Journal of HOSPITAL MEDICINE www.journalofhospitalmedicine.com

BRIEF REPORTS

Crowdsourcing Medical Expertise in Near Real Time

Max H. Sims[1], Jeffrey Bigham, PhD[2], Henry Kautz, PhD[3], Marc W. Halterman, MD, PhD[1]*

[1]Department of Neurology, University of Rochester Medical Center, Rochester, New York; [2]Human Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania; [3]Department of Computer Science, University of Rochester Medical Center, Rochester, New York.

Given the pace of discovery in medicine, accessing the literature to make informed decisions at the point of care has become increasingly difficult. Although the Internet creates unprecedented access to information, gaps in the medical literature and inefficient searches often leave healthcare providers' questions unanswered. Advances in social computation and human computer interactions offer a potential solution to this problem. We developed and piloted the mobile application DocCHIRP, which uses a system of point-to-multipoint push notifications designed to help providers problem solve by crowdsourcing from their peers. Over the 244-day pilot period, 85 registered users logged 1544 page views and sent 45 consult questions. The median initial first response from the crowd occurred within 19 minutes. Review of the transcripts revealed several dominant themes, including complex medical decision making and inquiries related to prescription medication use. Feedback from the post-trial survey identified potential hurdles related to medical crowdsourcing, including a reluctance to expose personal knowledge gaps and the potential risk for "distracted doctoring." Users also suggested program modifications that could support future adoption, including changes to the mobile interface and mechanisms that could expand the crowd of participating healthcare providers. *Journal of Hospital Medicine* 2014;9:451–456. © 2014 Society of Hospital Medicine

The volume of existing knowledge and the pace of discovery in medical science challenge a clinician's ability to access relevant information at the point of care. Knowledge gaps that arise in practice usually involve matters related to diagnosis, drug therapy, or treatment.[1] In the clinical setting, healthcare providers (HCPs) answer questions using a variety of online and print resources. Ironically, HCPs often lack the training required to find details regarding uncommon disorders or complex medical decisions that are not easily found or well represented in the published literature.[2] Instead, HCPs turn to trusted colleagues who possess the necessary expertise.[3]

Closing the "knowledge-to-practice gap" involves a range of factual information and data derived from published evidence, anecdotal experience, as well as organization- and region-specific practices.[4] The inability to codify both explicit and tacit information has been linked to variability in prescription practices, excessive use of surgical services, and delayed decisions involving the appropriate provision of end-of-life care.[5] Although electronic medical record systems are not configured to support peer collaboration,[6] alternative strategies including crowdsourcing has been used successfully in other domains to tap collective intelligence of skilled workers.[7] Crowdsourcing allows organizations to explore problems at low cost, gain access a wide range of complementary expertise, and capture large amounts of data for analysis.[8,9] Although an increasing number of physicians use either smartphones or tablets on the job,[10] peer-to-peer medical crowdsourcing has not been investigated, despite the fact that processes involving team-based clinical decision making are associated with better outcomes.[11] Here we field tested the mobile crowdsourcing application DocCHIRP (Crowdsourcing Health Information Retrieval Protocol for Doctors) and assessed user opinion regarding its utility in the clinical setting.

MATERIALS AND METHODS

DocCHIRP Program Design

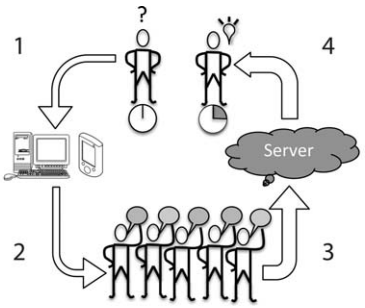
The authors (M.W.H., J.B., H.K.) conceptualized and designed DocCHIRP for mobile (iOS [Apple Inc., Cupertino, CA] and Android [Google Inc., Mountain View, CA]) and desktop use. Email prompts and push notifications, which were modeled after the application VizWiz (Rochester Human Computer Interaction Group, University of Rochester, Rochester, NY), supported near real-time communication between HCPs. According to recent US Food and Drug Administration guidelines, DocCHIRP is considered a medical reference,[12] intended to share domain-specific knowledge on diagnosis, therapy, and other medically relevant topics. Devices were password protected and encrypted according to university standards. A typical workflow involves an index provider faced with a clinical question that sends a consult question to 1 or more trusted providers. The crowd receiving the notification responds when available using either free-text

