes - 75% user accounts male, 25% female	No	Yes	Yes	No	No	NP	Yes	No	875 unique visitors : 5 unique contributors (0.57%)
6	771	No	No	No	No	No	No	No	No	No	No	No
7	No	No	No	No	No	No	No	No	No	No	No	No
8	721	No	No	No	No	No	No	No	No	No	No	No
9	330 and 442	Yes - all 18 to 62	Yes - 62% male, 37% female	No	Yes	No	No	Yes - research relevant to crowd's community	No	No	No	No
10	No	No	No	No	No	No	No	No	No	No	No	No
11	1,190	NP	NP	No	Yes	No	NP	No	No	No	No	No
12	12,252 registered players, 2,905 regular players	No	No	No	No	No	No	Yes - players work on specific disease they feel connection to (e.g., cancer)	No	No	No	42% of registered users failed to complete a single puzzle
13	>6,000	No	No	No	Yes	No	No	No	Yes - social media was main source	No	No	No
14	31	Yes - all 18 to 40	No	No	No	No	No	No	No	No	No	No
15	2,150	No	No	No	Yes	No	No	No	No	No	No	No
16	414	No	No	No	No	No	No	No	No	No	No	No
17	150 and 102	No	No	No	No	No	No	No	No	No	No	No
18	399	Yes - 79% 18 to 40	Yes - 53% female	No	Yes	No	Yes - 55% college or graduate degree	No	No	No	No	No

Table 3. (Continued)

Reference no.	Size	Age (years old)	Gender	Racial/ethnic background	Geographic location	Occupation	Education	Relationship to the research question	Referral source	Conflicting interests	Motivation	Viewer to participant ratio
19	855 total (varied from 32-234 per experiment)	Varied between 32-234 per experiment and NP	No	No	Yes	No	No	No	No	No	No	Dropout rates reported
20	No	No	No	No	No	No	No	No	No	No	No	No
21	313 teams and individuals	Yes - 33% ≥ 41	No	No	Yes	Yes	No	Yes - some participants cite "personal" as motivation	No	Yes - excluded from participation: the investigator team, project vendors, and project sponsors	Yes - contribution to an important cause, fun, money, personal, education	No

Reference number key

- 1- (Cooper et al., 2010)¹²
- 2- (Cooper et al., 2010)¹³
- 3- (Freifeld et al., 2010)¹⁴
- 4- (Behrend et al., 2011)¹⁵
- 5- (Bender et al., 2011)¹⁶
- 6- (Cooper et al., 2011)¹⁷
- 7- (Khatib et al., 2011)¹⁸
- 8- (Khatib et al., 2011)¹⁹
- 9- (Chunara et al., 2012)²⁰
- 10- (Eiben et al., 2012)²¹
- 11- (Jarmolowicz et al., 2012)²²
- 12- (Kawrykow et al., 2012)²³
- 13- (Luengo-Oroz et al., 2012)²⁴
- 14- (Mavandadi et al., 2012)²⁵
- 15- (Mavandadi et al., 2012)²⁶
- 16- (McKenna et al., 2012)²⁷
- 17- (Nguyen et al., 2012)²⁸
- 18- (Turner et al., 2012)²⁹
- 19- (Crump et al., 2013)³⁰
- 20- (Hipp et al., 2013)³¹
- 21- (Merchant et al., 2013)³²

*"No" indicates that data were not explicitly reported in the paper; "NP" indicates that data were collected but not published

papers we identified successfully used crowdsourcing to solve protein structure problems,¹⁸ improve alignment of promoter sequences,²³ track H1N1 influenza outbreaks in near real time,¹⁴ classify colonic polyps,^{27,28} and identify RBCs infected with *Plasmodium falciparum* parasites.^{24–26} Furthermore, as Mavandadi et al. point out, one way around the problem of involving lay people in making a medical diagnosis is to use crowdsourcing to distill the data for a medical professional, who can then make the final decision. For example, a pathologist must look at more than 1,000 RBCs to rule a sample negative, but if crowdsourcing identifies RBCs that are infected, all a pathologist has to do is officially confirm the diagnosis with a single image.²⁵ This could be especially useful in resource-poor areas.

