Usability Testing Medical Devices: A Practical Guide to Minimizing Risk and Maximizing Success

Chris Hass and Dan Berlin

Mad*Pow, Boston, MA
{chass,dberlin}@madpow.net

Abstract. This experience-based paper provides an introduction to U.S. regulations, example methodology documents, and practical advice for planning and executing medical device usability studies.

Keywords: Medical Device usability testing, procedures, practical guide, minimizing risk.

1 Introduction

Usability practitioners must be prepared to study not only websites and software but also physical devices and human interactions in a variety of settings. Rarely are the procedural and interaction stakes higher than when conducting usability research to envision, evaluate, or enhance a medical device. The safety risks for participants, the regulatory oversight, and often the scope of studies are magnified by the importance of and risks associated with medicinal delivery device interactions, populations who may have chronic or acquired medical conditions and/or diseases, and clients who view the approximations of usability research as a poor substitute for the precision of clinical trials.

Usability practitioners who are used to performing "discount" research techniques [2] may discover that these techniques are insufficient (or that client partners may deem them inappropriate) when human safety is in the balance. With this in mind, we have structured this paper to provide a practical, lessons-learned introduction to, and overview of, preparing for and conducting medical device usability testing. Our hope is that this information will inspire other practitioners and lower the learning curve associated with this important usability research arena.

2 A Meeting of the Minds

The process of conducting a successful medical device usability test begins by understanding the clients typically associated with medical device design and manufacturing. As a heavily regulated industry closely affiliated with the biological and medicinal sciences, medical device manufacturers, pharmaceutical companies, and healthcare institutions have a tendency to expect that any usability study involving a

A. Marcus (Ed.): DUXU/HCII 2013, Part II, LNCS 8013, pp. 407–416, 2013.

medical device will be conducted with exacting precision, scientific procedures, finely defined decision making, and the execution associated with clinical trials. Clinical trials may take months or years to plan and execute, and are both minutely documented and tightly regulated. Moreover, a client with a medical device interest may have invested enormous sums of money in the product being evaluated. There is little room, in their minds, for approximation.

Usability work, however, is often rife with approximation. Studies are often more loosely defined, faster to execute, may be conducted without pre-defined moderators' guides or note taking, and typically involve small numbers of participants per activity. Usability studies often stress simulation (of device use, for example) rather than field deployment. Usability practitioners are often unable to determine medical efficacy or clinical impact, focusing rather on ergonomics, safety (according to highly precise definitions), ease of use, efficiency of use, and device appeal. Furthermore, usability practitioners are rarely qualified to handle or administer medicinal agents, so conducting studies with active pharmaceutical agents can raise concerns.

It becomes contingent upon usability practitioners to recognize that some "typical" practices may seem to clients like casual or bad science - even when practitioners' experience indicates otherwise. Far from knocking the nature of usability work, when evaluating ergonomics, ease of use, efficiency, and device appeal (among other traits), usability techniques have much to offer and clients associated with medical device design and evaluation are well served to collaborate with usability practitioners.

2.1 Bridging the Usability Gap

In the United States, the U.S. Food and Drug Administration (FDA) regulates medical devices and determines whether or not a medical product or medicinal agent may be offered to the U.S. market. The FDA often requires proof of human factors/usability procedures in the development or (re)design of the medical devices it approves. In our experience, the relationship between medical device manufacturers and clients who "don't speak usability" can be at best confusing and at worst tempestuous.

As a result, practitioners are well advised, especially where regulatory bodies are concerned, to learn about the client's relationship to these governing bodies. It may be advantageous to ask:

- What are the relevant regulations governing this product?
- Who is the client's regulatory liaison?
- What is the client's relationship to regulating bodies (such as the FDA)?
- Has the client met with them? If so, when? What transpired?
- Have they received Comments (written feedback or directives)?
- What criteria will the regulators use to evaluate the usability study's validity and findings?
- Does the client understand usability research?
- Is it permissible to talk directly with the regulatory bodies?

There are other agencies that can help bridge the gap between usability research and clinical trials as well. Internal Review Boards (IRBs) can be invaluable (and sometimes required) resources to help practitioners and clients define regulation-

appropriate methodologies and to select research best practices. IRB involvement can also boost confidence among clients that a usability team's seemingly "casual" approaches are valid and based on industry best practices. IRBs may be found within research organizations or hired on a consulting basis to provide document review, procedural oversight, and ultimately to certify that research conducted with human participants is carried out properly and humanely.