*Address for correspondence and reprint requests: Marc W. Halterman, MD, Department of Neurology, Center for Neural Development & Disease, University of Rochester Medical Center, 601 Elmwood Avenue, Box 645, Rochester, NY 14642; Telephone: 585-273-1335; Fax: 585-276-1947; E-mail: marc_halterman@urmc.rochester.edu

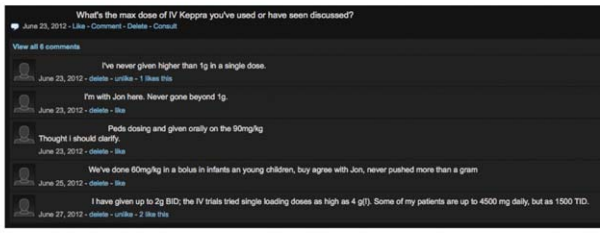
Additional Supporting Information may be found in the online version of this article.

Received: February 12, 2014; **Revised:** March 20, 2014;
Accepted: March 30, 2014
2014 Society of Hospital Medicine DOI 10.1002/jhm.2204
Published online in Wiley Online Library (Wileyonlinelibrary.com).

A



B



C

Device	# Visits	Pages Per Visit	Average Visit Duration (Min:Sec)
Tablet	147	8.6	4:34
Desktop	519	16.80	13:10
Mobile	878	5.5	5:15
Total	1544		

Mobile Breakdown	Visits by Device	Pages/Device	Average Visit Duration (Min:Sec)
Apple iPod Touch	30	3.8	3:37
Apple iPhone	838	5.5	5:14
Apple iPad	147	8.6	4:34
Droid	4	5.5	7:04
Unknown	6	7.44	10:28
Total	1025		

FIG. 1. Architecture of the DocCHIRP platform. (A) Schematic of the DocCHIRP workflow. The provider formulates the initial consult (1) and sends the information request to the crowd using either a mobile device at the point of care or Web interface on a desktop computer. (2) The crowd is selected based on provider preferences, receives consult, and replies if they possess the necessary expertise and are available to respond. (3) DocCHIRP captures feedback from the cloud consultants (4) and returns the data to the index provider in near real time. (B) Screen shot of the user interface. Discussion threads are time stamped and clustered with the initial consult question. Users can respond with a free-text reply or simply vote on the comment. In this example, the headshots and names of the field trial participants have been edited to preserve anonymity. (C) Analysis of the devices used to engage the DocCHIRP server and information regarding server time grouped by device type. Abbreviations: BID, twice daily; IV, intravenous.

responses or agree/disagree prompts (Figure 1A,B). Providers use preference settings to manage crowd membership, notification settings, and demographics describing their expertise.

Trial Recruitment

The University of Rochester Research Subjects Review Board approved the study, in which prospective users were required to review and agree to a statement regarding potential liability as part of the consent process. In this pilot study, we invited a cross-section of providers (n = 145) from the Departments of Neurology (including the Division of Pediatric Neurology), Pediatrics, Neuroradiology, Psychiatry, Orthopedics, Emergency Medicine, Internal Medicine, and Family Medicine to participate. E-mail invitations were sent to HCPs in 3 phases in April (phase I), June (phase II), and August (phase III) over 244 consecutive days. At the conclusion of the trial, 85 HCPs (59%) had created accounts including attending physicians (n = 63), residents (n = 13), fellows (n = 1), and nurse practitioners (n = 8). We did not seek parity in either age or gender representation.

Data Analysis

Mobile device and network usage data, question and response strings, as well as data regarding hardware and browser identity were collected using Google Analytics (Google Inc., http://www.google.com/analytics), and discussion threads were recovered from the DocCHIRP user logs. After the trial was completed, we invited participants to complete a 10-minute, anonymous, online survey consisting of 21 open- and closed-ended questions (www.surveymonkey.com). Here we report the open responses regarding the use of crowdsourcing.