Crowdsourcing used for health research has been employed to accomplish one of four main categories of tasks: problem solving, data processing, surveillance/monitoring, and surveying. Data from existing studies show that crowdsourcing has the potential to beneficially address the following points: quality, cost, volume, speed, and novel science. Crowdsourcing has been demonstrated to be a viable way to increase the accuracy of computer recognition of RBCs infected with malaria parasites²⁵ (quality), be a low cost alternative to more traditional behavioral research and epidemiology studies^{15,20,22,30} (cost), engage over one hundred thousand people in a research problem¹⁴ (volume), allow research to progress much faster than if processed by investigators alone^{14,15,20,22–24,26–32} (speed), and produce new scientific discoveries^{13,18,19,21,23} (novel science).

Additionally, there is the advantage of an untapped expertise of the crowd. Even though these crowdsourced projects are not asking scientific experts to participate, participants have been found to be experts at puzzles and problem solving, which would make them specifically adept at solving protein structures in Foldit^{12,13,17–19,21} and solving multiple sequence alignment with Phylo.²³ Presumably, among members of the public one could find experts at many different tasks, especially when the task is presented as a game that benefits science. Finally, crowdsourced projects raise public awareness about the project and about science in general.

The papers identified in this study varied widely in the amount and type of data that were reported about the crowd and the experimental setup. Crowdsourcing articles rarely reported data about the demographics of the crowd participating, including information standard to most clinical trials such as the size of the cohort, age, gender, and geographic location.¹¹ These data and others such as motivation for participation, education, and occupation are crucial for understanding the people involved in the research. Ideally, all papers that use crowdsourcing should include at a minimum data regarding the demographics and logistics of

data collection (Appendix Table 5). If collected in the future, these data would provide crucial information to help the scientific community understand how the crowd works and how to best maximize the use of crowdsourcing. However, the crowdsourcing methodology for health research is in an early phase of development, and there is additional work to be done to develop these methods and related reporting standards.

When thinking of conducting a crowdsourced study, it may be difficult to choose or create the most appropriate platform. In some cases researchers used their own custom platforms and in others they employed AMT. We examined studies that employed the crowd to perform crowdsourced tasks; however, there are also platforms that allow one to post a “challenge” and offer a monetary reward for the best solution. Kaggle (kaggle.com) allows scientists and others to post complex data analysis problems along with monetary rewards for the best solution,³³ and InnoCentive (innocentive.com) is a more general platform that allows prizes to be posted for any sort of research and development problem.³⁴ It is also worth noting that crowdfunding websites have become popular and may be a potential way to fund research projects. Examples include RocketHub (rockethub.com) and Petridish (petridish.org).³⁵

Although we conducted a review of references and review articles in addition to a crowdsourced search for literature, our results do not include all articles. Posting a survey on the Internet (collecting research data about Internet users themselves by having them answer questions or perform tasks) has been around for >15 years¹⁰ and therefore is not a novel research method. Our review only includes a sample of projects that used the Internet to survey participants. This specific type of crowdsourcing has been reviewed elsewhere.³⁶ Additionally, there are other types of research that are sometimes referred to as crowdsourcing but do not meet our definition of crowdsourcing. They involve investigators mining data that have been generated by users, but generally not for research purposes, such as estimating influenza activity by analyzing Twitter (twitter.com) posts.³⁷ Other examples reviewed elsewhere include investigators mining data or surveying online communities such as PatientsLikeMe (patientslikeme.com) and 23andMe (23andme.com).³⁸ Our search did not include gray literature or searching the Internet using a search engine. The goal of the study, however, was to characterize primary peer-reviewed health research. While several projects may be featured on the World Wide Web and science magazines, it is imperative that projects are published in academic journals so that the scientific community can validate the methods and demonstrate the varied, interesting, and successful uses of crowdsourcing in health and medical research.

CONCLUSION

Crowdsourcing has been used to help answer important health-related research questions. Utilizing crowdsourcing can improve the quality, cost, and speed of a research project while engaging large segments of the public and creating novel science. This methodology serves as an alternative approach for studies that could benefit from large amounts of manual data processing, surveillance conducted by people around the world, specific skills that members of the public may have, or diverse subject pools that can be surveyed at low cost. In this systematic review, we identified four types of research needs that have been addressed by crowdsourcing and specify criteria that future studies should meet in order to help standardize the use of crowdsourcing in health and medical research.