Generally, we have found within the U.S. that engaging an IRB to review documentation adds roughly $2,000-$5,000 in costs to a study. Practitioners may expect to add the IRB's review time to a project timeline, and should be cognizant of any non-disclosure agreements in place before sharing details of the study with external partners. IRBs typically require formal document versioning and use signature pages to ensure that all parties reviewing a document are identified. They may physically inspect your usability lab and study setup.

It is especially important to note that once an IRB has approved a document or series of documents (such as a Usability Specification, consent forms, etc.) they may NOT be amended or altered during the course of the study without the changed documents being re-reviewed and approved by the IRB. This can come as a surprise for practitioners who may be used to adjusting study methodologies during a study to streamline session time, improve procedural clarity, or adjust recruitment criteria. IRB-approved documents are often stamped to identify the final approved version.

3 Planning

Communicating the value of this collaboration may prove challenging. This communication happens not only in client discussions, but through the generation of precise documentation. In the planning stages of a medical device usability study we recommend the generation of "Usability Specification" documents that define and present, among other things:

- Study goals and methodologies
- Recruitment screener text and images
- Recruitment flyer text and images
- The moderator's guide including: tasks, their associated steps; what should be observed; questions to be asked of participants; observer ratings scales, tally methods.
- Consent and Assent forms for adults and children, respectively
- Safety concerns outlining potential risks associated with each study activity
- Steps to be taken to ameliorate or reduce risks
- Procedures for safeguarding participants in the event of accident or incident
- Procedures for identifying and reporting adverse events
- Device background detailing the purpose of the device to be evaluated including salient device features
- Device instructions for use (IFU) if they are to be included in the study
- Explanations of how the study will validate the device's usability including:

- The rationale behind every quantitative metric used including ease of use ratings scales, instructions for use clarity, design, ergonomics, observed data such as usability, help needed, number of errors, *etc*. (see Fig. 1)
- Definitions of use errors, how they will be identified, noted, and prioritized
- Each question to be asked, metric to be used, how it will be tallied/captured, and how it will be analyzed, interpreted, and presented
- Detailed task breakdowns of each action associated with utilizing the device
- Detailed explanations of data NOT to be captured with exclusion rationales

- Notes grid(s) for consolidating data that include:
  - Clear delineation of "micro tasks" involved in device use
  - Each task step and question on its own spreadsheet row
  - Clear identification of study sections (Introduction, Task 1, Task 2, etc.)
  - Greyed out areas that may be ignored
  - Identification of roles (*e.g.*, a moderator will focus on the participants' device use, a note taker will capture participant comments)

Rating	No Help Needed	Little Difficulty	Some Help	Much Help	Task Failure
Explanation	Completed with no help	1-2 small errors	3-5 errors	Moderator gets involved	Failure
Small errors:	Doesn't press green button Doesn't remove cap properly				

Fig. 1. Example Participant/Moderator Help Metric and Error Definition

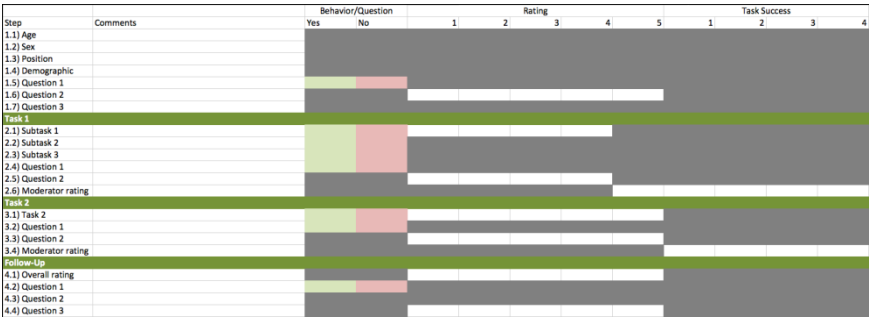


Fig. 2. Example Notes Grid

These documents are frequently required and are typically referred to as a "Usability Validation" (Risk assessment and amelioration) document and a "Usability Engineering File" or "Design Efficiency File" (Methods and procedural rationale).

Typically these documents are informed by Federal and international regulations related to the design and evaluation of medical devices. For US medical device studies the following documents are highly recommended:

- FDA Quality System Regulation (US FDA, 2011) [5]
- HE: 75 2009 (AAMI. 2009) [1]
- ISO/IEC 62366:2007 (ISO, 2007) [3]
- The Health Insurance Portability and Accountability Act of 1996 (HIPAA) (US H&HS, 1996) [6]

Both established and emerging standards are essential reading and clients will often expect a thorough understanding of them. Practitioners should expect that the drafting, review, and ratification of usability specification documents may be a significantly time-consuming process and involve many parties including: the client, regulators associated with the client's work, Internal Review Board(s), compliance personnel, Adverse Event personnel, other subcontractors, clinical personnel, subject matter experts, and legal teams.

