Innovations in Technology in the Development and Collection of Patient-Reported Outcomes

Hilary Wilson, PhD, *Senior Research Associate, Outcomes Research;*
Sonya Eremenco, MA, *Director ePRO New Products, Outcomes Research;*
Karin Coyne, PhD, MPH, *Senior Research Leader and Scientific Director, Outcomes Research*

Advances in technology have had a significant impact on data collection in all phases of the drug development process, including the process of developing and implementing a patient-reported outcome (PRO) measure. Although the most frequently noted technological advance in the data collection of PROs is the adoption of electronic PRO (ePRO) devices, which allows for real-time collection of patient outcomes, there are also technological approaches in other phases of the PRO instrument development process. The purpose of this article is to briefly review and consider the innovative approaches to data collection for both qualitative and quantitative data used in the development and validation of PROs.

HYPOTHESIS GENERATION OF THE CONCEPTUAL FRAMEWORK: NATURAL LANGUAGE PROCESSING OF WEB CONTENT

The PRO instrument development process typically begins with a literature review in the therapeutic area of interest to inform the hypothesized conceptual framework for the relevant concept of interest. Often interviews with therapeutic area experts provide content expertise to refine the hypothesized conceptual framework. Analysis of Web content (e.g., patient blogs) through natural language processing or other qualitative software analysis approaches may offer an additional complementary tool to inform the hypothesized conceptual framework in this early concept elicitation phase of PRO development.

At the most basic level, an analysis of Web content using natural language processing provides a summary of the frequency of various text fields within a selected sample of text. Examples of sample text include generic blog and microblog sites, such as Wordpress[1] and Twitter[2], patient support forums or Facebook pages that are organized by relevant patient interest groups,[3,4] patient-centric platforms designed specifically for patients to connect,[5,6] or more broadly all Web content available through a keyword search (i.e., "fibromyalgia") on a search engine. More complex models provide interpretative structure to the text data, and the more sophisticated software applications include data visualization approaches to summarizing the frequency and structure of the text data. Several commercial and open source software applications are available to perform text analysis through natural language processing of Web content (see link for text mining software examples: http://en.wikipedia.org/wiki/List_of_text_mining_software).

Figure 1 presents a foam tree diagram generated through a Web content analysis using Carrot[2,7] an open-source Web content analysis tool using the search term "fibromyalgia." In this case, the sample is the 86 most relevant of 2,380,000 website hits through the Carrot[2] search engine, and as evident by the diagram, the most common content addresses "treatment of fibromyalgia." From the perspective of early concept elicitation for a fibromyalgia symptom questionnaire, further examination of the diagram highlights the potential importance of muscle pain, sleep disturbance, soft tissues, and the possible relationship with chronic fatigue and arthritis. This example is provided to demonstrate the type of exploratory analysis that might be conducted. A natural language processing-based content analysis of a more select sample of text— for example, fibromyalgia patient forums or blogs written by fibromyalgia patients—may provide insights more directly relevant to a symptom measure.
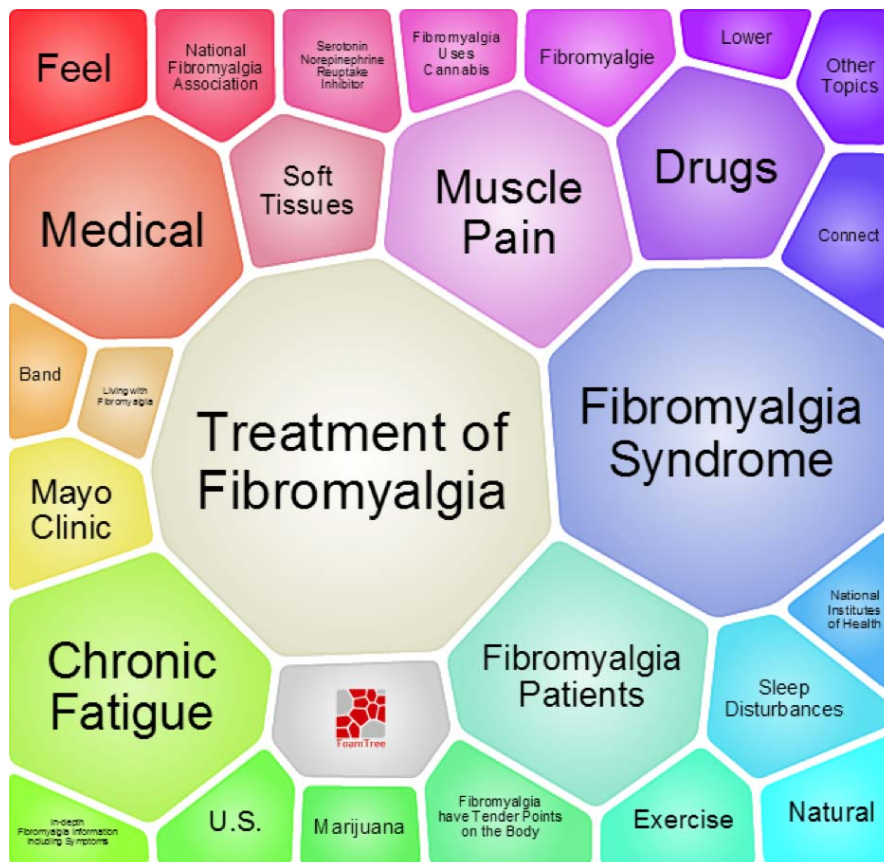
figure 1—*Extracted on April 15, 2014 (Source: Weiss D, Osinski S. 2014)[13]*

