

ePROs: Practical Issues in Pen and Touchscreen Systems

A number of factors influence the successful electronic collection of patient data, including screen size.

Mar 2, 2007

By: [Brian Tiplady](#)

Applied Clinical Trials

Electronic methods such as electronic patient reported outcomes (ePROs) are in increasing use as a tool for collecting diary and questionnaire data directly from patients. There are many advantages to using such methods: only valid, in-range data options can be entered; missing data within a questionnaire can be prevented; feedback and reminders can be given to patients to help them comply with study procedures; all entries are date and time stamped, so it is not possible for patients to back-fill diary data; data can be immediately transmitted to a secure central server; and, finally, question branching can be invisible to the patient, making it easier than paper questionnaires.



Despite these clear advantages, concerns about the use of electronic systems have been raised in two main areas. The first is user acceptance, in particular with patients who are unfamiliar with computers, some of whom may be uncomfortable with the idea of using such systems. The second area of concern has to do with the equivalence of electronic versions and standardized paper methods.

Electronic reporting

There are a number of ways that ePRO systems may differ from paper. Computer screens are generally smaller than sheets of paper, limiting the amount of information that can be presented to the patient at any one time—and moving between pages on a computer can be less straightforward than turning over paper sheets. On the other hand, changing a selected option on an electronic scale is simpler and neater than with paper.

A number of studies have evaluated ease of use and acceptability of handheld electronic diaries. It has generally been found that electronic data collection systems are highly acceptable to patients, who in many cases prefer them to paper systems. This was found to be true for all patient populations, including older patients, patients with no previous computer experience, and younger patients familiar with computers and mobile phones.^{1,2}

Comparisons between paper and electronic versions of patient self-report scales have generally showed equivalence of the data collected from the different versions.³ Achieving such acceptability and validity requires careful design and configuration. This article will highlight some of the important issues that should be addressed when designing or specifying an ePRO system, including selection of an appropriate device, the different issues that arise with diaries and questionnaires, and ways of avoiding potential bias in the data collection process. The focus here will be on computer systems with touch or pen screens.

PCs, handhelds & other devices

There are two main types of computers with pen or touch interfaces available. The first is the tablet PC, which is similar in size and weight to a small laptop (see Figure 1a). These are standard PCs as far as hardware and software are concerned.

At the other end of the size range are palm-top or handheld devices (see Figure 1b). They are not full PCs, having much smaller processing and memory capabilities. However, they have ample computing power to run patient diaries and questionnaires, and can support wireless or wired data transfer. One limitation is screen size, typically 8 to 9 cm diagonal (about 2.5 inches), but in some cases considerably smaller. They are, however, very light and portable, and much less expensive than tablet PCs.

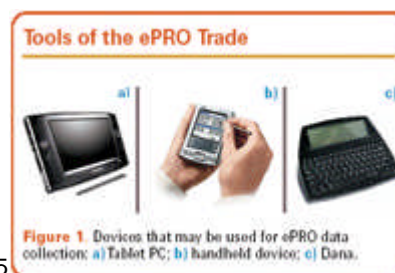


Figure 1: Tools of the ePRO Trade

There is very little on the consumer market between these two extremes. There are, however, specialist devices targeted toward educational or vertical markets that are intermediate in size and also in price; an example is shown

in Figure 1c. This has a screen about twice the size of the larger handheld devices and accepts both pen and keyboard input. When used for ePRO, the patient would use only the pen.

Screen layout

PRO instruments use a variety of question types with different frequencies of assessment. For example, many patient diaries use daily assessments (or assessments several times a day) with short and simple questions. Quality of life (QoL) assessments, on the other hand, are typically carried out in the clinic, use lengthier questionnaires, and are administered less frequently. (An example of a question from a patient diary is shown in Figure 2.) Another type of question used in diaries is the visual analogue scale. These are generally shorter on handhelds than on paper, but equivalence studies have shown comparable results.^{4,5} Even on smaller, palm-size devices, patients have no difficulty with these types of diary questions.

In other situations, screen space is of vital importance. Many quality of life scales include definitions of terms, which may take up substantial amounts of space. For example, a rating of heartburn may specify the location so that the pain can be distinguished from other symptoms. (Without these definitions and clarifications, the validity of the items is suspect.) Generally, such questionnaires have more response options, often seven, which may be specified in more detail, adding to the amount of text that must be displayed (see Figure 3).