However, given that these usability specification documents provide a precise roadmap for all pre-, intra-, and post-research activities, they are well worth the focus, collaboration, and foresight required to shape and execute a reliable, unbiased, and repeatable study. Identifying risks associated with the study, for example, is an invaluable activity, given that safety protocols may be exacting: even simulated product use may put participants at risk, target audiences may be physically or cognitively vulnerable, activity environments may be unusual for practitioners (a care clinic, a patient's home), discussions of patients' health may become personal or emotional – and defining appropriate safeguards may involve consultation with clients and medical care specialists.

When preparing for data analysis during the creation of the Usability Specification documents, it may be helpful to think in terms of a "micro task" scale: Every individual step of using the device should be observed and recorded. However, it's difficult to watch and report on every step of a complex interaction and to document it perfectly, especially when dealing with a device that participants are manipulating quickly. Tallying procedures need to be streamlined and specific, but also realistic. Pilot testing potential approaches will provide invaluable insights for practitioners and their clients to establish reasonable data capture, and therefore data analysis approaches.

When planning, it is also helpful to be mindful of capturing and organizing the most salient qualitative data. We recommend summarizing qualitative data immediately after a given study session, before the moderator and note taker's memory of the session fades.

4 Recruiting

Recruiting participants for medical device usability studies can pose unique challenges. These are understandably related to the type of device to be evaluated, the

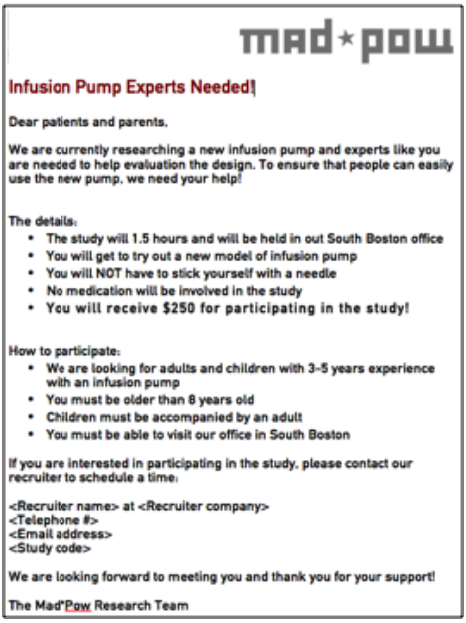
nature of its use, the number of participants, and the health of the target participants. Medical device usability tests frequently involve moderately large sets of users, in our experience from 50 to 150 participants is not unusual. In addition, users of medical devices may be elderly, infirm, may have more than one disease, condition, or contributing factor (co-morbidity) that affects their treatment regimen and/or use of a medical device.

The criteria associated with identifying users of a specialized device, or even one intended for the general public may be complex and may take significant time to identify and enumerate with clients. Notifying a recruiting partner as early as possible that a study may require patient participants with a specific diagnosis is extremely helpful, so that partner can begin laying groundwork for reaching out to potential study participants. Moreover, identifying patient populations in significant numbers may be difficult, involve unforeseen travel, or require higher honoraria to ensure participation. Recruitment partners, such as external recruiting agencies may also have insights into geographic regions where a greater population of the target audience may be found. Recruiting specialized populations may take significantly longer in the medical arena and can cause unexpected study delays. As always, we recommend open communication with clients about outreach efforts, organizations contacted, and any difficulties encountered. In many cases clients may be legally unable to directly recommend clinicians, institutions, or patients associated with a particular disease, condition, or disability.

When recruiting patients it can be helpful to cultivate relationships with local acute and ambulatory care organizations, as can reaching out to groups that specialize in serving the audience you are seeking. These may include: healthcare institutions, local and national support groups, federal agencies, community support groups, social media, and online crowd-sourcing sites. When reaching out to potential participants, you may discover that acute and ambulatory care providers may be unable to "promote" a study that is not associated with their organization, which can limit access to potential participants. We have also encountered situations where clinical practitioners request a "referral fee" for sharing recruitment flyers with patients. While this appears to be a relatively common practice, it is illegal in the United States and not recommended.