RESULTS

Attending and resident physicians represented the majority of DocCHIRP account holders (91%), with nurse practitioners accounting for the remaining sample (9%). There were 50 male and 35 female participants, with an age range of 28 to 78 years (median age, 43 years). Departmental affiliations included Pediatrics (n = 28, 33%), Neurology (n = 27, 32%), Internal Medicine (n = 10, 12%), Psychiatry (n = 4, 5%), the Division of Pediatric Neurology (n = 11, 13%), and others (n = 5, 6%). Of the 1544 total visits to the DocCHIRP site, providers favored using smart phones (56.8%) and tablets (9.5%) over the desktop interface (33.6%; Figure 1C). iPhone use (81.7%) surpassed the other platforms combined. Desktop users visited twice as many pages (16.8 pages/visit) compared to those using smart phones (5.5 pages/visit) or tablets (8.6 pages/visit). Desktop users remained engaged longer than mobile users (~13 vs ~5 minutes). In the post-trial user survey, we received 72 valid surveys from 85 potential participants (85% response rate).

We used a tiered enrollment design, sending invitations to potential participants in 3 phases to study the relationship between the size of the HCP crowd and sustained use as reported in other social networks.[13] Using a cutoff of >3 visits per week to demarcate active periods of use, we saw during the initial phase of enrollment that 20 providers generated a total of 170 visits over 22 days (Figure 2A). The addition of 28 members (phase II, n = 48 total) extended active use by 28 days, with a total of 476 page visits. The addition of 32 members (phase III, n = 85 total) resulted in 56 days of active participation with 612 visits to the site. When plotted (Figure 2B), the relationship between crowd size (total number of registered users) and cumulative visits (R^2 = 0.951), as well as crowd size and days of high activity (R^2 = 0.953) were linear and direct. We also investigated