Clearly, there are much greater challenges in setting up such questionnaires on small devices than the layout shown in Figure 2, and these challenges are critical, as it is important that all types of patients are able to use ePRO systems. If not, the population included in the research may be biased, and thus not representative of the population under consideration.

ePRO Systems for all

We like to think that technology has become pervasive in our society, and we cite senior citizens who keep in touch with their grandchildren by text messaging and email.⁷ These people are real enough, but there are also substantial numbers who have never used a computer before.² So we have to make sure that everyone can use ePRO systems effectively, irrespective of their familiarity with computer technology.

Experienced computer users often underestimate the task facing novices, and regard things like scroll bars, drop-down menus, double-clicking, and so on as intuitive. But for someone encountering them for the first time, these things are not in the least obvious, and it may take considerable time to become familiar with them and to use them effectively.

Question response

The information needed to answer a question has a number of components, including the question itself, the available responses, any necessary explanation or clarification of terms, and information applying to the whole questionnaire or to groups of questions. All of this information must be used by the patient for results to be consistent and valid.

A series of questions may refer to the same time period—such as one week or month—since the last visit. Response grades such as mild, moderate, and severe are often defined in terms of degree of impairment and may be the same for many or all questions. To save space and "clarify" displays, it is tempting to relegate such information to introductory instruction screens, and/or to supplementary "pop-ups." But this can create problems, as those not familiar with computers are less likely to make use of help or other optional facilities and may not remember information from earlier screens.

The concept of cognitive load is an important one, both in understanding the way patients respond to questions (whether in paper or electronic format) and in assessing the importance of differences in format when the content is identical. The load has two main components: the amount of material that must be held in short-term memory, and the amount of processing that has to be done to answer the question.

For example, let's say a patient is asked "How many times did you go to the bathroom today?" In this case, what is meant by going to the bathroom is straightforward, but some processing is necessary to come up with the frequency. If asked about heartburn pain right now, the time element is straightforward, but heartburn may need to be distinguished from other forms of pain or discomfort such as epigastric pain or regurgitation. There will be an even greater load if both the symptom and the time period require processing, as in: "How much have you been bothered by heartburn during the past week?"

Cognitive load



Figure 3: Quality-of-Life Questions Add Complexity



Figure 2: Typical Patient eDiary Question

Mode of administration may influence cognitive load in two main ways: The load may be increased if information is not immediately available. This can occur if the patient has to remember a symptom definition or response period from a previous screen rather than the immediate one. Cognitive load will also be increased if an additional action is required to obtain information, such as tapping on a help tab or navigating back to a previous screen. Such actions constitute concurrent tasks, which may interfere with the primary task of constructing a response to the question. Even very simple concurrent tasks, including holding information in memory, can show significant interference.⁸

However, adding a varying memory load to a task will have much greater impact than adding a constant memory load.⁹

Suppose we start with a questionnaire that requires recalling a series of symptoms over the past seven days. This recall period is specified on each question on the paper version, but to save space on a handheld, we might consider putting this information in an instruction screen at the beginning rather than on every screen. Thus, the patient must hold this information in memory—an example of a constant memory load.

Now consider another approach to space-saving: screen splitting. In this case the question is presented on one screen, and responses on another screen. Again, the patient is required to remember information from one screen to the next, but now the information changes with each question.

Both types of change will increase cognitive load, but the increase is substantially greater with screen-splitting. The load imposed by additional tasks such as scrolling or navigating through an application may also be considered in this way.

When concurrent tasks are automatic they have much less impact than tasks that require effort. Tasks become automatic with repetition and practice, but it can take longer to acquire this automaticity than to become reasonably proficient. Consider how an experienced driver will have no problem talking while changing gear. A new driver may be able to change gear without problem too, but not at the same time as having a conversation. Thus, even when a patient has learned to perform these tasks, there may well be an increase in cognitive load not found for experienced computer users.

Practical issues

These issues have parallel implications for the two domains we started with: user acceptance and the validity and equivalence of ePRO systems. From the perspective of user acceptance, Mikael Palmblad and I¹⁰ have recommended a number of measures to ensure that all types of patients can use ePRO systems effectively. One of these is to have all the information the patient needs immediately available on the application screen, thus avoiding scrolling or screen-splitting.