Acknowledgments:

Contributors: The authors thank Hope Lappen, MLIS, for her assistance with designing the systematic literature search.

Funders: NIH, K23 grant 10714038 (Merchant).

Prior presentations: None

Conflict of Interest: The authors declare that they do not have any conflicts of interest.

Corresponding Author: Raina M. Merchant, MD, MSHP; Perelman School of Medicine, University of Pennsylvania, 423 Guardian Drive, Blockley Hall, Philadelphia, PA 19104, USA (e-mail: raina.merchant@uphs.upenn.edu).

REFERENCES

1. Sobel D. *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*. New York: Walker Publishing Company; 1995.
2. National Audubon Society. History of the Christmas Bird Count [Internet]. 2012 [cited 2012 Dec 24]. Available from: <http://birds.audubon.org/history-christmas-bird-count>
3. ICT Data and Statistics Division. Measuring the Information Society 2012 [Internet]. Geneva, Switzerland: International Telecommunication Union; 2012 [cited 2012 Oct 14]. Available from: http://www.itu.int/ITU-D/ict/publications/idi/material/2012/MIS2012_without_Annex_4.pdf
4. Lintott C, Schawinski K, Bamford S, Slosar A, Land K, Thomas D, et al. Galaxy Zoo 1: data release of morphological classifications for nearly 900000 galaxies. *Monthly Notices of the Royal Astronomical Society*. 2011;410(1):166–78. doi:10.1111/j.1365-2966.2010.17432.x.
5. Sayigh L, Guick N, Hastie G, Tyack P. Repeated call types in short-finned pilot whales, *Globicephala macrorhynchus*. *Marine Mammal Science*. 2012;29(2):312–24. doi:10.1111/j.1748-7692.2012.00577.x.
6. Nature Publishing Group. Scientific American launches Citizen Science Whale-Song Project, Whale FM [Internet]. 2011 [cited 2012 Oct 14]. Available from: http://www.nature.com/press_releases/sciam-whale.html
7. Sullivan BL, Wood CL, Iliif MJ, Bonney RE, Fink D, Kelling S. eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*. 2009;142(10):2282–92. doi:10.1016/j.biocon.2009.05.006.
8. Marris E. Supercomputing for the birds. *Nature*. 2010;466(7308):807. doi:10.1038/466807a. PubMed PMID: 20703280.
9. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19–22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948. [cited 2012 Jul 21]. Available from: <http://www.who.int/about/definition/en/print.html>
10. Reips U-D, Birnbaum MH. Behavioral research and data collection via the Internet. In: Proctor RW, Vu K-PL, eds. *Handbook of human factors in Web design*. 2nd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2011:563–85.
11. Higgins JPT, Deeks JJ (editors). Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [Internet]. The Cochrane Collaboration; 2011 [updated 2011 Mar; cited 2012 Aug 1]. Available from www.cochrane-handbook.org
12. Cooper S, Treuille A, Barbero J, Leaver-Fay A, Tuite K, Khatib F, et al. The challenge of designing scientific discovery games. *California: Proceedings of the Fifth International Conference on the Foundations of Digital Games*; Monterey; 2010:40–7. 1822354: ACM.
13. Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, et al. Predicting protein structures with a multiplayer online game. *Nature*. 2010;466(7307):756–60. doi:10.1038/nature09304. PubMed PMID: 20686574; PubMed Central PMCID: PMC2956414.
14. Freifeld CC, Chunara R, Mekar SR, Chan EH, Kass-Hout T, Ayala Iacucci A, et al. Participatory epidemiology: use of mobile phones for community-based health reporting. *PLoS medicine*. 2010;7(12):e1000376. doi:10.1371/journal.pmed.1000376. PubMed PMID: 21151888; PubMed Central PMCID: PMC2998443.
15. Behrend TS, Sharek DJ, Meade AW, Wiebe EN. The viability of crowdsourcing for survey research. *Behavior research methods*. 2011;43(3):800–13. doi:10.3758/s13428-011-0081-0. PubMed PMID: 21437749.
16. Bender J, O'Grady L, Desphande A, Cortino A, Saffie L, Husereau D, et al. Collaborative authoring: a case study of the use of a wiki as a tool to keep systematic reviews up to date. 2011.
17. Cooper S, Khatib F, Makedon I, Lu H, Barbero J, Baker D, et al. Analysis of social gameplay macros in the Foldit cookbook. *France: Proceedings of the 6th International Conference on Foundations of Digital Games*; Bordeaux; 2011:9–14. 2159367: ACM.
18. Khatib F, DiMaio F, Foldit Contenders G, Foldit Void Crushers G, Cooper S, Kazmierczyk M, et al. Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature structural & molecular biology*. 2011;18(10):1175–7. doi:10.1038/nsmb.2119. PubMed PMID: 21926992.
19. Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popovic Z, et al. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(47):18949–53. doi:10.1073/pnas.1115898108. PubMed PMID: 22065763; PubMed Central PMCID: PMC3223433.
20. Chunara R, Chhaya V, Bane S, Mekar SR, Chan EH, Freifeld CC, et al. Online reporting for malaria surveillance using micro-monetary incentives, in urban India 2010–2011. *Malaria journal*. 2012;11:43. doi:10.1186/1475-2875-11-43. PubMed PMID: 22330227; PubMed Central PMCID: PMC3305483.
21. Eiben CB, Siegel JB, Bale JB, Cooper S, Khatib F, Shen BW, et al. Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nature biotechnology*. 2012;30(2):190–2. doi:10.1038/nbt.2109. PubMed PMID: 22267011.
22. Jarmolowicz DP, Bickel WK, Carter AE, Franck CT, Mueller ET. Using crowdsourcing to examine relations between delay and probability discounting. *Behavioural processes*. 2012;91(3):308–12. doi:10.1016/j.beproc.2012.09.001. PubMed PMID: 22982370.
23. Kawrykow A, Roumanis G, Kam A, Kwak D, Leung C, Wu C, et al. Phylot: a citizen science approach for improving multiple sequence alignment. *PLoS One*. 2012;7(3):e31362. doi:10.1371/journal.pone.0031362. PubMed PMID: 22412834; PubMed Central PMCID: PMC3296692.
24. Luengo-Oroz MA, Arranz A, Frean J. Crowdsourcing malaria parasite quantification: an online game for analyzing images of infected thick blood smears. *Journal of medical Internet research*. 2012;14(6):e167. doi:10.2196/jmir.2338. PubMed PMID: 23196001; PubMed Central PMCID: PMC3510720.
25. Mavandadi S, Dimitrov S, Feng S, Yu F, Sikora U, Yaglidere O, et al. Distributed medical image analysis and diagnosis through crowd-sourced games: a malaria case study. *PLoS One*. 2012;7(5):e37245. doi:10.1371/journal.pone.0037245. PubMed PMID: 22606353; PubMed Central PMCID: PMC3350488.