In studies where an IRB is involved, recruitment flyer and screener text will need to be defined early in the project and be approved by the IRB before recruitment may begin. It is worth noting that once a study is in progress, any change in recruitment screener criteria (or methodology) will have to be documented along with a rationale for the change then re-submitted to the IRB for approval. This can add time delays and in some cases additional costs. IRBs will often "stamp" an approved recruitment flyer (and consent form) and the stamp must be present on documents used to support the study, indicating that the documents have not been altered.

The contents of a successful study flyer (to be distributed to clinician's offices or electronically) will include: the purpose of the study, why potential participants would want to participate, compensation details, and recruiter contact information.



mad*pow

Infusion Pump Experts Needed!

Dear patients and parents,

We are currently researching a new infusion pump and experts like you are needed to help evaluation the design. To ensure that people can easily use the new pump, we need your help!

The details:

- The study will 1.5 hours and will be held in out South Boston office
- You will get to try out a new model of infusion pump
- You will NOT have to stick yourself with a needle
- No medication will be involved in the study
- You will receive $250 for participating in the study!

How to participate:

- We are looking for adults and children with 3-5 years experience with an infusion pump
- You must be older than 8 years old
- Children must be accompanied by an adult
- You must be able to visit our office in South Boston

If you are interested in participating in the study, please contact our recruiter to schedule a time:

<Recruiter name> at <Recruiter company>
<Telephone #>
<Email address>
<Study code>

We are looking forward to meeting you and thank you for your support!

The Mad*Pow Research Team

Fig. 3. Sample Recruitment Flyer

5 Conducting the Study

5.1 Participant Safety

As with any usability study, safeguarding participants' comfort and safety is of paramount importance. Where medical device usability studies are concerned practitioners should be especially cognizant of participants' health. Will participants need accessibility accommodations? Will their condition/disease affect participation? Will they arrive with an aide or caretaker? Moreover, we recommend ensuring their physical safety by having a robust first aid kit (and in some cases medical personnel) nearby, making water, juice, sugary, and non-sugary snacks available, ensuring that they know whether they will be working with a placebo rather than a medicinal agent, and providing accessible food that may be easily grasped and consumed. Most importantly we establish a plan of action ahead of time for the (hopefully unlikely) event someone is injured. This plan should include a path of notification in the event of injury: reporting it internally, to the client, the IRB, Adverse Event personnel, a regulatory agency (as appropriate), and any other relevant governing agencies.

To prevent participant injury, stress, dehydration, confusion, disorientation, and fatigue, knowledge of the relevant diseases/conditions associated with the medical device is invaluable. Advance reading, interviewing clinical personnel, and consulting subject matter experts may help practitioners to better safeguard their participants.

5.2 Usability Lab Setup

Medical device usability testing can take place in a formal usability lab, on-site at a clinic or care facility, in participants' homes, or any other relevant location. Without delving too deeply into technical setups, we have found the following rules of thumb helpful when conducting medical device usability testing of a physical device:

- Set up cameras to enable data capture from two different angles
  - A close-up view of the device interaction (over the shoulder or capturing hand interactions)
  - Directly in front of the participant to capture the entire field of interaction (including the participant's face as appropriate)
- Have all device supplies at the ready, within reach, and neatly organized
  - Alcohol wipes, extra needles, batteries, loading cartridges, etc.
  - You're likely to need more supplies than you would initially expect
- Have a first-aid kit and nitrile gloves nearby (some people are allergic to latex)
- Have a first aid plan and practice enacting it

5.3 Adverse Event Reporting

One procedural aspect that may be novel to practitioners unfamiliar with medical device usability testing is Adverse Event reporting. An Adverse Event (AE) occurs when a study participant shares information about a side effect or other unusual or negative experience related to device or drug use. Such information must be reported to clients and regulatory bodies immediately. Encountering and reporting an AE in the US is a tightly regulated process. AEs are critical information for medical institutions, pharmaceutical companies, and the FDA. Identifying an AE can be complex and it is not unusual for clients to require usability practitioners to participate in AE training prior to the commencement of research activities. Practitioners should identify to whom AEs should be reported (most often to the client, who will notify their regulatory agencies). A valuable rule of thumb when determining whether a bit of information is an AE or not is: "When in doubt, report it." Typical AE reports may be as simple as an email or phone call communicating:

- The comment or action comprising the AE
- A general description of the scenario where the AE was discovered
- The device or medication involved
- The date and time of the AE
- AE physical symptoms and corrective action taken
- Whether there was injury to a person or not
- Official study codes

Some examples of adverse events are:

- A usability study participant reports that she stopped taking a drug because the blister packs and pills sometimes arrived crushed.
- While testing an infusion pump a participant mentions that the tube "always fell out of" his previous pump (the client's product).