From the perspective of validity/equivalence, Shields et al.¹¹ propose a hierarchical approach, with increasing requirements for validation as greater changes are made from the paper questionnaire. A trivial change, such as instructing a patient to tap a choice on a screen rather than circle it on paper, would have minimal implications; changes that required increased memory load or increased navigation choices might require the use of methods such as cognitive interviewing or psychometric revalidation. Thus, design choices that improve usability will likely reduce the validation requirement, and vice versa.

These considerations bring us back to device selection as well as to user-interface design. To meet the goal of setting up ePRO applications that all patients can use effectively and be comfortable with, it is important to start with an adequate screen size. What is adequate will depend very much on the questions presented—but if fitting everything on the screen is only possible at the expense of user ease (even for some patients), then the screen is probably too small.

Some patients are likely to find ePRO more challenging than others. Obvious examples include people with poor eyesight or tremor and those anxious about using new technology. For all patients to be able to use ePRO systems, the systems need to be set up with all patients in mind. If we design for a 75-year-old who has never seen a handheld PC, then a 17-year-old who has been using computers for years is not likely to have a problem. The reverse is not true. This principle of inclusive design should be applied to all aspects of the ePRO system.

Conclusion

The observation that patients in general like ePRO (and often prefer it to paper) is not something that is an automatic property of electronic methods. It happens because a lot of care has gone into system design. Some basic principles are implicit in everything that has been said up until now¹⁰ :

- Think the system through from the patient's point of view
- Design for patients who are most likely to find the system challenging
- Do not make it harder than it needs to be for anyone.

These principles apply to selecting a platform or planning the way ePRO systems will be used in a trial as much as they apply to the actual system design. If consistently applied, they will ensure that all types of patients can use

ePRO effectively, be comfortable with it, and provide good quality data.

Brian Tiplady is a senior clinical scientist with *invivodata Ltd.*, Regal House, 70 London Road, Twickenham TW1 3QS, United Kingdom, email: btiplady@invivodata.com

References

1. P.R. Yarnold, M.J. Stewart, F.C. Stille, G.J. Martin, "Assessing Functional Status of Elderly Adults via Microcomputer," *Perceptual & Motor Skills*, 82 (2) 689–690 (1996).
2. A. Begg, G. Drummond, B. Tiplady, "Assessment of Postsurgical Recovery after Discharge Using a Pen Computer Diary," *Anaesthesia*, 58 (11) 1101–1105 (2003).
3. C.J. Gwaltney, A.L. Shields, S. Shiffman, "Equivalence of Electronic and Paper-and-Pencil Administration of Patient Reported Outcome Measures: A Meta-Analytic Review," *Value in Health* (Submitted for Publication).
4. R.N. Jamison, R. Gracely, S. Raymond, J. Levine et al., "Comparative Study of Electronic vs. Paper VAS Ratings: A Randomized, Crossover Trial Using Healthy Volunteers," *Pain*, 99 (1-2) 341 (2002).
5. D. Kreindler, A. Levitt, N. Woolridge, C.J. Lumsden, "Portable Mood Mapping: The Validity and Reliability of Analog Scale Displays for Mood Assessment via Handheld Computer," *Psychiatry Research*, 120 (2) 165–177 (2003).
6. D.A.Revicki, M. Wood, I. Wiklund, J. Crawley, "Reliability and Validity of the Gastrointestinal Symptom Rating Scale in Patients with Gastroesophageal Reflux Disease," *Quality of Life Research*, 7, 75-83 (1998).
7. J. McBeth, "R u Redy 4 Uni's Txt Msg Classes?" *The Scotsman*, 17 August 2005.
8. A. Baddeley, *Working Memory* (Clarendon, Oxford, 1986).
9. M. Venturino, "Interference and Information Organization in Keeping Track of Continually Changing Information," *Human Factors*, 39 (4) 532–539 (1997).
10. M. Palmblad and B. Tiplady, "Electronic Diaries and Questionnaires: Designing User Interfaces that are Easy For All Patients to Use," *Quality of Life Research*, 13, 1199–1207 (2004).
11. A.L. Shields, C.J. Gwaltney, B. Tiplady, J. Paty, S. Shiffman, "Grasping the FDA's PRO Guidance," *Applied Clinical Trials*, 15 (8) 69–83 (2006).