What advantage does analyzing diverse Web content provide to early stages of concept elicitation beyond what a structured literature review and therapeutic and patient experts might provide?

- One key advantage is that due to the emergent nature of the data, there is no unintended influence of the interaction between the researcher and participant responses. Patients may be more authentic in speaking about sensitive issues and may discuss issues that they would not discuss with their clinicians.

- Given the potential for analysis of large amounts of data, concepts or domains that are relevant to only a subset of participants that may be missed in a more focused qualitative analysis of only a few patients may be uncovered.

- The approach may be particularly useful in an indication where there is limited existing evidence characterizing the disease, or in cases where the disease is rare and access to patients is limited.

- Published manuscripts and the opinions of therapeutic area experts introduce their own bias, and by combining these approaches with a patient-centric, conversational data element, a richer picture of the conceptual framework may be realized.

While there are a variety of advantages, the following limitations should also be considered with this approach:

- As of 2012, over half of American adults aged 65 years and older are online and growing, but despite increasing socio-demographic reach of Internet use, the online population is still biased in favor of young, educated, and white participants[8].

- Sample bias may also be introduced by the selection of materials to be analyzed (e.g., differences among Twitter users relative to those in a specific disease forum).

- The importance of certain concepts may be over-estimated by one or two users or websites that focus on a specific concept.

- The analysis is a combination of computationally driven analysis of the text and user guidance, so it is subject to interpretation bias and error that is introduced as part of the models used in the natural language processing software.

- As an emerging field, there is no standard analytic approach or guidance related to ethical considerations surrounding the privacy and confidentiality of analyzing this type of data.

Despite these challenges, the approach offers an efficient way to gather information related to a concept of interest, and may offer a unique perspective that is not readily available through the traditional literature reviews, patient surveys, and qualitative interview approaches used at this stage in PRO development.

NOVEL APPROACHES TO ENGAGING PATIENTS FOR CONCEPT ELICITATION AND COGNITIVE INTERVIEWING

In-person interviews are the "gold standard" in qualitative research, however this approach is arguably the most expensive and time-consuming, and often recruitment is limited to narrow geographic locations. Telephone, video-conferencing, and Web-based interview mediums are alternatives that offer potential cost and time savings and broader geographical reach.