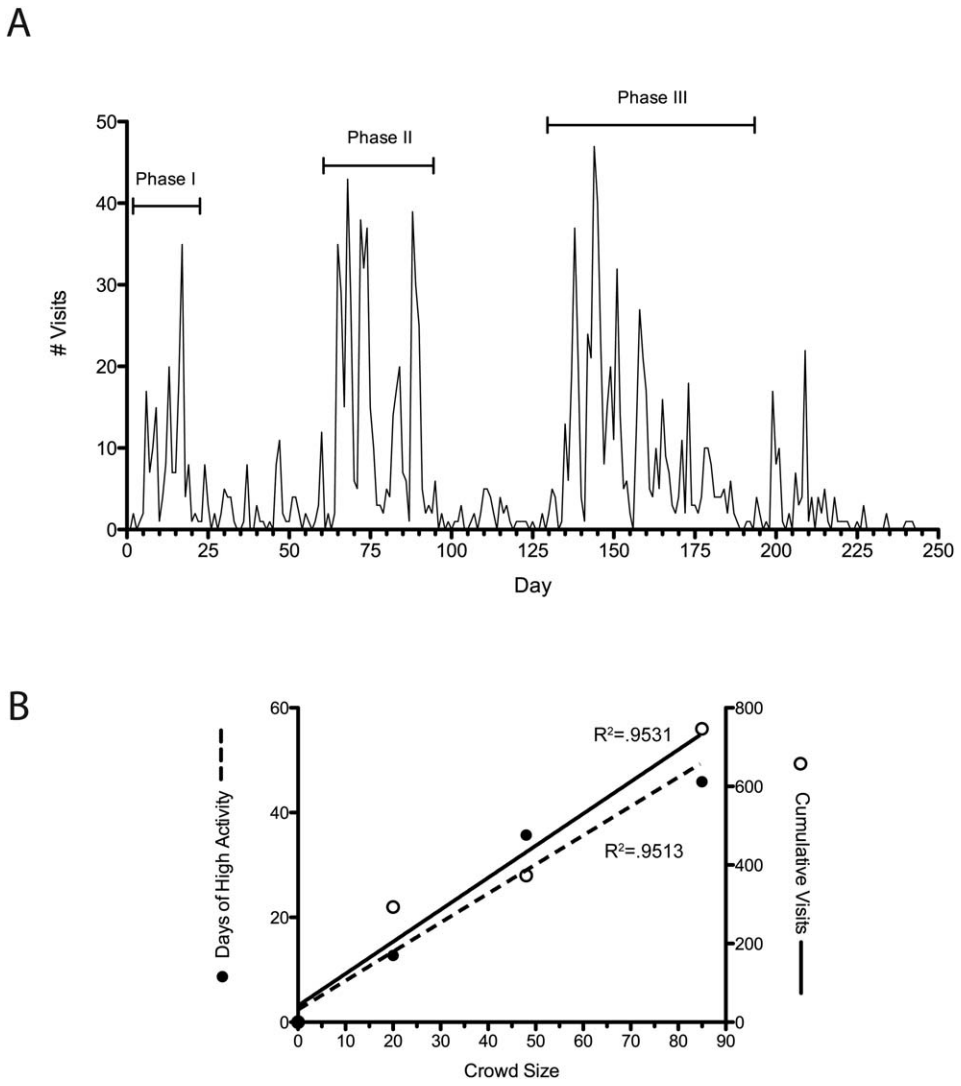


FIG. 2. Activity of provider engagement during the 3 phases of the DocCHIRP field trial. (A) Providers were recruited to participate in the field trial in 3 distinct phases between April 1, 2012 and November 30, 2012. Periods of significant use were determined in each phase as described in the methods. (B) Plot demonstrating the relationship between days of high activity (dashed line), cumulative visits (solid line), and crowd size.

the timing of user engagement by pooling the data and breaking down use by time of day and day of the week (Figure 3A,B). In addition to observing peak engagement during the midmorning and afternoon, times of anticipated physician-patient contact, we observed a third use peak in the evening. With the exception of sporadic weekend use, DocCHIRP use clustered during midweek.

DocCHIRP users generated 45 questions. The fastest first response was returned in less than 4 minutes, with a median first response time of 19 minutes (Figure 3C). Analysis of the consult requests received revealed a clustering of 7 principal question-response groups: (1) the effective use of medications, (2) complex medical decision making, (3) use of the application itself, (4) guidance regarding the standard of care, (5) selection and interpretation of diagnostic tests, (6) differential diagnosis, and (7) patient referral (Figure 3D). Consults regarding medication use and complex decision making were dominant themes (63%). Several consults generated multiple responses, broadening the scope of the original query or requesting additional information (Table 1).

To better understand factors influencing use of the mobile crowdsourcing application, we surveyed users, receiving 68 comments related to the overall approach and barriers to adoption among other aspects (Table 2). The 40 comments regarding the use of medical crowdsourcing were divided evenly between supporters and critics. Enthusiasm for cross-discipline collaboration, having tools to codify expert knowledge, and discovering consensus opinion from the expert crowd was offset by concerns that push notifications would distract providers, compromise efficiency, and potentially lead providers to act on inaccurate information.

DISCUSSION

In the current study, we developed and field-tested the application DocCHIRP, which helps HCPs crowdsource information from each other in near real time.