26. **Mavandadi S, Dimitrov S, Feng S, Yu F, Yu R, Sikora U, et al.** Crowdsourced BioGames: managing the big data problem for next-generation lab-on-a-chip platforms. *Lab on a chip*. 2012;12(20):4102–6. doi:10.1039/c2lc40614d. PubMed PMID: 22918378; PubMed Central PMCID: PMC3477593.
27. **McKenna MT, Wang S, Nguyen TB, Burns JE, Petrick N, Summers RM.** Strategies for improved interpretation of computer-aided detections for CT colonography utilizing distributed human intelligence. *Medical image analysis*. 2012;16(6):1280–92. doi:10.1016/j.media.2012.04.007. PubMed PMID: 22705287; PubMed Central PMCID: PMC3443285.
28. **Nguyen TB, Wang S, Anugu V, Rose N, McKenna M, Petrick N, et al.** Distributed Human Intelligence for Colonic Polyp Classification in Computer-aided Detection for CT Colonography. *Radiology*. 2012;262(3):824–33. doi:10.1148/radiol.11110938. PubMed PMID: 22274839; PubMed Central PMCID: PMC3285221.
29. **Turner AM, Kirchoff K, Capurro D.** Using crowdsourcing technology for testing multilingual public health promotion materials. *Journal of medical Internet research*. 2012;14(3):e79. doi:10.2196/jmir.2063. PubMed PMID: 22664384.
30. **Crump MJ, McDonnell JV, Gureckis TM.** Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS One*. 2013;8(3):e57410. doi:10.1371/journal.pone.0057410. PubMed PMID: 23516406; PubMed Central PMCID: PMC3596391.
31. **Hipp JA, Adlakha D, Eyler AA, Chang B, Pless R.** Emerging technologies: webcams and crowd-sourcing to identify active transportation. *American journal of preventive medicine*. 2013;44(1):96–7. doi:10.1016/j.amepre.2012.09.051. PubMed PMID: 23253658.
32. **Merchant RM, Asch DA, Hershey JC, Griffis HM, Hill S, Saynisch O, et al.** A crowdsourcing innovation challenge to locate and map automated external defibrillators. *Circulation Cardiovascular quality and outcomes*. 2013;6(2):229–36. doi:10.1161/CIRCOUTCOMES.113.000140. PubMed PMID: 23481522.
33. **Carpenter J.** May the best analyst win. *Science*. 2011;331(6018):698–9. doi:10.1126/science.331.6018.698. PubMed PMID: 21311005.
34. **Hussein A.** Brokering knowledge in biosciences with InnoCentive. Interview by Semahat S. Demir. *IEEE engineering in medicine and biology magazine: the quarterly magazine of the Engineering in Medicine & Biology Society*. 2003;22(4):26–7. PubMed PMID: 14515687.
35. **Wheat RE, Wang Y, Byrnes JE, Ranganathan J.** Raising money for scientific research through crowdfunding. *Trends in ecology & evolution*. 2013;28(2):71–2. doi:10.1016/j.tree.2012.11.001. PubMed PMID: 23219380.
36. **Mason W, Suri S.** Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods*. 2012;44(1):1–23. doi:10.3758/s13428-011-0124-6. PubMed PMID: 21717266.
37. **Signorini A, Segre AM, Polgreen PM.** The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS One*. 2011;6(5):e19467.
38. **Swan M.** Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *Journal of medical Internet research*. 2012;14(2):e46. doi:10.2196/jmir.1988. PubMed PMID: 22397809.