Telephone interviews may be conducted at a considerable cost-savings, and may also allow access to geographically disparate subjects;[9] however, despite these advantages, they are less frequently utilized than face-to-face interviews in qualitative research.[10] The primary limitation noted for this modality is the lack of visual cues, which is perceived to lead to the loss of important nonverbal and contextual data, although the empirical evidence for this frequently cited rationale is lacking with limited comparison of telephone and in-person qualitative interview modalities.[10]

Video teleconferencing provides the added benefit of observation of facial visual cues for both the participant and interviewer. However, most often the image only covers the head, so other body language is not observable. Either audio, or both audio and video, may be recorded dependent on the software that is used. Participants may travel to a video-conferencing center, which may be available and rented on an hourly basis, or alternately they may utilize at home, high-speed Internet, Web browser, and HIPAA-compliant video-conferencing software.[11] The availability and accessibility of HIPAA-compliant video conferencing software does impact the socio-demographic reach of this medium relative to teleconferencing. However, in cases where face-to-face interviews might offer an advantage in rapport development and observation of facial cues, it is an important medium to consider.

Online forums created for the purpose of the research study, where participants can discuss specific topics through posting a series of messages and a researcher moderates the discussion is another, although little used, Web-based option for conducting qualitative research.[11] The posts made by participants are the unit of analysis and are analyzed in a similar fashion to transcripts. Participants may be recruited from other relevant Internet communities, or through community settings, or alternately a patient-centric research platform designed specifically for patients to connect with researchers, such as Patients LikeMe.[5] In contrast to other mediums, these types of interactions are asynchronous—such that participants log on at different times and dialogue is not conversational at a set point in time. This offers an advantage to collection and participation across geographical time zones, but it also does not allow for prompt and immediate response, which may be considered a disadvantage in some contexts.[10] Given the sensitivity of the health information that is being discussed, security measures should be considered, although
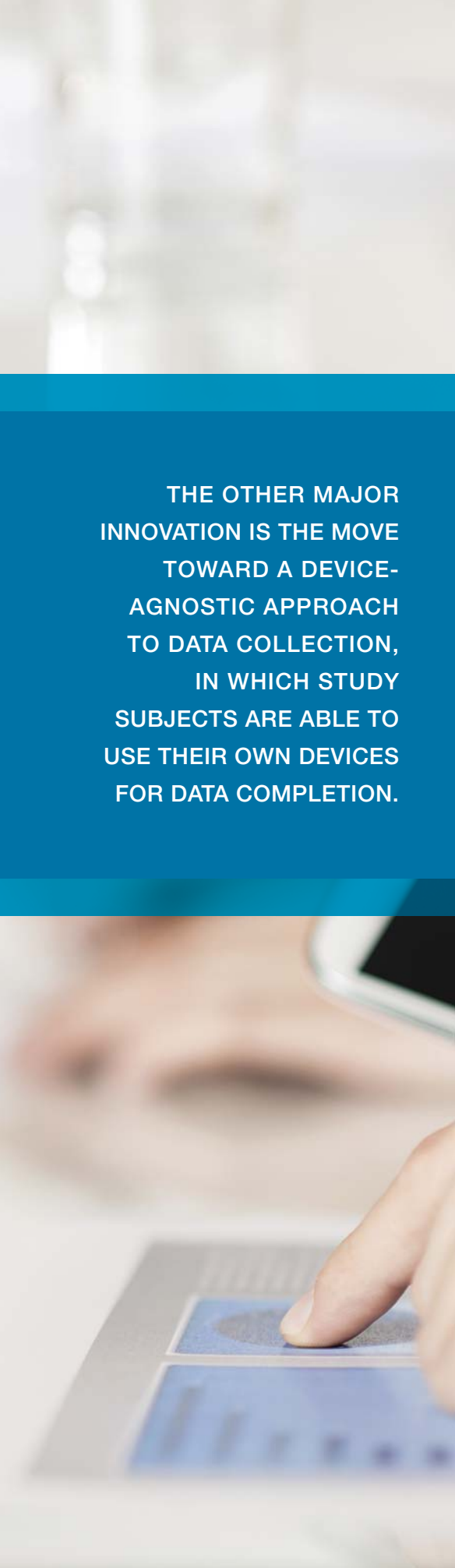
no clear guidelines exist around this issue. The inability to confirm diagnosis is a further limitation to the documentation of content validity in the target population of interest. However, it is possible for patients to consent to the release of their medical records for diagnosis confirmation.

