# Crowdsourcing in medical research: theory and practice

**Joseph Tucker** [Corresp., 1, 2] , **Suzanne Day** [1] , **Weiming Tang** [1] , **Barry Bayus** [3]

[1] University of North Carolina at Chapel Hill, Chapel Hill, United States

[2] London School of Hygiene & Tropical Medicine, University of London, London, United Kingdom

[3] Kenan-Flagler School of Business, University of North Carolina at Chapel Hill, Chapel Hill, United States

Corresponding Author: Joseph Tucker
Email address: jdtucker@med.unc.edu

The field of crowdsourcing for medicine has substantially expanded. We define crowdsourcing as an organization having a large group attempt to solve community problem, then share the solution with the broader public. Large groups of individuals can participate in medical research through open contests, hackathons, and related activities. The purpose of this literature review is to examine the definition, theory, and practice of crowdsourcing in medicine in order to facilitate crowdsourcing research. This multi-disciplinary review defines crowdsourcing for health, identifies theoretical antecedents (collective intelligence and open source models), and explores implications of the approach. Several critiques of crowdsourcing are also examined. Although several crowdsourcing definitions exist, there are two essential elements: (1) having a large group of individuals, including experts and non-experts, propose potential solutions; (2) sharing solutions with the public through implementation or open access materials. The public can be a central force in framing a common problem and developing feasible and compelling solutions. Crowdsourcing is related to, but distinct from other participatory research approaches. Crowdsourcing can be a useful for informing medical research, programs, and policy. A growing evidence base suggests that crowdsourcing in medicine can result in high-quality outcomes, broad community engagement, and more open science.

# Crowdsourcing in medical research: theory and practice

Joseph D. Tucker,[1,2,3] Suzanne Day,[1,4] Weiming Tang,[1,2,5] Barry Bayus[6]

[1] Institute for Global Health and Infectious Diseases, University North Carolina at Chapel Hill, Chapel Hill, USA

[2] Social Entrepreneurship to Spur Health (SESH) Global, Guangzhou, China

[3] Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[4] Department of Social Medicine, UNC School of Medicine, University North Carolina at Chapel Hill, Chapel Hill, USA

[5] Department of STD Control, Dermatology Hospital of Southern Medical University, Guangzhou, China

[6] Kenan-Flagler School of Business, University North Carolina at Chapel Hill, Chapel Hill, USA

Corresponding Author Contact Information: Dr Joseph Tucker, UNC Project-China, 2 Lujing Road, Guangdong Provincial STD Control Department, Guangzhou, 510095; Tel: (86) 13560294997   Email: jdtucker@med.unc.edu

## Abstract

The field of crowdsourcing for medicine has substantially expanded. We define crowdsourcing

as an approach to problem solving which involves an organization having a large group attempt

to solve a community problem, then sharing the solution with the broader public. Crowdsourcing

allows large groups of individuals to participate in medical research through open contests,

hackathons, and related activities. The purpose of this literature review is to examine the

definition, theory, and practice of crowdsourcing in medicine in order to facilitate crowdsourcing

research. This multi-disciplinary review defines crowdsourcing for health, identifies theoretical

antecedents (collective intelligence and open source models), and explores implications of the

approach. Several critiques of crowdsourcing are also examined. Although several

crowdsourcing definitions exist, there are two essential elements: (1) having a large group of

individuals, including experts and non-experts, propose potential solutions; (2) sharing solutions

with the public through implementation or open access materials. The public can be a central

force in framing a common problem and developing feasible and compelling solutions.

Crowdsourcing is related to, but distinct from other participatory research approaches.

Crowdsourcing can be a useful for informing medical research, programs, and policy. A growing

evidence base suggests that crowdsourcing in medicine can result in high-quality outcomes,

broad community engagement, and more open science.

49

## Introduction

51 Crowdsourcing is an approach to problem solving that has gained momentum in the past

52 decade.[1,2] Crowdsourcing involves having a group of non-experts and experts to propose

53 solutions to a problem, then sharing solutions with the broader public.[3] This concept has spurred

54 diverse health programs, including challenge contests (also called innovation contests,

55 inducement prize contests, and open contests), hackathons, online systems for collaboration, and

56 other activities (Table 1). Many crowdsourcing programs focus on medicine.[1,4,5] However, to

57 date descriptions of crowdsourcing for medicine have neglected the theoretical foundations of

58 this approach.

59

60 The diversity of crowdsourcing approaches complicates attempts to achieve a single overarching

61 theory.[6,7] Some have suggested that the bottom-up nature of engaging the public is antithetical to

62 theory-driven interventions. Others argue that the relatively brief history of crowdsourcing

63 makes it premature to develop a robust theory.[8] However, the conceptual basis of crowdsourcing

64 reaches well beyond the first use of the term. This history alongside more recent data on

65 collective intelligence and open-source models pave the way for a crowdsourcing theory.

66

## Survey Methodology

68 This literature review examined the peer-reviewed and gray literature on crowdsourcing

69 approaches related to health. We searched PubMed and Google Scholar to identify potential

70    studies for inclusion. This manuscript defines crowdsourcing for health, identifies theoretical

71    antecedents, considers relationships with other approaches, and examines common critiques.