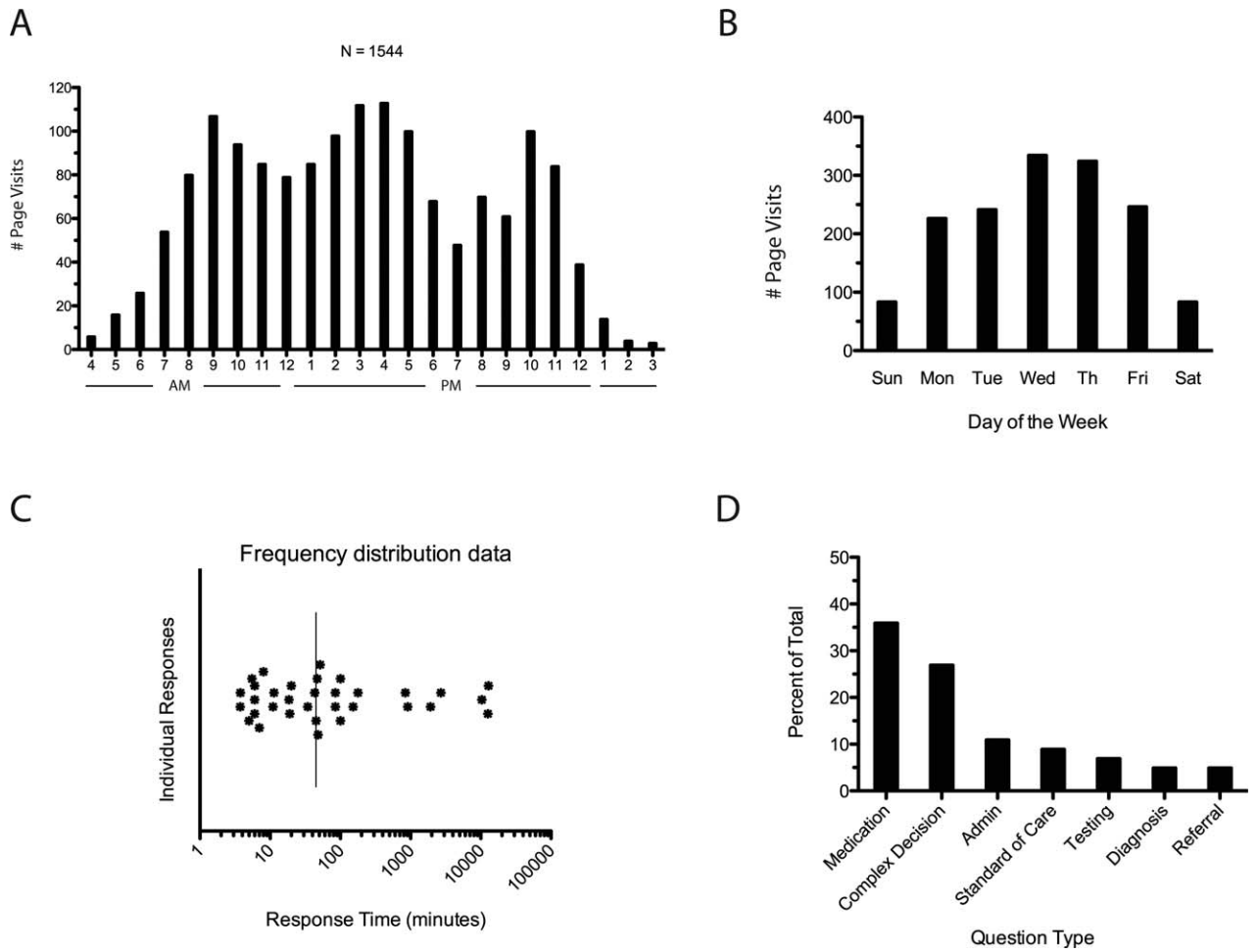


FIG. 3. Analysis of provider visits to the DocCHIRP server. The data from the 3 trial periods were combined and plotted according to: (A) the frequency of user engagement by time of day, and (B) by the day of the week. (C) Frequency distribution of response latencies observed in the field trial showing the number of discrete queries against the response time in minutes. The median response time is shown as a vertical line. (D) Histogram demonstrating the content of the initial consult questions submitted (n = 45).

The average response latency in this pilot trial was 20 minutes, which was unexpectedly fast given the relatively small size of the participating crowd. Additionally, nearly one-third of users accessed the server in the evening using the web interface rather than their mobile phone. This suggests that although HCPs liked having direct access to colleagues in near real time, the also valued the opportunity to connect asynchronously after hours.