WEB-BASED DATA COLLECTION

Historically, ePRO technologies have included personal digital assistants (PDA), interactive voice response (IVR), Web-based systems, smartphones, tablets, purpose-built devices such as peak flow meters with integrated diary capabilities, and digital pens. As technology permeates every aspect of daily life, further innovation in ePRO is taking place in Web-based data collection and in "device agnostic" data collection, also termed "Bring Your Own Device" (BYOD).

Web-based data collection has been associated with large screen devices like desktop or laptop computers. Thus, Web-based questionnaires were designed with these browsers and screen sizes in mind, and assumed peripherals including a keyboard and mouse were available for response entry. However, mobile devices are now capable of accessing the Web through specific mobile Web browsers, which have a very different look and feel compared to large screen devices and require touchscreens or navigation buttons for response entry. Therefore, the Web-based approach has been more broadly defined to include the use of a wide range of devices with access to the Internet, including mobile browsers.

Smaller, mobile Web-browsing devices have the portability of a smartphone or PDA device, and may provide larger screen sizes which allow for longer questions, longer responses, and can accommodate translations more easily than smaller handheld devices.

Web-based data collection requires that the user interface be optimized to work with and be validated on a combination of the operating system (i.e., Windows, iOS, Android) and the browser (i.e., Internet Explorer, Firefox, Chrome, Safari, etc.), but device-specific validation is not required.[12] In most cases, a choice has to be made to optimize the interface for a larger screen browser or a mobile browser. Another concern is the need for uninterrupted access to the Internet during questionnaire completion. In some cases, the questionnaire can be saved and resumed later, but unexpected interruptions can lead to loss of data already entered and would require the study subject to start over from the beginning.

BRING YOUR OWN DEVICE (BYOD) APPROACH

The other major innovation is the move toward a device-agnostic approach to data collection, in which study subjects are able to use their own devices for data completion. The BYOD approach appeals to study sponsors because it reduces the cost of providing devices to all study subjects, the logistics of deploying devices internationally, training issues, maintenance and help desk issues during a study,[12] and the need to maintain the devices after study completion. Study subjects who have their own devices may prefer to use a device with which they are familiar rather than carrying a second device around with them. Two approaches to BYOD are currently in use: "Apps" and Mobile Web.

"APPS"

An "App" version of the PRO questionnaire can be downloaded to the study subject's own device to be accessed for data collection during a trial. The App is programmed to work on a specific operating system, most prominently Apple's iOS or Android. Advantages include consistency in display across devices within a given operating system, the ability to answer the questionnaire offline and then transmit the data when completed, and using the device's own alarm feature to remind the subject to complete the questionnaire, critical for daily diaries with limited completion windows.

The main disadvantage to the App approach is the need for a compatible smartphone that can accept the App and the need to download it to the device. The subject's device must be assessed to ensure it has the right operating system version and screen size, which puts the burden on sites to determine if a study subject's device is acceptable. Provisioning backup devices to subjects who do not have a phone or compatible device must be considered. Security and privacy are also major concerns. The App must be 21CFR Part 11 compliant and requires validation on every type of mobile phone, tablet and computer used in the trial.[12] Data entered on the device may not be as secure as on a standalone device because the patient's own device is used for many other purposes. There are also concerns regarding data loss if the device is lost or fails.[12] Finally, there are cost considerations because subjects must pay for the data transmission using their own mobile service plans, while in traditional ePRO these costs are covered by the sponsor.

MOBILE WEB

Advantages of the Mobile Web approach to BYOD are that accessing the questionnaire is much simpler as only a link to the website is needed, there is "zero footprint on the patient device and no need for local installation,"[12] and no data reside on the device as it is merely an interface to access the Web-based browser. Device-specific compatibility may be less of an issue although the questionnaire still needs to be optimized to work with Mobile Web

> THE OTHER MAJOR INNOVATION IS THE MOVE TOWARD A DEVICE-AGNOSTIC APPROACH TO DATA COLLECTION, IN WHICH STUDY SUBJECTS ARE ABLE TO USE THEIR OWN DEVICES FOR DATA COMPLETION.

browsers in general. A broader range of devices may be used with Mobile Web than with the App-based approach.