72

73    **Crowdsourcing: A Definition**

74    There have been many definitions of crowdsourcing since Jeff Howe coined the term in 2005.[9]

75    The term is a portmanteau composed of "crowd" and "outsourcing." The original definition was

76    applied to describe companies outsourcing tasks to a group of individuals who worked

77    collectively or individually. Howe himself realized that this initial definition was overly narrow

78    and later expanded it to include the application of open-source principles to fields outside of

79    software. However, this definition fails to capture the subsequent obligation to share the solution

80    with the public. Henk van Ess suggested that crowdsourcing involves experts and non-experts

81    attempt to solve a problem, then freely sharing the solutions with the public.[3] This definition is

82    more relevant to medical applications and provides a clearer outline of the essential aspects of

83    crowdsourcing – group participation in problem solving and sharing solutions widely.

84

85    First, an organization has a group (including experts and non-experts) attempt to solve a

86    problem. The group could be working independently or collaborating as a team. The rationale for

87    sourcing solutions from a group rather than select individuals includes the following: (1) the

88    potential for groups to have expert knowledge and experiences in a related field; (2) the

89    importance of public participation and community consultation in health services; (3) the

90    potential for local end-users, patients, and others to be more actively engaged in the process of

91    developing new ideas; (4) the inclusion of people from the community to assist in designing

92    interventions that would be feasible and relevant in the local community. The group participation

93    component of crowdsourcing has been used by states, international organizations, and non-

94    profits for centuries. For example, in 1714, the British government wanted to find an accurate

95    method to measure a ship's longitudinal position. They offered a cash prize to whomever

96    developed a solution that met pre-specified benchmarks. This spurred many groups to focus on

97    enhanced methods for measuring longitude, resulting in important advances in this field.[10]

98

99    The second key component of crowdsourcing involves sharing the solution with the public. This

100   could be accomplished through implementing the solution in a local community[11] or creating

101   open access materials for public use.[12] For example, the rights to an exceptional crowdsourced

102   image would not be held by the crowdsourcing organizers, but would instead be made widely

103   available through creative commons attribution. Additionally, given that crowdsourcing takes a

104   solution from a group, there is an ethical responsibility to share the solution with the public.[13]

105   Crowdsourcing approaches may generate a range of materials and products that can be shared in

106   both digital and in-person formats. Some examples of ways that crowdsourced materials have

107   been shared include: providing crowdsourced images, concepts, and logos to the public through

108   an open access website;[12] widely distributing images through social media;[14] evaluating the

109   effectiveness of the crowdsourced output through a trial;[7,11,15] holding a series of in-person

110   workshops to communicate crowdsourced findings with key stakeholders.[16]

111

112   These two crowdsourcing components – group participation and sharing of solutions – are each

113   indebted to earlier multidisciplinary theory on collective intelligence and open source models,

114   respectively. The next two section explores these related concepts as they inform crowdsourcing.

115

116   **Collective Intelligence**

117   Collective intelligence suggests that in certain settings, a group is better able to solve difficult

118   problems than an individual working alone. The concept is not a universal statement about

119   groups being wiser than individuals, but rather that there are certain contexts wherein this is true.

120   The collective intelligence concept has a history in political science, philosophy, social science,

121   and biology. Perhaps the earliest mention of this concept was in 1785 when Marquis de

122   Condorcet published a theorem about the relative probability of a given group of individuals

123   arriving at a correct decision.[17] The theorem examines the optimal number of voters when

124   engaging in a group decision. The number is greater when there is a higher probability of each

125   voter making a correct decision; the number is small when there is a lower probably of each

126   voter making a correct decision. This provides a theoretical basis for democracy and has been

127   widely used in political science.[18,19] Within a health context, Condorcet's theorem has been used

128   in clinical diagnostic imaging[20] and reviewing organ transplant eligibility.[21]

129

130   Philosophers and others have contributed to the development of a collective intelligence concept.

131   The French philosopher Pierre Lévy defined collective intelligence as "a form of universally

132   distributed intelligence, constantly enhanced, coordinated in real time, and resulting in effective

133   mobilization of skills."[22] Social reformers have also used collective intelligence as a key guiding

134    principle. H. G. Wells described a "World Brain" concept that would help citizens to share

135    information as a group,[23] benefiting from local knowledge and experience within a common

136    platform. He envisioned the platform as a non-commercial resource that would span political

137    boundaries and help expand knowledge.[23] The crowdsourced encyclopedia, Wikipedia, echoes

138    some of the structures and functions of Wells' original world brain concept.

139

140    Empirical evidence from humans suggests that there is a convergent collective intelligence factor

141    that explains a group's performance on several tasks.[24] Further empirical evidence supporting

142    collective intelligence is summarized in James Surowiecki's *The Wisdom of Crowds*.[25] He argues

143    that four elements are necessary for collective intelligence – diversity of opinion, independence

144    of individual ideas, decentralization of ideas, and a way to aggregate individual ideas.