Relative to the total number of page views, the number of HCPs using the technology for peer-to-peer consultation was low. Feedback provided in the post-trial survey suggested several reasons for this effect. Some providers viewed the application without posting because they were reluctant to disclose knowledge gaps to their peers. Several users suggested implementing a system that supports anonymous posting, but others thought this would undermine the value of the information provided. Additionally, users recognized the potential for crowdsourcing to adversely effect HCP's productivity and daily workflow. This is relevant given growing concerns about "distracted doctoring" and association with reduced safety and quality of medical care.[14] This concept is further echoed in a paper by Wu et al. demonstrating that frequent interruptions offset the perceived benefit of increased mobility afforded by the use of mobile technology.[15] However, it is worth considering that if implemented properly, study participants believe crowdsourcing could have a net neutral impact on clinical workflow by improving the efficiency of provider communication and saving time otherwise spent problem solving. Participants also felt the approach could infringe on an already threatened work-life boundary, and could also lead to unprofessional and antisocial behaviors.[16] Collectively, these problems are not unique to medical crowdsourcing, and prior experience in this area may offer several viable solutions. First, because crowd burnout is inversely proportional to crowd size, successful adoption in practice will require growing a provider base of sufficient depth and expertise to handle the consult demand. With the expansion of accountable care organizations across the United States, this will not likely be a limiting

TABLE 1. Sample Consults and Responses From the DocCHIRP Community

Question Type	Consult	Response(s)
Medication	How do you treat headache from viral meningitis?	R1: Any analgesic will work; need to clarify that the headache is not post-LP, which may require blood patch.
	Anyone know how oral fluconazole (liquid) tastes? We needed to prescribe for a young 13 year old.	R1: We should get a pharmacist on the chat. I would call the pharmacy and see if they can compound it with flavoring.
	How frequently do your patients complain of myalgias on statins? Have you prescribed coenzyme Q in this situation?	R1: Did you see the editorial in the Green Journal yesterday?Took the position that statins were not to blame. I usually give a trial off to make sure symptoms resolve. Usually I try them on a different statin.Have not routinely rx'd Q10.
Complex medical decision making	Has anyone seen tapeworm infection from raw pork? Do we need to report this? We treated with mebendazole.	R1: You can check with CDC here: http://www.cdc.gov/parasites/cysticercosis. R2: First-line treatment for *Tsolium* is praziquantel or albendazole.However, mebendazole has also been used to successfully treat *T solium*. R3: Whipworm is another common pork tapeworm.It is also covered by mebendazole
	What are the current guidelines regarding the use of statins in patients with a history of lobar hemorrhage.	R1: Larger studies (SPARCL, HPS) both showed higher hemorrhage risks in statin treated patients.Cohort studies generally don't show an obvious risk to statins. I've generally taken patients off their statins when they come in with lobar ICH, and more neutral when it's a hypertensive bleed.
Standard of care	How often would someone have to fall before you felt uncomfortable anticoagulating for AFib?	R1: The risk of falls alone should not automatically disqualify a person from being treated with warfarin. R2: I recall reading a meta-analysis that suggested ~300 falls/year would start to favor not anticoagulating, but short of that, falls were not an important factor.
	Anyone used IVIG for any of the following: autoimmune encephalopathy, NMO, paraneoplastic limbic encephalitis, PANDAS?	R1: We had a patient recently with a history of autoimmune encephalopathy who was treated with IVIG.
Administrative	What medical apps do you have on your phone?	R1: DocCHIRP, Epocrates, NIH stroke calculator. R2: I have Merck Medicus, Micromedex drugs, growth charts, and shots. . .all those are free.I also have Red Book from AAPand Sanford Guide, which I paid for. R3: Instant ECG, ACLS Advisor, 10-Second EM.
Testing	What would be considered a normal vitamin D level in a 2 year old?	R1: We typically treat at a level less than 30, with likely greater treatment if less than 21. I'm sure our phone nurses would be willing to share [our protocol].
	I have an obese 13-year-old AA girl with acanthosis nigricans. Do you check HbA1c?	R1: Yes. Sign of insulin resistance. HbA1c along with fasting blood glucose are a good start.Close monitoring indicated regardless. Endo may have more insight as to whether or not other labs are useful, such as fasting C-peptide.
Referral	Has anyone ever seen preteen or teen patients with ADHD-like symptoms and poor sleep referred for a sleep study for possible restless leg syndrome?	R1: RLS seen in kids, but criteria are different for children than adults.Sleep studies may be warranted. R2: I've also heard about a link between restless leg and iron deficiency. Is it a girl? R3: Checking CBC, ferritin, and iron is a good start.