APPENDIX

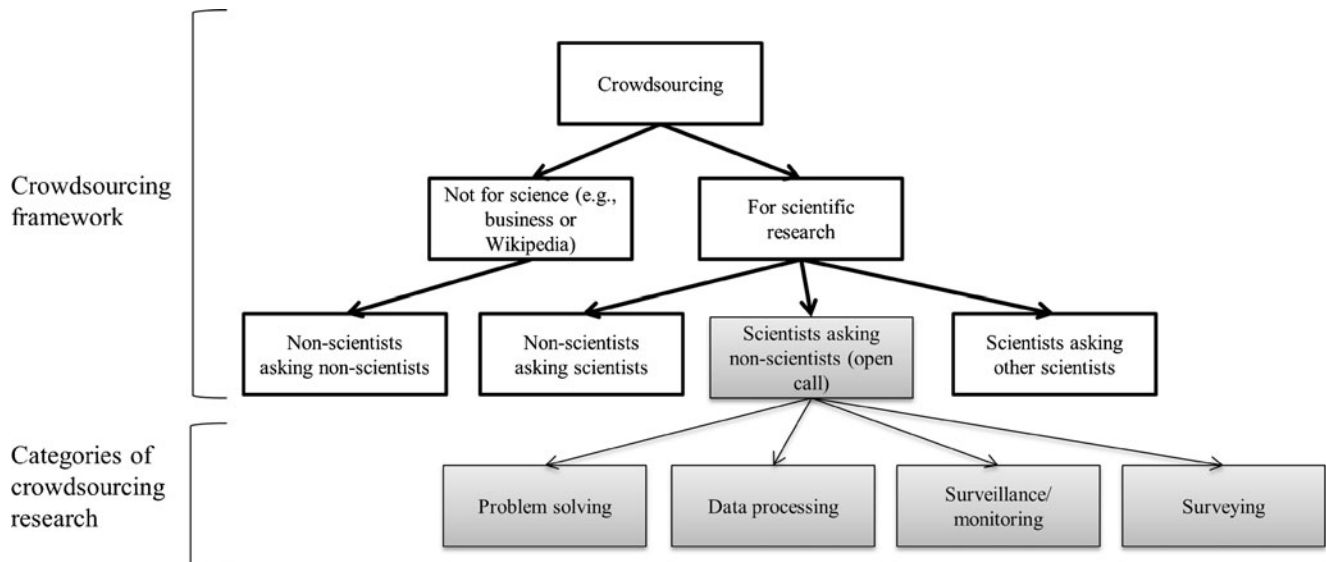


Figure 2. Crowdsourcing framework and categories.