145    Surowiecki shows how collective intelligence has been used in many different contexts, ranging

146    from prediction markets to the Delphi method. The Delphi method has a group of individuals

147    iteratively answer questions and converge on a single answer. The method has been widely used

148    to achieve group consensus in health guidelines.[26]

149

150    Collective intelligence approaches have been evaluated in several medical settings. Research

151    among medical students suggests that groups of medical students have increased diagnostic

152    performance compared to individual medical students.[27,28] Similar approaches have been

153    evaluated in the context of physician diagnosis of skin cancer[29] and breast cancer.[30]

154

155    **Open Source Model**

156     Open source models can inform the second important component of crowdsourcing – sharing the

157     solution with the public. Open source refers to a decentralized structure that facilitates

158     collaboration and online sharing. Open source models were developed in the 1960s and 1970s as

159     a way to collaboratively develop software and share code.[31] In 1969, the United States Advanced

160     Research Project Agency created the first large, high-speed computer network. This extended

161     opportunities for sharing code among broader online groups. For example, the Linux operating

162     system is one of the first open source operating systems, shared online and available for free to

163     anyone. Linus Torvalds developed the source code for this operating system by sending it to

164     other internet users who helped improve it on a volunteer basis. The collective development of

165     open source products, such as Linux, demonstrate how large, diverse groups working together

166     can iteratively enhance a product that is openly available, to the benefit of all.

167

168     This trend also led to the development of Creative Commons, a non-profit organization that

169     allows individuals to legally change and share creative works. The organization has a series of

170     copyright licenses that clarify the terms of sharing. There are currently approximately 1.4 billion

171     works that have been licensed through Creative Commons.

172

173     Open source models have increasingly appeared in medicine. For example, several drug

174     development projects have used open source models.[32-34] A project called Open Source Drug

175     Discovery focuses developing drugs for neglected tropical diseases through open source

176     methods.[35] Thousands of volunteers from over 100 countries have helped with micro-tasks to

177     develop more effective drugs for tuberculosis, schistosomiasis, and other infectious diseases.[36]

178    The open source platform has resulted in high-quality research, including advances related to the

179    development of schistosomiasis drugs.[37] Other open source models for drug discovery have been

180    developed for Huntington's disease,[38] malaria,[39] and other diseases.[35]

181

182    Open source models have also been used within genomics. A Shiga-toxin producing *E. coli*

183    outbreak occurred in Germany in 2011, infecting 3000 individuals. Scientists used an open

184    source model to organize the analysis of a genome sequence from a single individual. The

185    collaborative effort brought together volunteers from around the world, creating the genome

186    sequence within two weeks of receiving the DNA samples.[40] In addition, the DREAMS

187    Challenge team has organized many open source challenge contests.[41] These typically involve

188    volunteers collaboratively working together to solve a problem related to big data and genomics.

189    Several evaluations of this approach have found it to be effective in developing prognostic

190    models based on clinical data.[42-44] Both collective intelligence and open source models reveal

191    some of the theoretical antecedents of crowdsourcing.

192

193    **Relationship to Other Research Approaches**

194    Crowdsourcing as an approach is distinct from, but related to community-based participatory

195    research, participatory action research, and community-driven research.  Each of these different

196    approaches has a theoretical framework, methods, and assumptions. At the same time, each of

197    these three approaches can be used to inform crowdsourcing research and programs.

198

199    Community-based participatory research actively engages the community in all stages of the

200    research process, contributing to shared decision making and community ownership.[45] The

201    community plays a central force in setting the agenda, implementing the study, and evaluating

202    the results, such that local community members and researchers iteratively collaborate to

203    improve the health of the community. Similarities between community-based participatory

204    research and crowdsourcing include the following: a focus on listening to and partnering with

205    local communities; a potential to increase healthy equity; an acknowledgement that communities

206    can be a powerful source of new ideas. These areas of convergence suggest that community-

207    based participatory research could be a useful complement to crowdsourcing. For example,

208    community-based participatory research was used to increase community engagement in an HIV

209    cure research project.[46]

210

211    Other related approaches include participatory action research and youth participatory action

212    research. Participatory action research focuses on partnering with communities to participate in

213    research and achieve social change.[47] Youth participatory action research provides youth with

214    opportunities to learn about social problems that affect their lives and then propose actions to

215    address these problems.[48-51] The participatory action approach considers youth as potential

216    experts and co-creators of knowledge.[52] Shared elements of crowdsourcing and participatory

217    research approaches include the emphasis on participation, local community partnerships, and

218    empowerment of non-experts. Participatory action research has been used to complement

219    crowdsourcing projects related to environmental health[53] and to design crowdsourcing

220    approaches for HIV self-testing.[54]

221

222    Finally, community-driven research is another approach related to crowdsourcing. Community-

223    driven research has community members and researchers collaboratively design, implement,

224    analyze, interpret, and disseminate research findings.[55] Community-driven research starts with an

225    assessment of local priorities from the perspective of the community. Both community-driven

226    research and crowdsourcing focus on community-led research, developing ideas and programs

227    from the bottom-up for the community.[56]  All three of these approaches have been used in health

228    research. We now turn to examine crowdsourcing specifically in the context of health.

229

230    **Crowdsource-able Challenges in Health**

231    There is evidence demonstrating that some health challenges are more likely to produce wise

232    groups that others. These specific settings are instances in which crowdsourcing would likely be

233    more useful and settings where it would be less useful.[2] Issues that are more amenable to

234    crowdsourcing approaches include specific behavioral or social issues (e.g., changing condom

235    use behaviors), topics that have champions and are timely, topics with robust ally networks, and

236    topics where the public has a responsibility to be engaged. Given the limited research literature

237    on crowdsourcing in health, the evidence should be interpreted with caution. At the same time,

238    lessons learned from previous examples of successful crowdsourcing for addressing health-

239    related challenges may help individuals or groups to develop additional applications for this

240    approach. We will now examine the four areas where crowdsourcing is more useful - specific

241    behavioral or social issues, topics that have champions and are timely, topics with robust ally

242    networks, and topics where the public has a responsibility to be engaged.

