Viewpoint

Software Testing of eHealth Interventions: Existing Practices and the Future of an Iterative Strategy

Oyinda Obigbesan[1]; Kristen Graham[2], BSc, MBA; Karen M Benzies[1,3], RN, PhD

[1]Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[2]Haskayne School of Business, University of Calgary, Calgary, AB, Canada

[3]Faculty of Nursing, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Karen M Benzies, RN, PhD
Faculty of Nursing
University of Calgary
2500 University Drive NW
Calgary, AB, T2N 1N4
Canada
Phone: 1 403 220 2294
Email: benzies@ucalgary.ca

Abstract

eHealth interventions are becoming a part of standard care, with software solutions increasingly created for patients and health care providers. Testing of eHealth software is important to ensure that the software realizes its goals. Software testing, which is comprised of alpha and beta testing, is critical to establish the effectiveness and usability of the software. In this viewpoint, we explore existing practices for testing software in health care settings. We scanned the literature using search terms related to eHealth software testing (eg, "health alpha testing," "eHealth testing," and "health app usability") to identify practices for testing eHealth software. We could not identify a single standard framework for software testing in health care settings; some articles reported frameworks, while others reported none. In addition, some authors misidentified alpha testing as beta testing and vice versa. There were several different objectives (ie, testing for safety, reliability, or usability) and methods of testing (eg, questionnaires, interviews) reported. Implementation of an iterative strategy in testing can introduce flexible and rapid changes when developing eHealth software. Further investigation into the best approach for software testing in health care settings would aid the development of effective and useful eHealth software, particularly for novice eHealth software developers.

Introduction

eHealth interventions are becoming a part of standard care, with mobile apps or software solutions being created for patients and health care providers. For example, we designed a digital audit software for a new model of care called Merge (which is better known in research and development as Alberta Family Integrated Care) that improves outcomes by integrating families of newborns with critical illness into the neonatal intensive care unit (NICU) team [1]. Merge was adopted in all 14 NICUs in Alberta, Canada, through a process that involves quarterly fidelity audits, which are central to implementation success and sustainability. We replaced the manual and labor-intensive audits with this fidelity audit software to increase efficiency and reduce personnel costs.

Appropriate software testing is important to ensure that an eHealth intervention realizes its goals, which may include improving access, efficiency, and quality of care [2]. Software testing, which includes alpha and beta testing, is critical to establish the effectiveness and usability of an eHealth software. In this viewpoint, we define alpha testing in accordance with the Guide to the Software Engineering Body of Knowledge [3], as internal software testing, which occurs first and is often performed within the development team [4]. Subsequently, beta testing is performed, which is external and involves a larger testing sample representative of the end-user population [3,4].

XSL•FO
RenderX

When we set out to test the Merge fidelity audit software, we wanted to identify software testing practices for novice eHealth software developers. Thus, we performed an exploratory scan of the literature to understand and identify existing methods for testing software in health care settings.

For our exploratory scan, we used Google Scholar and search terms related to eHealth software testing (eg, "health alpha testing," "eHealth testing," and "health app usability"). We included research articles that described their process of either alpha or beta testing and were either patient- or health care provider–facing. We excluded articles if they were nonpeer-reviewed literature published before 2015 from a non-Organisation for Economic Co-operation and Development member country because we were interested in recent practices conducted in similar health care contexts. We identified 34 articles that met our inclusion criteria; among these, we selected 7 articles for this viewpoint about software testing practices (Multimedia Appendix 1), including 5 primary research articles [5-9], 1 review article [10], and 1 framework [11].