Table 4. Category of Research Task Crowdsourcing was Utilized to Accomplish

Ref no.	Title	Category
1	The challenge of designing scientific discovery games	Problem solving
2	Predicting protein structures with a multiplayer online game	Problem solving
3	Participatory epidemiology: use of mobile phones for community-based health reporting	Surveillance/ monitoring
4	The viability of crowdsourcing for survey research	Survey
5	Collaborative authoring: a case study of the use of a wiki as a tool to keep systematic reviews up to date	Data processing
6	Analysis of social gameplay macros in the Foldit cookbook	Problem solving
7	Crystal structure of a monomeric retroviral protease solved by protein folding game players	Problem solving
8	Algorithm discovery by protein folding game players	Problem solving
9	Online reporting for malaria surveillance using micro-monetary incentives, in urban India 2010-2011	Surveillance/ monitoring
10	Increased Diels-Alderase activity through backbone remodeling guided by Foldit players	Problem solving
11	Using crowdsourcing to examine relations between delay and probability discounting	Survey
12	Phylo: a citizen science approach for improving multiple sequence alignment	Problem solving
13	Crowdsourcing malaria parasite quantification: an online game for analyzing images of infected thick blood smears	Data processing
14	Distributed medical image analysis and diagnosis through crowd-sourced games: a malaria case study	Data processing
15	Crowd-sourced BioGames: managing the big data problem for next generation lab-on-a-chip platforms	Data processing
16	Strategies for improved interpretation of computer-aided detections for CT colonography utilizing distributed human intelligence	Data processing
17	Distributed human intelligence for colonic polyp classification in computer-aided detection for CT colonography	Data processing
18	Using crowdsourcing technology for testing multilingual public health promotion materials	Survey
19	Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research	Survey
20	Emerging technologies: webcams and crowd-sourcing to identify active transportation	Data processing
21	A crowdsourcing innovation challenge to locate and map automated external defibrillators	Surveillance/ monitoring

Reference Number Key

- 1- (Cooper et al., 2010)^{1,2}
 2- (Cooper et al., 2010)^{1,3}
 3- (Freifeld et al., 2010)^{1,4}
 4- (Behrend et al., 2011)^{1,5}
 5- (Bender et al., 2011)^{1,6}
 6- (Cooper et al., 2011)^{1,7}
 7- (Khatib et al., 2011)^{1,8}
 8- (Khatib et al., 2011)^{1,9}
 9- (Chunara et al., 2012)^{2,0}
 10- (Eiben et al., 2012)^{2,1}
 11- (Jarmolowicz et al., 2012)^{2,2}
 12- (Kawrykow et al., 2012)^{2,3}
 13- (Luengo-Oroz et al., 2012)^{2,4}
 14- (Mavandadi et al., 2012)^{2,5}
 15- (Mavandadi et al., 2012)^{2,6}
 16- (McKenna et al., 2012)^{2,7}
 17- (Nguyen et al., 2012)^{2,8}
 18- (Turner et al., 2012)^{2,9}
 19- (Crump et al., 2013)^{3,0}
 20- (Hipp et al., 2013)^{3,1}
 21- (Merchant et al., 2013)^{3,2}

Table 5. Potential Data to Report for Health Research Using Crowdsourcing

Demographics of the crowd	Logistics of the crowdsourcing
Size of crowd	Length of time crowdsourcing was conducted
Age	Web platform and the use of a mobile platform
Gender	Use of individuals versus teams and intracrowd sharing techniques
Racial/ethnic background	Data collected or processed
Geographic location	Complexity of the task
Occupation	Time given to do the task
Education	Advertisement of project
Relationship to the research question	Skill set required
Referral source	Incentives offered
Conflicting interests	Data validation techniques
Motivation	
Viewer to participant ratio	