243

244  First, behavioral and social science-related health challenges are more conducive to a

245  crowdsourcing approach compared to biomedical ones. Systematic reviews[5,57] and a World

246  Health Organization practical guide on crowdsourcing in health and health research[2] suggest that

247  crowdsourcing is used more frequently in the context of behavioral science. Having input from

248  community groups may be particularly useful in cases where social relationships or related

249  behavioral issues are at stake. For example, a challenge related to promoting HIV testing

250  behaviors may be more feasible to crowdsource compared to a challenge related to developing

251  new clinical HIV drugs. More specifically, within the broad fields of behavioral and social

252  sciences, many crowdsourcing projects have asked groups to develop health communications

253  (e.g., images, videos) or marketing tools.[11,14,58]

254

255  Second, topics that have champions and are sensitive to local communications trends are more

256  readily taken on by crowdsourcing. Here were refer to champions as individuals who represent a

257  cause and are willing to openly endorse a project. In the context of a social media challenge

258  contest, this is the extent to which the call for submissions is forwarded. For example, our

259  women leaders in global health call for suggestions on increasing women's participation in a

260  WHO/TDR mid-career fellowship was highly forwarded. We speculate that this was related to

261  the increasing number of women champions of this cause and a favorable social media

262  environment. Ultimately the contest received 311 submissions from 65 countries.

263

264  Third, groups are more likely to be wise in the context of strong partnerships and networks. The

265  nature of crowdsourcing depends on a call for widespread participation to be circulated widely.

266    These networks could be online or in-person, but the steering committee needs to be able to

267    activate the networks. Networks could include student groups, patient advocacy organizations,

268    professional associations, social media groups, or other groups. With broader networks and

269    stronger partnerships, a crowdsourcing contest is more likely to reach creative people who are

270    willing to contribute, resulting in more high-quality entries.

271

272    Fourth, crowdsourcing is particularly useful in settings where the public has a responsibility to

273    be engaged. The rationale for public engagement varies widely in research studies. Some studies

274    actively seek public input and opinion,[46,59] while others discourage it. Given the strong force of

275    the public in crowdsourcing, this is an important issue.

276

277    **Critiques of Crowdsourcing**

278    There are three main critiques of crowdsourcing that merit consideration – the madness of groups

279    concept, the problem of low-quality submissions, and cognitive fixation on examples. We will

280    examine each of these critiques generally and then in the context of crowdsourcing as it applies

281    to medicine.

282

283    First, the madness of groups refers to the potential for groups to create and disseminate popular

284    delusions, contributing to panic and moral outrage.[60] The nineteenth century journalist Charles

285    Mackay remarked, "Men, it has been said, think in herds; it will be seen that they go made in

286    herds, while they only recover their senses slowly, and one by one." Psychologists have

287    examined how individual behaviors contribute to and diverge from the collective behavior of the

288    groups. Group behavior may be associated with a loss of responsibility. This is illustrated in the

289    case of Boaty McBoatface, a shipping vessel name that received the largest number of public

290    votes.[61] However, crowdsourcing as an approach does not suggest that all groups are wise at all

291    times, but rather that there are specific conditions that can allow for wise groups. In addition,

292    several individuals have made rebuttals and clarified the concept of a mad group. Clark McPhail

293    has shown how mad groups are primarily the result of individuals, rather than a group

294    disposition.[62] Empirical data on whether group behavior results in a loss of responsibility has

295    been mixed.[63] Within the context of medicine, unmoderated online platforms could help to

296    propagate myths or misunderstandings about a disease.[64] Yet appropriate moderation of online

297    platforms in the context of health interventions decreases this risk.[65] Hence, the risk of mad

298    groups in health interventions appears limited.

299

300    Second, crowdsourcing projects are sometimes associated with many low-quality outputs. A

301    systematic review of crowdsourcing suggests that only a subset of outputs are excellent[1]. Having

302    non-experts contribute to a more complex medical project will result in a wide range of outputs,

303    especially when mass engagement translates into hundreds of submissions. However, the ability

304    to prompt a large number of submissions is an advantage of crowdsourcing and suggests that a

305    wider group of individuals is actively participating. Several techniques for judging have been

306    developed to assess large numbers of crowdsourcing contributions,[2] including group judging

307    (having a group of non-experts evaluate),[11] panel judging (having a diverse group of individuals

308    evaluate),[14] and artificial intelligence.[66,67]  Several systematic reviews of crowdsourcing in

309  medicine suggest that crowdsourcing allows a broad range of quality, including both low and

310  high-quality submissions.[5,57,68]

311

312  Finally, the problem of cognitive fixation on prior ideas has been described in crowdsourcing.[69]

313  This refers to the phenomenon when providing an example or reference limits the diversity of

314  ideas solicited. This concept is similar to groupthink, which occurs when a group of individuals

315  converges on a single solution.[70] There are several technical ways of designing a crowdsourcing

316  project that could limit cognitive fixation, including the following: limiting the use of examples

317  when calling for innovative ideas; drawing on different groups of individuals or different topics

318  (avoiding serial contests focused on the same topic); and having a submission system in which

319  those who submit do not view other submissions.