Lack of a Standardized Testing Framework

There was not a single standardized framework for software testing that was used across the 5 primary research articles [5-9]; 2 articles [7,8] cited two different frameworks, while the other 3 articles did not cite any. Fishbein et al [8] used Darlow and Wen's [12] best practices to guide the development of their mobile health (mHealth) intervention. The practices relevant for software testing included using mixed methods (eg, questionnaires, semistructured interviews), engaging stakeholders (eg, software designers, subject matter experts, health professionals, patients), and publishing the results of testing to facilitate learning from successes and difficulties [12]. Cho et al [7] used the 3-level stratified view of health IT usability evaluation [13]. Levels of testing, as defined by this framework, consisted of (1) a user-centered design, which incorporated users' needs in the development phase; (2) a usability evaluation in a laboratory setting, where testing occurred in a regulated environment; and (3) usability in a real-world setting, which assessed user experience in practical applications [7,13]. In Cho et al [7], levels 2 and 3 most closely correspond to beta testing, as testers are representative of the end-user population [7]. While Darlow and Wen's [12] best practices provide guidelines regarding what the testing should include, the stratified view of health IT usability evaluation [13] provides more specific guidance on how to conduct testing and what each phase may entail. However, both frameworks involve multiple methods of testing to gain comprehensive user feedback and aim to create standardization in testing to ensure the rigor of eHealth interventions. The lack of standardization in software testing raises issues concerning consistency in quality assurance, which has the potential to compromise patient safety. Additionally, the lack of clear guidelines for testing can lead to inefficiencies in the allocation of the development team's time and resources.

Objectives and Methods of Testing

Although few articles used a specific framework for testing, there were some similarities in the general process among the 5 primary research articles[5-9] (Multimedia Appendix 1). For both alpha and beta testing, this involved defining the objectives of testing, selecting testers and methods of evaluation, collecting and analyzing data, and refining the software.

The desired result of testing was the evaluation of different aspects of the eHealth software. Only 2 articles [5,6] clearly stated their objectives for testing. Ahonen et al [5] planned to identify content acceptability, feasibility, and technical issues during alpha testing of their eLearning intervention. For their mHealth intervention, Athilingam et al [6] aimed to evaluate design and functionality during alpha testing, followed by evaluations of helpfulness, usability, and design during beta testing. In addition to these objectives, software testing was used to evaluate safety, reliability, effectiveness, satisfaction, and accessibility, and to determine if the software fulfilled its intended purpose [14]. While the objectives of testing may overlap between alpha and beta testing, alpha testing may have a greater focus on software-specific objectives (eg, safety and bug fixes) and beta testing may focus on user experience–related objectives (eg, design and satisfaction) [4].

Selecting testers was an important testing step mentioned in all articles that described testing; however, the distinction between alpha and beta testing was often obscured at this stage. Several authors [5,6] stated that they were conducting alpha testing, yet the testing sample included members of the end-user population, which is more appropriately aligned with the definition for beta testing. For example, Athilingam et al [6] described alpha testing of their mobile app for patients with heart failure, but the testing sample consisted of the target end users (ie, people with a history of heart failure). Additionally, there was no clear justification for the size of the testing sample, which ranged from 2 to 76 testers [5-9]. In a scoping review, the sample size for software testing was found to vary according to the method of testing, with studies based on qualitative methods having fewer testers than those based on quantitative methods [10]. Most studies opted for a larger number of testers and tested their software only once. Hoffiman et al [9] had only 2 alpha testers who tested their Public Open Space Tool software on 55 different green spaces. The feasibility of this strategy depended on the same tester generating new insights during each round of alpha testing.

The most common methods of testing were questionnaires, "think-aloud" techniques, interviews, and focus groups [10]. The literature suggests using qualitative methods of testing over quantitative methods because qualitative methods generally result in a deeper understanding of user experience and gaps in the software [5,10]. Mixed methods were used in several studies [5,7,8,10], where quantitative methods (eg, questionnaires) were used to gain initial feedback followed by in-depth qualitative data collection (eg, interview or focus group) to ensure that the feedback was valuable and instructive. Using mixed methods allowed development teams to discover latent information and gain a better understanding of their participants' experiences and needs [7,10].