The need for constant Internet connectivity is a major disadvantage because it is required to access the questionnaire initially, and mobile access can drop suddenly during questionnaire completion.[12] Variability across devices and screen sizes is also a concern; it is impossible to test all possible variations of browsers and devices to ensure that the instrument displays consistently. Although some question the need for device specific validation,[12] different screen sizes may lead to incorrect responses if text is not visible on the screen without scrolling. The subject must also have a mobile data plan and therefore has to bear the cost of accessing the Internet to participate in the study. Reminders in the Mobile Web approach may be sent via email or text messaging/short messaging service with a link to access the system, but the audible approach of an alarm on the device is not as feasible. Therefore, the reminder could be easily missed if the subject is not near the device, resulting in lower compliance due to inadequate reminders.

CONCLUSION

This year (2014) marks the 25th anniversary of the World Wide Web.[8] With the advent of Web 2.0 in the last decade—a medium which allows users to interact and collaborate with each other, versus passively consume Web content—people are now able to engage in ongoing, interactive dialogue through various social media networking sites, blogs, and communities. The changing landscape of Internet access and engagement is shaping health care and the research process. Technological innovations in data collection have the potential to improve and streamline the PRO development process, from hypothesis generation to data collection in clinical trials, and to facilitate patient engagement on many levels. However, when considering newer technology options, it is important to consider some general limitations noted above. Perhaps even more importantly, clear guidelines and approaches to managing the privacy and security of these Web-based approaches are needed. When considering novel approaches to data collection, it is also important to balance the costs, sampling bias, logistical challenges and the patient's desire for convenience against privacy and security concerns.

For more information, please contact Hilary.Wilson@evidera.com, Sonya.Eremenco@evidera.com or Karin.Coyne@evidera.com.

References

1 Wordpress. Wordpress. [Cross-platform]. 2003; 3.8.3:http://wordpress.com/.

2 Twitter. [Social network service, microblogging]. 2006; www.twitter.com.

3 Oral Cancer Foundation Support Forum. [Article]. 2001; http://www.oralcancerfoundation.org/forum/.

4 NET Patient Foundation Forum. [Online Forum]. 2006; http://www.netpatientfoundation.org/forum/.

5 PatientsLikeMe. [Social Media]. 2004; http://www.patientslikeme.com/.

6 The Patient Forum. [Online Forum]. 1996; http://www.thepatientforum.com/.

7 Osiński S, Weiss D. Carrot[2] User and Developer Manual. 2013; http://download.carrot2.org/head/manual/.

8 Zickuhr K, Madden M. Older adults and internet use. 2012; http://www.pewinternet.org/2012/06/06/older-adults-and-internet-use/.

9 Sturges JE, Hanrahan KJ. Comparing Telephone and Face-to-Face Qualitative Interviewing: a Research Note. *Qualitative Research.* 2004 Apr; 4(1):107–118.

10 Opdenakker R. Advantages and Disadvantages of Four Interview Techniques in Qualitative Research. Forum: *Qualitative Social Research.* 2006 Sep; 7(4). At http://www.qualitative-research.net/index.php/fqs/article/view/175/391. Accessed April 18, 2014.

11 Im EO, Chee W. An Online Forum as a Qualitative Research Method: Practical Issues. *Nurs Res.* 2006 Jul-Aug; 55(4):267–273.

12 Yeomans A. The Future of ePRO Platforms. *Applied Clinical Trials Online.* 2014 Jan 28. At http://www.appliedclinicaltrialsonline.com/appliedclinicaltrials/article/articleDetail.jsp?id=833920&pageID=1. Accessed April 18, 2014.

13 Weiss D, Osiński S. Carrot[2] Foam Tree-Fybromyalgia Search. [Search Results Clustering Engine]. 2014; http://search.carrot2.org/stable/search?source=web&view=foamtree&skin=fancy-compact&query=fibromyalgia&results=100&algorithm=lingo&ETools DocumentSource.country=ALL&EToolsDocumentSource.language=ALL&EToolsDocumentSource.customerId=&EToolsDocumentSource.safeSearch=false. Accessed April 15, 2014.