Towards Crowd-Based Needfinding in Medical Device Development[1]

Cory R. Schaffhausen
Department of Mechanical Engineering,
University of Minnesota,
111 Church Street SE,
Minneapolis, MN 55455

Timothy M. Kowalewski
Department of Mechanical Engineering,
University of Minnesota,
111 Church Street SE,
Minneapolis, MN 55455

1 Background

Public and private institutions invest over $100 billion annually in biomedical research [1]. These funds can be misused and the value diminished if the engineers and clinicians are not aligning work with areas of clinical need [2]. As medical device developers have pointed out, "Get (the clinical need) right and you have a chance, get it wrong and all further effort is likely to be wasted" [3, p. 3]. Needfinding is a formal tool within the design cycle used to fully understand the problem, particularly when involving human behavior [4]. While effective when used, these methods are reported to be underutilized in healthcare practice [5–7] due to unique challenges in healthcare [8,9]. This lack of early stage input can lead to unnecessary investment in new projects already destined to fall short of the potential clinical improvements motivating the work.

The results reported here represent a preliminary feasibility study on a crowd-based method of needfinding. The long-term objective is to collect a broad, unfiltered list of potential unmet clinical needs from the point-of-view of technology users. This method can compliment data from existing observation-based methods currently used [3,4], but will have several advantages over existing methods, including: needs can be collected from a wider, more diverse sample of a surgeon population; individual time commitments can be minimized relative to in-depth interviews and observations; and the same data collection infrastructure can be leveraged to later prioritize unmet needs through formalized review and input from clinicians and professional medical societies.

The study is designed to test the effectiveness of methods to overcome existing limitations of nonobservation-based interactions with users. Specifically, users are often unable to directly articulate how a technology fails to meet their needs because they have adapted to shortcomings and rarely consciously think of them [3].

2 Methods

This feasibility study is a first phase in a multipart study to validate data collection methods, filter data to remove duplications using automated textual semantic similarity analysis, and rate need quality from the perspective of users. In order to avoid overwhelming clinician collaborators with a long series of tests of multiple variables, the first feasibility tests are performed on nonmedical applications and recruit participants from the general public. These results will inform later phases to adapt the methods specifically to medical device applications.

Our study tests the possibility of improving a users' ability to articulate needs using a wide variety of narrative prompts derived from best-practice design methods [10,11]. These prompts may help users view interactions with technology with new perspectives. Figure 1 represents a simplified matrix outlining these perspectives. On one axis, interactions with products or devices can be viewed on a spectrum ranging from emotional to technological. On the other axis, a user perspective can range from first person (internal) to third person (empathy).

Each cell of the matrix is populated with one or more narrative prompts of 3-6 sentences directing the reader to think about a specific perspective. For example, when viewing medical devices from a perspective of habits and an internal focus, a shortened summary of a prompt could be "I tried to use that once, but I don't use it anymore." In this one example, such a perspective may elicit an experience of an unsuccessful attempt to adopt a new and potentially beneficial clinical process. This identifies a need to reduce barriers to adoption, and the contextual information we collect will identify what the specific barriers are. The remaining areas of the matrix may elicit very different types of experiences, and the most effective prompts might be tunable based on the user demographics and experiences. When needs are collected from a wide range of perspectives and from a large group of diverse clinicians, the resulting data will potentially provide a finely detailed picture of the problem space.

In order to test the feasibility of collecting open-ended data from an online interface, we conducted a pilot of 25 individuals. Participants were recruited using Amazon Mechanical Turk, an online marketplace to hire workers for brief "human intelligence" tasks often for problems poorly suited to computational methods or machine learning. The requesters hiring workers can stipulate various inclusion criteria (such as a United States IP address) and then pay nominal amounts, often less than $1 for a completed task. Early phases of this study will continue to use more available, nonmedical platforms to refine our methods before studying the approach in broad medical contexts. We expect methods with broad effectiveness in general products to be most effective in diverse medical technologies, similar to broad effectiveness of iterative and rapid prototyping methods.

In our study, workers viewed a summary of the task, in this case to answer questions about problems with common products and services. Each participant was randomly assigned to one of four focus areas (such as cooking and household chores), reviewed informed consent information, and participated in a training exercise to understand the difference between an unmet need and an invention. The instructions stated that complete descriptions are required, and inventions were explicitly discouraged. Participants were told to focus on generating as many problem or need statements as they could regardless of whether they currently thought the cost of finding a solution would exceed the benefit, and they were told they would be paid small bonuses for higher quantities (for example, $0.20 for each five problem descriptions). Future applications to medical devices will not use fee-based incentives; instead, incentive structures will rely on alternatives, such as professional society recognition.