When selecting a method of testing, development teams should particularly consider their specific testing sample. Methods of testing that may be effective for alpha testers may be unsuccessful when used with beta testers. Factors affecting the accuracy of testing data include the length of testing, question clarity and interpretation, question sequence, and the context in which testing takes place [15]. The general approach to software testing culminated in the collection and analysis of data, which were then used to create an iteration of the eHealth intervention.

Throughout the testing process of eHealth software, it is essential to uphold ethical principles due to the sensitive nature of health data and the potential impact on patient well-being. This includes ensuring patient privacy and confidentiality in compliance with national health care regulations, obtaining informed consent prior to participation, and adopting data protection and security measures to protect sensitive information [16,17].

An Iterative Strategy

The success of eHealth development depends on the ability to adapt to the rapidly evolving nature of the digital world and changing user needs. Development teams interested in building eHealth interventions must consider how to ensure that testing occurs quickly, while still guaranteeing rigor and protecting privacy. Wilson et al [11] proposed an "mHealth Agile and User-Centered Research and Development Lifecycle" that combines an agile approach with traditional clinical trial phases to create high-quality mHealth interventions. Using an agile approach involves continuous iterative cycles that prioritize providing feedback on the software at frequent intervals. Compared to a linear approach to development, continuous evaluation of the developing software enhanced team collaboration, improved the quality of the software, and allowed for ongoing improvement with emphasis on end-user feedback [11].

Wilson et al [11] further suggested that alpha testing can be used to gain first-impression, surface-level insights from the development team, focusing on usability, desire to use, fulfillment of intended purpose, and safety. External testers can then be involved in beta testing, where usefulness, feasibility, and acceptability are evaluated [11]. When developing software for health care settings, several iterative rounds of alpha testing before proceeding to iterative rounds of beta testing may be most effective [11]. To further expand the breadth of feedback and reduce recall bias, iterative rounds of beta testing may be conducted with different samples of the end-user population. Depending on the revisions to the eHealth software after beta testing, a development team may decide to return to alpha testing to begin the process again. This highlights a within-and-between iterative strategy when conducting alpha and beta testing, offering the development team the flexibility to make continuous improvements to their eHealth software.

Conclusion

Currently, there is great variation in how software testing is conducted in health care settings. However, we found that an iterative approach to testing is compatible with the need for an agile development technique for eHealth software. Qualitative methods of testing tend to yield more in-depth user experience feedback during beta testing and researchers benefit when two or more methods are used throughout the beta iteration. Yet, it remains unclear what prompts the transition from alpha to beta testing and when a repeat of testing is required. In the rapidly developing field of eHealth and mHealth interventions, it would be useful to have an agreed upon definition of what constitutes alpha and beta testing. Additionally, with the various frameworks of testing available, the best approach to software testing in health care settings remains unclear. To create clarity in this process, we suggest conducting a systematic review to understand and appraise the full scope of software testing practices within health care settings.

Authors' Contributions

KMB conceptualized the study. OO and KG conducted the scan of the literature and wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Conflicts of Interest

KMB is the founder and CEO of Liminality Innovations. KG is an employee of Liminality Innovations. OO has no conflicts of interest to report.

Multimedia Appendix 1

Characteristics of included studies.
[DOCX File , 25 KB-Multimedia Appendix 1]

References

1. Benzies KM, Aziz K, Shah V, Faris P, Isaranuwatchai W, Scotland J, et al. Alberta FICare Level II NICU Study Team. Effectiveness of Alberta Family Integrated Care on infant length of stay in level II neonatal intensive care units: a cluster randomized controlled trial. BMC Pediatr. Nov 28, 2020;20(1):535. [FREE Full text] [doi: 10.1186/s12887-020-02438-6] [Medline: 33246430]

2. Meier CA, Fitzgerald MC, Smith JM. eHealth: extending, enhancing, and evolving health care. Annu Rev Biomed Eng. 2013;15:359-382. [doi: 10.1146/annurev-bioeng-071812-152350] [Medline: 23683088]

3. Bourque P, Fairley RE. Guide to the Software Engineering Body of Knowledge (SWEBOK). IEEE Computer Society. 2014. URL: https://www.computer.org/education/bodies-of-knowledge/software-engineering [accessed 2024-07-11]

4. Hai-Jew S. Alpha testing, beta testing, and customized testing. In: Designing Instruction For Open Sharing. Cham. Springer International Publishing; 2019:381-428.