320

321  **Crowdsourcing Applications in Health**

322  Crowdsourcing approaches have already been applied to health research, programs, and policy

323  (Table 2). From a research perspective, crowdsourcing approaches have recruited research

324  participants, contributed to systematic reviews, and informed clinical trials. Many studies have

325  used Amazon Turk or other platforms to recruit study participants into online randomized

326  controlled trials.[7,71-73] While such approaches are often rapid and save money, there are concerns

327  about generalizability.[74] Crowdsourcing has been used to identify potentially relevant citations as

328  part of systematic reviews. This approach has been found reliable[75] and is being piloted as part of

329  a Cochrane program.[76] Finally, crowdsourcing has been used as a tool to inform clinical trials.

330  Several studies have used crowdsourcing to develop study names, logos, and related materials.[59]

331    Crowdsourcing has also helped to tailor health programs and make them more people-centered.

332    Crowdsourcing approaches may localize health program materials, increasing community uptake

333    and contributing to subsequent scale-up. Crowdsourcing has been used to develop locally

334    responsive health communication materials (e.g., images, videos, slogans).[1] Finally,

335    crowdsourcing has several potential applications for health policy. The open, transparent nature

336    of crowdsourcing contests could increase trust and facilitate community feedback on future

337    health policies. Open calls for community feedback have helped to inform youth HIV policy[77]

338    and WHO hepatitis testing guidelines.[78] More research and programs are needed to build out the

339    evidence base supporting crowdsourcing approaches.

340

341

342

343

344

345

346

347

348

349

350

**Table 1. Crowdsourcing activities used to improve health: structure and function**

| Crowdsourcing activity | Structure | Function |
|---|---|---|
| Challenge contests | Open solicitation and promotion to the public for contest submissions; evaluation, celebration, and sharing of contest submissions | Generate innovative ideas, logos, images, or videos (e.g., images to increase HIV testing,[7,14,79] strategies to promote hepatitis testing[78]) |
| Hackathons | Short (often 3 days) event that brings together individuals around a common cause | Design a technological solution or a health service (e.g., design an HIV testing service[80]) |
| Online collaboration systems | Websites or portals that allow individuals to solve a problem | Solve micro-tasks for a small amount of money (e.g., evaluation of surgical skills[68]) |

351

352

**Table 2. Crowdsourcing applications in health research, programs, and policy**

| Crowdsourcing Application | Purpose of Crowdsourcing | Examples |
|---|---|---|
| Designing health research studies | Recruiting study participants; undertaking formative research for clinical trials | Online research studies;[81] assist with systematic reviews;[75] solicit community feedback on HIV cure research[46,82] |
| Developing health programs | Tailor health programs to be more people-focused and inclusive | Developing logos, taglines, or videos for community health programs[14] |
| Informing health policy | Directing health policy towards community-identified needs | Soliciting youth input on HIV policy;[77] soliciting hepatitis cases for global guidelines[78] |

353

354

355

356

357

358

359

## References

361

362 1.  Pan SW, Stein G, Bayus B, et al. Systematic review of design contests for health:
363 Spurring innovation and mass engagement. *BMJ Innovations* 2017.
364 2.  Han L, Chen A, Wei S, Ong JJ, Iwelunmor J, Tucker JD. Crowdsourcing Contests in
365 Health and Health Research: A Practical Guide. Geneva: World Health Organization, 2018.
366 3.  van Ess H. Crowdsourcing: How to find a crowd. 2010.
367 https://www.slideshare.net/searchbistro/harvesting-knowledge-how-to-crowdsource-in-2010.
368 4.  Brabham DC, Ribisl KM, Kirchner TR, Bernhardt JM. Crowdsourcing applications for
369 public health. *Am J Prev Med* 2014; **46**(2): 179-87.
370 5.  Ranard BL, Ha YP, Meisel ZF, et al. Crowdsourcing--harnessing the masses to advance
371 health and medicine, a systematic review. *J Gen Intern Med* 2014; **29**(1): 187-203.
372 6.  Ringh M, Rosenqvist M, Hollenberg J, et al. Mobile-phone dispatch of laypersons for
373 CPR in out-of-hospital cardiac arrest. *N Engl J Med* 2015; **372**(24): 2316-25.
374 7.  Tang W, Han L, Best J, et al. Crowdsourcing HIV Testing: A Pragmatic, Non-Inferiority
375 Randomized Controlled Trial in China. *Clinical Infectious Diseases* 2016; **62**: 1436-42.
376 8.  Geiger D, Rosemann M, Fielt E. Crowdsourcing information systems: a systems theory
377 perspective.  22nd Australian Conference on Information Systems; 2011; Sydney, Australia;
378 2011.
379 9.  Howe J. The Rise of Crowdsourcing. Wired. 2006.
380 10.  And the winner is...: Capturing the promise of philanthropic prizes. 2009.
381 http://mckinseyonsociety.com/downloads/reports/Social-Innovation/And_the_winner_is.pdf.
382 11.  Tang W, Wei C, Cao B, et al. Crowdsourcing to expand HIV testing among men who
383 have sex with men in China: A closed cohort stepped wedge cluster randomized controlled trial.
384 *PLoS Med* 2018; **15**(8): e1002645.
385 12.  Wu D, Best LL, Stein G, Tang W, Tucker JD, Healthy Cities Contest T. Community
386 participation in a Lancet Healthy Cities in China Commission. *Lancet Planet Health* 2018; **2**(6):
387 e241-e2.
388 13.  Tucker JD, Pan SW, Mathews A, Stein G, Bayus B, Rennie S. Ethical Concerns of and
389 Risk Mitigation Strategies for Crowdsourcing Contests and Innovation Challenges: Scoping
390 Review. *Journal of medical Internet research* 2018; **20**(3): e75.
391 14.  Zhang Y, Kim J, Liu F, et al. Creative contributory contests (CCC) to spur innovation in
392 sexual health: Two cases and a guide for implementation. *Sexually transmitted diseases* 2015;
393 **42**: 625-8.
394 15.  Tang W, Mao J, Liu C, et al. Reimagining Health Communication: A Non-Inferiority
395 Randomized Controlled Trial of Crowdsourcing in China. *Sex Trans Dis* 2016.