Each participant was required to pass a training quiz with 100% correct answers. In the quiz, participants identified incomplete descriptions and also identified inventions from a list of five examples. The final step before continuing was an optional set of

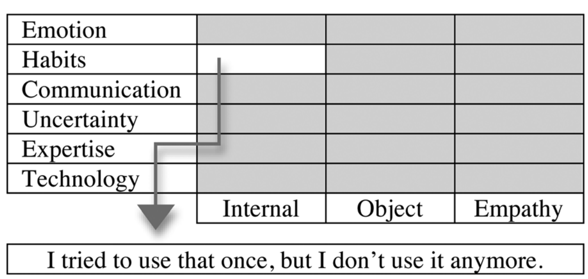

Fig. 1 Matrix of narrative prompts

[1]Accepted and presented at the Design of Medical Devices Conference (DMD2014), Minneapolis, MN, April 7–10, 2014.
DOI: 10.1115/1.4027085
Manuscript received February 21, 2014; final manuscript received March 3, 2014; published online July 21, 2014. Editor: Arthur G. Erdman.

demographics questions as well as questions on personality traits. This information will be used to analyze if any particular prompt or row or column of the matrix is most beneficial to a particular type of person.

Participants then began submitting written problem statements. After each statement, he or she had the option of providing a background story for additional context about the problem. As described above for an example of trying a new technique once and not continuing, a problem statement might indicate there are barriers to adoption, and the contextual information would identify what the specific barriers are.

When participants had exhausted readily available problems that came to mind, they were able to ask to for help and were presented with a single, randomly assigned narrative prompt. Afterwards, they were able to reflect on this and continue to submit any new problem or need descriptions. This count of needs submitted after the prompt indicates whether the prompt was effective. The analysis of the complete narrative prompt study will be a regression analysis comparing the count of needs submitted after reviewing a prompt accounting for the quantity of needs submitted before the prompt. The current feasibility study data were reviewed for evidence of participants requesting narratives and submitting additional needs and stories.

3 Results

In order to include 25 participants with a passing training quiz, 49 individuals began the task, and 24 failed the quiz. The total time to recruit 25 participants who passed training was 3 days. The range of final compensation per individual was $0.30 to $1.10. All individuals who passed the quiz answered all optional demographic and personality questions. A total of 77 needs or problems were submitted, and 62% included optional stories to provide detailed context. Of the 11 individuals requesting a narrative, the range of needs submitted before the prompt was 0–7, mean 2.4. The range of needs submitted after the prompt was 0–6, mean 1.5.

4 Interpretation

The results of this feasibility study suggest that open-ended needs and problems data can be successfully collected via online crowd-based platforms. The results also suggest that individuals who had indicated they did not have additional information to share are able to identify more needs after reviewing a narrative prompt focusing on one perspective of interacting with products.

A larger sample will be required to fully analyze differences based on demographics, personality type, type of product, and type of narrative prompt. Further study is also warranted to study the effects of multiple narrative prompts and prompts of other kinds. While these results are encouraging to further evaluate the use of prompts, the long-term value of the data collected will be dependent on the variability or degree of redundancy in responses as well as the quality of the needs based on metrics that would reflect the likelihood of impacting patient care should the need be translated into a final product or device. Additional research in these areas will be critical to fully capitalize on the data collection methods tested here.

References

[1] Dorsey, E. R., de Roulet, J., Thompson, J. P., Reminick, J. I., Thai, A., White-Stellato, Z., Beck, C. A., George, B. P., and Moses, H., 2010, "Funding of US Biomedical Research, 2003–2008," JAMA, **303**(2), pp. 137–143.

[2] Moses, H., Dorsey, E. R., Matheson, D. H. M., and Thier, S. O., 2005, "Financial Anatomy of Biomedical Research," JAMA, **294**(11), pp. 1333–1342.

[3] Zenios, S. A., Makower, J., Yock, P. G., Brinton, T. J., Kumar, U. N., Denend, L., and Krummel, T. M., 2010, *Biodesign: The Process of Innovating Medical Technologies*, Cambridge University Press, Cambridge, UK.

[4] Patnaik, D., and Becker. R., 1999, "Needfinding: The Why and How of Uncovering People's Needs," Design Manage. J. (Former Series), **10**(2), pp. 37–43.

[5] Money, A., Barnett, J., Kuljis, J., Craven, M., Martin, J., and Young. T., 2011, "The Role of the User Within the Medical Device Design and Development Process: Medical Device Manufacturers' Perspectives," BMC Med. Inform. Decision Making, **11**(2), p. 15.

[6] Shah, A., and Alshawi, S., 2010, "The Role of User Requirements Research in Medical Device Development," European, Mediterranean & Middle Eastern Conference on Information Systems (EMCIS2010), Abu Dhabi, UAE, April 12–13, Paper No. C100.

[7] Poulton, B. C., 2001, "User Involvement in Identifying Health Needs and Shaping and Evaluating Services: Is it Being Realised?" J. Adv. Nursing, **30**(6), pp. 1289–1296.

[8] Martin, J., and Barnett, J., 2012, "Integrating the Results of User Research Into Medical Device Development: Insights From a Case Study," BMC Med. Inform. Decision Making, **12**(7), p. 74.

[9] Vredenburg, K., Mao, J. Y., Smith, P. W., and Carey, T., 2002, "A Survey of User-Centered Design Practice," SIGCHI Conference on Human Factors in Computing Systems (CHI 2002), Minneapolis, MN, April 20–25, pp. 471–478.

[10] IDEO, 2003, *Method Cards: 51 Ways to Inspire Design*, Palo Alto, CA.

[11] Mattelmäki, T., 2006, *Design Probes*, University of Art and Design Helsinki, Helsinki, Finland.