- During a study a participant mentions that her arms started to itch when she stopped taking the client's medication.
- A male participant mentions that his wife became pregnant two weeks after he started taking the client's medication, even though she used birth control pills.

6 Analysis and Reporting

6.1 Data Analysis

One clear benefit of the up-front time spent defining the Usability Specification documents at the start of the study is that if data capture methodologies are well conceived and executed, data analysis should be very straightforward. The process becomes one of distilling the notes grid into summary findings. We cannot understate the value of a cogent data capture methodology. Without it, inevitable data capture inconsistencies and human error may cause significant data analysis difficulties, particularly with clients versed in the precision of clinical trials.

Frequent questions pertaining to missing data include: Was the question asked? Was the behavior observed? Will the recordings verify or find data? (In this case it is vital that you HAVE recordings, of course, and that they offer clear images and sound.) It is not unusual for a client seeking regulatory approval to insist on a perfect notes grid, *i.e.*, that it is entirely complete, reliably reflects the study sessions, and contains no missing data.

For practitioners used to more approximate data capture, this may come as a shock. The amount of time a client may require practitioners to invest in ensuring the perfection of the notes grid is difficult to underestimate. When conducting study research sessions, ensuring that every question is asked, every observation is recorded correctly, and that each session is executed uniformly may be exceedingly difficult. But, this consistency is without a doubt the best defense against a lengthy data analysis phase.

When analyzing quantitative measures practitioners typically benefit from capturing data that answers the following questions:

- What percentage of participants successfully completed the task?
- What ease of use ratings did participants give?
- How many participants committed the most common errors?

6.2 Reporting

Defining reporting requirements up front in the usability specification can bring great benefits as reporting guidelines, formats, and industry standards (e.g., CIF a.k.a. ISO/IEC 25062:2006, other specialized formats [4]) may be extensive and/or complex and raw data may be reviewed many, many, many times by clients paying careful attention to methodology and procedural precision. Moreover, once the Usability Specification documents have been created, reviewed, and ultimately approved, in studies governed by a formal regulatory body, often they may not be changed during the course of the study without being re-examined and approved by all contributing

parties, be they client, Internal Review Board, or other regulatory agency. As a result, pre-defining the study risks, procedures, data solicitation and capture methodologies, and the like will often be required.

In our experience, clients sometimes express interest at the close of a study in using the positive study findings as marketing claims, which can raise concerns. To avoid later confusion, be sure to discuss potential data outcomes and follow-on data uses early in the study.

When reporting findings, the reports themselves may be relatively straightforward: a cogent walk-through of the methods utilized and summaries (graphical, statistical, and textual) of the study findings and the researchers' conclusions. However, some clients and regulatory bodies require reports in standardized industry formats (e.g., CIF) or that match existing product documentation. Determining in advance the preferred reporting format again, will save confusion, time, and effort.

7 Conclusion

Conducting a usability study for a medical device is similar to more "typical" studies, except with much stricter protocols and less room for procedural flexibility. Practitioner and participant safety remains paramount, and methodologies that focus on quantitative measures (reporting percentages of participants that did/did not perform an action) are likely to be successful. Careful planning, knowledge of national and international regulations, and open communication with client partners will ensure a smooth and effective research study.

References

1. Association for the Advancement of Medical Instrumentation. ANSI/AAMI HE75:2009: Human Factors engineering – Design of Medical Devices. AAMI, Arlington (2009)
2. Dumas, J.S., Redish, J.C.: A Practical Guide to Usability Testing. Intellect, Portland (1993)
3. International Standards Organization. Medical Devices – Application of usability engineering to medical devices. Geneva, Switzerland: International Electrotechnical Commission (2007)
4. National Institute of Standards and Technology, Common Industry Format for Usability Test Reports. Washington, DC: U.S. Government Printing Office (1999), http://zing.ncsl.nist.gov/iusr/documents/cifv1.1b.htm (retrieved) United States Food & Drug Administration, Quality System (QS) Regulation/Medical Device Good Manufacturing Practices (2011), http://www.fda.gov/medicaldevices/deviceregulationandguidance/postmarketrequirements/qualitysystemsregulations/default.htm (retrieved)
5. United States Department of Health & Human Services. Health Insurance Portability and Accountability Act of 1996. (US H&HS publication f:publ191.104). Washington, DC: U.S. Government Printing Office (1996)