396  16.    Zhang A, Pan X, Wu F, et al. What Would an HIV Cure Mean to You: Qualitative
397  Analysis from a Crowdsourcing Contest in Guangzhou, China. *AIDS Res Hum Retroviruses*
398  2017.
399  17.    de Condorcet M. Essai sur l'application de l'analyse à la probabilité des décisions rendues
400  à la pluralité des voix. 1785. http://gallica.bnf.fr/ark:/12148/bpt6k417181.
401  18.    Ladha KK. The Condorcet Jury Theorem, Free Speech, and Correlated Votes. *American*
402  *Journal of Political Science* 1992; **36**(3): 617-34.
403  19.    Austen-Smith D, Banks JS. Information aggregation, rationality, and the Condorcet Jury
404  Theorem. *American Political Science Review* 1996; **90**(1): 34-45.
405  20.    Gottlieb K, Hussain F. Voting for image scoring and assessment (VISA)--theory and
406  application of a 2 + 1 reader algorithm to improve accuracy of imaging endpoints in clinical
407  trials. *BMC Med Imaging* 2015; **15**: 6.
408  21.    Koch T, Ridgley M. The Condorcet's Jury Theorem in a Bioethical Context: The
409  Dynamics of Group Decision Making. *Group Decision and Negotiation* 2000; **9**: 379-92.
410  22.    Lévy P. Collective intelligence : mankind's emerging world in cyberspace. New York:
411  Plenum Trade; 1997.
412  23.    Wells HG. World brain. Garden City, N.Y.,: Doubleday, Doran & Co., Inc.; 1938.
413  24.    Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a
414  Collective Intelligence Factor in the Performance of Human Groups. *Science* 2010; **330**: 686-8.
415  25.    Surowiecki J. The wisdom of crowds : why the many are smarter than the few and how
416  collective wisdom shapes business, economies, societies, and nations. 1st ed. New York:
417  Doubleday :; 2004.
418  26.    Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review
419  recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol* 2014; **67**(4):
420  401-9.
421  27.    Hautz WE, Kammer JE, Schauber SK, Spies CD, Gaissmaier W. Diagnostic performance
422  by medical students working individually or in teams. *JAMA* 2015; **313**(3): 303-4.
423  28.    Kammer JE, Hautz WE, Herzog SM, Kunina-Habenicht O, Kurvers R. The Potential of
424  Collective Intelligence in Emergency Medicine: Pooling Medical Students' Independent
425  Decisions Improves Diagnostic Performance. *Med Decis Making* 2017; **37**(6): 715-24.
426  29.    Kurvers RH, Herzog SM, Hertwig R, et al. Boosting medical diagnostics by pooling
427  independent judgments. *Proc Natl Acad Sci U S A* 2016; **113**(31): 8777-82.
428  30.    Wolf M, Krause J, Carney PA, Bogart A, Kurvers RH. Collective Intelligence Meets
429  Medical Decision-Making: The Collective Outperforms the Best Radiologist. *PLoS One* 2015;
430  **10**(8): e0134269.
431  31.    von Hippel E, von Krogh G. Open Source Software and the "Private-Collective"
432  Innovation Model: Issues for Organization Science. *Organization Science* 2003; **14**: 209-23.
433  32.    Munos B. Can open-source R&D reinvigorate drug research? *Nat Rev Drug Discov* 2006;
434  **5**(9): 723-9.
435  33.    Munos B. Can open-source drug R&D repower pharmaceutical innovation? *Clin*
436  *Pharmacol Ther* 2010; **87**(5): 534-6.
437  34.    So AD, Gupta N, Brahmachari SK, et al. Towards new business models for R&D for
438  novel antibiotics. *Drug Resist Updat* 2011; **14**(2): 88-94.
439  35.    Bagla P. Science in India. Crowd-sourcing drug discovery. *Science* 2012; **335**(6071):
440  909.