NOTE: **Abbreviations**:AA, African American; AAP, American Academy of Pediatrics;ACLS, Advanced Cardiovascular Life Support; ADHD, attention deficit hyperactivity disorder;CBC, complete blood count;CDC, Centers for Disease Control and Prevention; ECG, electrocardiograph; EM, emergency medicine; Endo, endocrinologist;HPS, Heart Protection Study; HbA1c, hemoglobinA1c; ICH, intracerebral hemorrhage;IVIV, Intravenous immunoglobulin; LP, lumbar puncture; NIH, National Institutes of Health; NMO, neuromyelitisoptica; PANDAS, Pediatric Autoimmune Neuropsychiatric Disorders Associated with Streptococcal Infections;RLS, restless legs syndrome; SPARCL, Stroke Prevention by Aggressive Reduction in Cholesterol Levels.

TABLE 2. Summary Response of Trial Participants Regarding Aspects of DocCHIRP

Category		Comments
Overall approach	Pro	"This is a process whose time has come; we need it to adapt to the exponential increase in information content that impacts our clinical decisionmaking" "I found [the application] it to be both useful and interesting."
	Con	"I just don't like these types of things. . .email already takes up too much time." "Curbside consults result in worse outcomes for the patient and the physician. I found myself uncomfortable using this approach." "My biggest concern is the interruption in one's thinking. . ..distractions are becoming increasingly common." "I do appreciate colleagues input; but ask for it verbally. . ..I am struggling to learn even texting. . .."
Barriers to adoption	Pro	"I think premise is great, it is just a matter of enough people participating to make it worthwhile to use." "There is power in numbers here-people won't use it unless there is lots of activity or feedback." "I think it will be very useful if the whole department or sections are involved in promoting and participating. . .."
	Con	"I did not test it much since the posts were not very frequent at the time that I tried it." "The barrier to use is quality control; how to substantiate the 'quality' of input provided is key."
Anonymous posting	Pro	"I would not have [posts] always be anonymous, but allow the user the option. . .." "Anonymity would be great. . .I was concerned that some of my questions were 'dumb'.
	Con	"Anonymous posting would increase the risk of trolling."
Suggested uses		"I see a role for this app in relaying questions to subspecialty groups for 'judgment call' questions." "Best place to talk about weird cases, odd presentations; to ask 'have you ever seen anything like this before'." "Consider rolling it out to entire family medicine department and/or primary care network."

factor. And although not implemented here, flexible notification settings, user-defined identity rules, and other higher-level software design elements should alleviate the issues related to provider reputation and workflow interruptions.

Overall, HCPs are optimistic that mobile handheld technologies will benefit their practice.[17] Yet, software-based approaches including expert decision support systems must overcome particular hurdles including lack of provider trust in the algorithms used in these approaches.[18] In the end, trust is ultimately a human phenomena; users will only trust the system if they know the information came from a trusted and highly reputable individual or institution. By tapping the expertise of a network of institutional colleagues, crowdsourcing addresses this issue of trust. Appropriately, providers were also concerned about the legality and personal risk of using crowdsourcing to discuss matters related to patient care. The technology was not intended to share protected health information, and as with other forms of digital communication, providers were cautioned during the consent process to monitor their behavior in this regard. Although soliciting advice from the medical crowd has an inherently higher level of risk compared to the use of crowdsourcing in education, research, or business, the index provider is ultimately responsible for considering all available information before making any treatment decision.

Though our pilot trial was not designed to assess effects on HCP efficiency or on the quality of care delivered, our work provides a unique window on the information-seeking behaviors HCPs and highlights potential modifications that could enhance the utility of future crowdsourcing programs. Because the trial was performed within the context of an academic health center, it remains to be seen how medical crowdsourcing will translate in private practice, rural clinics, and other clinical environments where peer-to-peer consultation is sought. Given the potential for high-stakes information exchanges, further study regarding the use of medical crowdsourcing in a controlled environment will be required before the technology can be disseminated to a broader audience. If future iterations of the mobile crowdsourcing application can address the aforementioned adoption barriers and support the organic growth of the crowd of HCPs, we believe the approach could have a positive and transformative effect on how providers acquire relevant knowledge and care for patients.