5. Ahonen EQ, Watson DP, Adams EL, McGuire A. Alpha test results for a Housing First eLearning strategy: the value of multiple qualitative methods for intervention design. Pilot Feasibility Stud. Oct 31, 2017;3(1):46. [FREE Full text] [doi: 10.1186/s40814-017-0187-y] [Medline: 29093826]

6. Athilingam P, Osorio RE, Kaplan H, Oliver D, O'neachtain T, Rogal PJ. Embedding patient education in mobile platform for patients with heart failure: theory-based development and beta testing. Comput Inform Nurs. Feb 2016;34(2):92-98. [doi: 10.1097/CIN.0000000000000216] [Medline: 26765655]

7. Cho H, Yen P, Dowding D, Merrill JA, Schnall R. A multi-level usability evaluation of mobile health applications: a case study. J Biomed Inform. Oct 2018;86:79-89. [FREE Full text] [doi: 10.1016/j.jbi.2018.08.012] [Medline: 30145317]

8. Fishbein JN, Nisotel LE, MacDonald JJ, Amoyal Pensak N, Jacobs JM, Flanagan C, et al. Mobile application to promote adherence to oral chemotherapy and symptom management: a protocol for design and development. JMIR Res Protoc. Apr 20, 2017;6(4):e62. [FREE Full text] [doi: 10.2196/resprot.6198] [Medline: 28428158]

9. Hoffimann E, Campelo D, Hooper P, Barros H, Ribeiro AI. Development of a smartphone app to evaluate the quality of public open space for physical activity. An instrument for health researchers and urban planners. Landsc Urban Plan. Sep 2018;177:191-195. [doi: 10.1016/j.landurbplan.2018.05.005]

10. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform. Jun 2019;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]

11. Wilson K, Bell C, Wilson L, Witteman H. Agile research to complement agile development: a proposal for an mHealth research lifecycle. NPJ Digit Med. Sep 13, 2018;1(1):46. [doi: 10.1038/s41746-018-0053-1] [Medline: 31304326]

12. Darlow S, Wen K. Development testing of mobile health interventions for cancer patient self-management: a review. Health Informatics J. Sep 2016;22(3):633-650. [FREE Full text] [doi: 10.1177/1460458215577994] [Medline: 25916831]

13. Yen P, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc. 2012;19(3):413-422. [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]

14. Shamsujjoha M, Grundy J, Khalajzadeh H, Lu Q, Li L. Developer and end-user perspectives on addressing human aspects in mobile eHealth apps. Inf Softw Technol. Feb 2024;166:107353. [doi: 10.1016/j.infsof.2023.107353]

15. Kelly B, Margolis M, McCormack L, LeBaron PA, Chowdhury D. What affects people's willingness to participate in qualitative research? An experimental comparison of five incentives. Field Methods. Jun 05, 2017;29(4):333-350. [doi: 10.1177/1525822x17698958]

16. Jokinen A, Stolt M, Suhonen R. Ethical issues related to eHealth: an integrative review. Nurs Ethics. Mar 2021;28(2):253-271. [FREE Full text] [doi: 10.1177/0969733020945765] [Medline: 32930041]

17. Skär L, Söderberg S. The importance of ethical aspects when implementing eHealth services in healthcare: a discussion paper. J Adv Nurs. May 29, 2018;74(5):1043-1050. [doi: 10.1111/jan.13493] [Medline: 29105151]

Abbreviations

mHealth: mobile health
NICU: neonatal intensive care unit

XSL·FO

RenderX

XSL•FO

RenderX