441  36.    Bhardwaj A, Scaria V, Raghava GP, et al. Open source drug discovery--a new paradigm
442  of collaborative research in tuberculosis drug development. *Tuberculosis (Edinb)* 2011; **91**(5):
443  479-86.
444  37.    Ardal C, Rottingen JA. Open source drug discovery in practice: a case study. *PLoS Negl*
445  *Trop Dis* 2012; **6**(9): e1827.
446  38.    Wilhelm M. Big Pharma Buys Into Crowdsourcing for Drug Discovery. Wired. 2017.
447  39.    Ardal C, Rottingen JA. An open source business model for malaria. *PLoS One* 2015;
448  **10**(2): e0117150.
449  40.    Rohde H, Qin J, Cui Y, et al. Open-source genomic analysis of Shiga-toxin-producing E.
450  coli O104:H4. *N Engl J Med* 2011; **365**(8): 718-24.
451  41.    Saez-Rodriguez J, Costello JC, Friend SH, et al. Crowdsourcing biomedical research:
452  leveraging communities as innovation engines. *Nat Rev Genet* 2016; **17**(8): 470-86.
453  42.    Guinney J, Wang T, Laajala TD, et al. Prediction of overall survival for patients with
454  metastatic castration-resistant prostate cancer: development of a prognostic model through a
455  crowdsourced challenge with open clinical trial data. *Lancet Oncol* 2017; **18**(1): 132-42.
456  43.    Allen GI, Amoroso N, Anghel C, et al. Crowdsourced estimation of cognitive decline and
457  resilience in Alzheimer's disease. *Alzheimers Dement* 2016; **12**(6): 645-53.
458  44.    Noren DP, Long BL, Norel R, et al. A Crowdsourcing Approach to Developing and
459  Assessing Prediction Algorithms for AML Prognosis. *PLoS Comput Biol* 2016; **12**(6): e1004890.
460  45.    Minkler M, Wallerstein N. Community based participatory research for health. San
461  Francisco, CA: Jossey-Bass; 2003.
462  46.    Mathews A, Farley S, Hightow-Weidman L, Muessig K, Rennie S, Tucker JD.
463  Crowdsourcing and community engagement: a qualitative analysis of the 2BeatHIV contest. *J*
464  *Virus Erad* 2018; **4**(1): 30-6.
465  47.    Bradbury H. The SAGE handbook of action research / edited by Hilary Bradbury. Third
466  edition ed. Los Angeles: SAGE Publications; 2015.
467  48.    Cammarota J, Fine M. Revolutionizing education : youth participatory action research in
468  motion. New York, NY: Routledge; 2008.
469  49.    Kirshner B. Productive Tensions in Youth Participatory Action Research. *Yearbook of*
470  *the National Society for the Study of Education* 2010; **109**(1): 238-51.
471  50.    Cammarota J, Fine M. Revolutionizing education: Youth participatory action research in
472  motion: Routledge; 2010.
473  51.    Ozer EJ, Piatt AA, Holsen I, Larsen T, Lester J, Ozer EM. INNOVATIVE
474  APPROACHES TO PROMOTING POSITIVE YOUTH DEVELOPMENT IN DIVERSE
475  CONTEXTS. *Positive Youth Development in Global Contexts of Social and Economic Change*
476  2016: 12.
477  52.    Ozer EJ. Chapter Seven-Youth-Led Participatory Action Research: Developmental and
478  Equity Perspectives. *Advances in child development and behavior* 2016; **50**: 189-207.
479  53.    English PB, Richardson MJ, Garzon-Galvis C. From Crowdsourcing to Extreme Citizen
480  Science: Participatory Research for Environmental Health. *Annu Rev Public Health* 2018; **39**:
481  335-50.
482  54.    ITEST: Innovative Tools to Expand Youth-Friendly HIV Self-Testing. 2018.
483  https://projectreporter.nih.gov/project_info_description.cfm?aid=9618360&icde=41536786&ddp
484  aram=&ddvalue=&ddsub=&cr=1&csb=default&cs=ASC&pball=.