Acknowledgements

The authors thank the physicians and nurse practitioners at the University of Rochester who participated in the trial. The authors also acknowledge Dr. Dan Goldstein at the Microsoft Research Group (New York, NY) for many helpful discussions.

Disclosures: This study was funded in part by grant support from the University of Rochester Robert B. Goergen Reach Fund (M.H.S.). Collaborative Informatics, LLC provided integrated mobile and server software used in this study. Dr. Halterman is co-owner of Collaborative Informatics, LLC and oversaw the specifications and construction of the software used in this study. Dr. Halterman has provided the necessary conflict of interest documentation in keeping with the requirements of the University of Rochester. The DocCHIRP study was reviewed by the institutional review board at the University of Rochester and received approval posing minimal risk.

References

1. Davies K, Harrison J. The information-seeking behaviour of doctors: a review of the evidence. *Health Info Libr J*. 2007;24(2):78–94.
2. Andrews JE, Pearce KA, Ireson C, Love MM. Information-seeking behaviors of practitioners in a primary care practice-based research network (PBRN). *J Med Libr Assoc*. 2005;93(2):206–212.
3. Perley CM. Physician use of the curbside consultation to address information needs: report on a collective case study. *J Med Libr Assoc*. 2006;94(2):137–144.
4. Kothari AR, Bickford JJ, Edwards N, Dobbins MJ, Meyer M. Uncovering tacit knowledge: a pilot study to broaden the concept of knowledge in knowledge translation. *BMC Health Serv Res*. 2011;11:198.
5. DeCato TW, Engelberg RA, Downey L, et al. Hospital variation and temporal trends in palliative and end-of-life care in the ICU. *Crit Care Med*. 2013;41(6):1405–1411.
6. McGinn CA, Grenier S, Duplantie J, et al. Comparison of user groups' perspectives of barriers and facilitators to implementing electronic health records: a systematic review. *BMC Med*. 2011;9:46.
7. Howe J. The Rise of Crowdsourcing. *Wired magazine*. 2006;14(6):1–4.
8. Hohman M, Gregory K, Chibale K, Smith PJ, Ekins S, Bunin B. Novel web-based tools combining chemistry informatics, biology and social networks for drug discovery. *Drug Discov Today*. 2009;14(5–6):261–270.
9. Ranard BL, Ha YP, Meisel ZF, et al. Crowdsourcing—harnessing the masses to advance health and medicine: a systematic review. *J Gen Intern Med*. 2014;29(1):187–203.
10. Katz-Sidlow RJ, Ludwig A, Miller S, Sidlow R. Smartphone use during inpatient attending rounds: prevalence, patterns and potential for distraction. *J Hosp Med*. 2012;7(8):595–599.
11. Shortliffe EH. Biomedical informatics in the education of physicians. *JAMA*. 2010;304(11):1227–1228.
12. Bakul P. Mobile medical applications: guidance for industry and Food and Drug Administration staff. Washington, DC: U.S. Department of Health and Human Services, Food and Drug Administration; 2013.
13. Rutherford A, Cebrian M, Dsouza S, Moro E, Pentland A, Rahwan I. Limits of social mobilization. *Proc Natl Acad Sci U S A*. 2013; 110(16):6281–6286.
14. Papadakos PJ. the rise of electronic distraction in health care is addiction to devices contributing. *J Anesth Clin Res*. 2013;4:e112.
15. Wu R, Rossos P, Quan S, et al. An evaluation of the use of smartphones to communicate between clinicians: a mixed-methods study. *J Med Internet Res*. 2011;13(3):e59.
16. Spiegelman J, Detsky AS. Instant mobile communication, efficiency, and quality of life. *JAMA*. 2008;299(10):1179–1181.
17. Prgomet M, Georgiou A, Westbrook JI. The impact of mobile handheld technology on hospital physicians' work practices and patient care: a systematic review. *J Am Med Inform Assoc*. 2009;16(6):792–801.
18. Alexander GL. Issues of trust and ethics in computerized clinical decision support systems. *Nurs Adm Q*. 2006;30(1):21–29.