485  55.     Orionzi DE, Mink PJ, Azzahir A, et al. Implementing a Community-Driven Research
486  Partnership: The Backyard Initiative Community Health Survey Methods and Approach. *Prog*
487  *Community Health Partnersh* 2016; **10**(4): 493-503.
488  56.     McElfish PA, Kohler P, Smith C, et al. Community-Driven Research Agenda to Reduce
489  Health Disparities. *Clin Transl Sci* 2015; **8**(6): 690-5.
490  57.     Crequit P, Mansouri G, Benchoufi M, Vivot A, Ravaud P. Mapping of Crowdsourcing in
491  Health: Systematic Review. *Journal of medical Internet research* 2018; **20**(5): e187.
492  58.     Zhang W, Schaffer D, Tso LS, et al. Innovation contests to promote sexual health in
493  china: a qualitative evaluation. *BMC Public Health* 2017; **17**(1): 78.
494  59.     Mathews A, Farley S, Blumberg M, et al. HIV cure research community engagement in
495  North Carolina: a mixed-methods evaluation of a crowdsourcing contest. *J Virus Erad* 2017;
496  **3**(4): 223-8.
497  60.     Mackay C. Memoirs of extraordinary popular delusions and the madness of crowds. 2d
498  ed. London: Office of the National Illustrated Library; 1852.
499  61.     Ellis-Petersen H. Boaty McBoatface wins poll to name polar research vessel. Guardian.
500  2016.
501  62.     McPhail C. The Myth of the Madding Crowd. *Social Forces* 1991.
502  63.     Manstead ASR, Hewstone M. The Blackwell encyclopedia of social psychology. Oxford,
503  UK ; Cambridge, Mass., USA: Blackwell; 1995.
504  64.     Lavorgna L, Lanzillo R, Brescia Morra V, Abbadessa G, Tedeschi G, Bonavita S. Social
505  Media and Multiple Sclerosis in the Posttruth Age. *Interact J Med Res* 2017; **6**(2): e18.
506  65.     Rice S, Robinson J, Bendall S, et al. Online and Social Media Suicide Prevention
507  Interventions for Young People: A Focus on Implementation and Moderation. *J Can Acad Child*
508  *Adolesc Psychiatry* 2016; **25**(2): 80-6.
509  66.     Mudie LI, Wang X, Friedman DS, Brady CJ. Crowdsourcing and Automated Retinal
510  Image Analysis for Diabetic Retinopathy. *Curr Diab Rep* 2017; **17**(11): 106.
511  67.     Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S, Navab N. AggNet: Deep
512  Learning From Crowds for Mitosis Detection in Breast Cancer Histology Images. *IEEE Trans*
513  *Med Imaging* 2016; **35**(5): 1313-21.
514  68.     Dai JC, Lendvay TS, Sorensen MD. Crowdsourcing in Surgical Skills Acquisition: A
515  Developing Technology in Surgical Education. *J Grad Med Educ* 2017; **9**(6): 697-705.
516  69.     Fu S, de Vreede G-J, Cheng X, Seeber I, Maier R, Weber B. Convergence of
517  Crowdsourcing Ideas: A Cognitive Load perspective.  ICIS 2017 PROCEEDINGS Association
518  for Information Systems; 2017; 2017.
519  70.     Janis IL. Victims of groupthink; a psychological study of foreign-policy decisions and
520  fiascoes. Boston,: Houghton; 1972.
521  71.     Tang W, Mao J, Liu C, et al. Crowdsourcing health communication about condom use in
522  men who have sex with men in China: a randomised controlled trial. *Lancet* 2016; **388 (Suppl**
523  **1)**: S73.
524  72.     Losina E, Michl GL, Smith KC, Katz JN. Randomized Controlled Trial of an Educational
525  Intervention Using an Online Risk Calculator for Knee Osteoarthritis: Effect on Risk Perception.
526  *Arthritis Care Res (Hoboken)* 2017; **69**(8): 1164-70.
527  73.     Jones RB, Goldsmith L, Hewson P, Williams CJ. Recruitment to online therapies for
528  depression: pilot cluster randomized controlled trial. *J Med Internet Res* 2013; **15**(3): e45.
529  74.     Wang C, Mollan KR, Hudgens MG, et al. Generalisability of an online randomised
530  controlled trial: an empirical analysis. *J Epidemiol Community Health* 2017.

531    75.      Mortensen ML, Adam GP, Trikalinos TA, Kraska T, Wallace BC. An exploration of
532    crowdsourcing citation screening for systematic reviews. *Res Synth Methods* 2017; **8**(3): 366-86.
533    76.      Cochrane Crowd. 2018. http://community.cochrane.org/tools/data-management-
534    tools/cochrane-crowd.
535    77.      Hildebrand M, Ahumada C, Watson S. CrowdOutAIDS: crowdsourcing youth
536    perspectives for action. *Reprod Health Matters* 2013; **21**(41): 57-68.
537    78.      Tucker JD, Meyers K, Best J, et al. The HepTestContest: a global innovation contest to
538    identify approaches to hepatitis B and C testing. *BMC infectious diseases* 2017; **17**(Suppl 1):
539    701.
540    79.      Group SS, Tucker JD. Crowdsourcing to promote HIV testing among MSM in China:
541    study protocol for a stepped wedge randomized controlled trial. *Trials* 2017; **18**(1): 447.
542    80.      Tucker JD, Tang W, Li H, et al. Crowdsourcing Designathon: A New Model for
543    Multisectoral Collaboration. *BMJ Innovations* 2018.
544    81.      Bartneck C, Duenser A, Moltchanova E, Zawieska K. Comparing the similarity of
545    responses received from studies in Amazon's Mechanical Turk to studies conducted online and
546    with direct recruitment. *PLoS One* 2015; **10**(4): e0121595.
547    82.      Zhang A, Pan X, Wu F, et al. What Would an HIV Cure Mean to You: Qualitative
548    Analysis from a Crowdsourcing Contest in Guangzhou, China. *AIDS Res Hum Retroviruses*
549    2018; **34**: 80-7.

550

# Table 1(on next page)

Table 1 and Table 2

**Table 1. Crowdsourcing activities used to improve health: structure and function**

| Crowdsourcing activity | Structure | Function |
|---|---|---|
| Challenge contests | Open solicitation and promotion to the public for contest submissions; evaluation, celebration, and sharing of contest submissions | Generate innovative ideas, logos, images, or videos (e.g., images to increase HIV testing,[7,14,79] strategies to promote hepatitis testing[78]) |
| Hackathons | Short (often 3 days) event that brings together individuals around a common cause | Design a technological solution or a health service (e.g., design an HIV testing service[80]) |
| Online collaboration systems | Websites or portals that allow individuals to solve a problem | Solve micro-tasks for a small amount of money (e.g., evaluation of surgical skills[68]) |

1

2

**Table 2. Crowdsourcing applications in health research, programs, and policy**

| Crowdsourcing Application | Purpose of Crowdsourcing | Examples |
|---|---|---|
| Designing health research studies | Recruiting study participants; undertaking formative research for clinical trials | Online research studies;[81] assist with systematic reviews;[75] solicit community feedback on HIV cure research[46,82] |
| Developing health programs | Tailor health programs to be more people-focused and inclusive | Developing logos, taglines, or videos for community health programs[14] |
| Informing health policy | Directing health policy towards community-identified needs | Soliciting youth input on HIV policy;[77] soliciting hepatitis cases for global guidelines[78] |

3

4