



PRO Recall Periods in Light of the Final FDA Guidance

Annabel Nixon PhD, Diane Wild MSc, Oxford Outcomes

Objectives:

The selection of the most appropriate recall periods for PROs has been a topic of much debate since the release of the draft FDA PRO guidance in February 2006. The final PRO guidance (December 2009) provides more insight into the way that the FDA will evaluate PRO recall periods.

The final PRO guidance (section D3) states that:

- The rationale and appropriateness of the recall period for a PRO instrument should be evaluated
- It is important to consider the patient's ability to validly recall the information requested
- Choice of recall period depends on:
 - measure's purpose and intended use
 - variability, duration, frequency and intensity of the concept measured
 - disease or condition's characteristics
 - treatment being tested
- Need to consider threats to content validity:
 - PRO's that rely on memory, especially if recall is over a long period of time
 - PRO's that compare current state with an earlier period
 - PRO's that ask the patient to average their response over time
- Preferred recall periods are:
 - Short recall periods
 - Current or recent state
- If detailed recall over a period of time is necessary:
 - Ask patients to think of their worst or best experience over the recall period
 - Use a diary method for data collection

This study reviews the literature around PRO recall periods in the light of the final guidance and provides recommendations to sponsors wishing to obtain FDA label claims on the basis of PRO endpoints.

Methods:

A literature review was conducted in Embase and Medline, with further searching in Google scholar. References from each of the relevant papers were hand searched. Forty four papers were reviewed with reference to section D3 of the FDA final PRO guidance, the research was summarized and a set of recommendations were developed.

Results:

Problems associated with recall

Psychology literature identifies that recall of complex information is problematic:

- Memory is limited and selective in terms of what information is encoded and subsequently available for recall
- Systematic biases have been identified in the way that experiences are recollected, more salient and intense events, and the most recent events are more likely to be recalled
- Respondents may experience interference (as an individual experiences an increasing number of events, the probability of recalling any one of those events specifically declines)
- If information required by a question is not available, the respondent may use other less relevant information to answer the question. For example, current pain intensity may be used if the respondent has difficulty remembering pain intensity levels over a longer period of time (DeWalt et al, 2007)
- Other factors that may be related to errors in recall are personal variables (such as age, gender), social desirability, interviewing technique and the motivation of the respondent (Coughlin, 1990)

Recall of Pain and Fatigue

The majority of empirical work with PROs focuses on the measurement of pain with some evidence from fatigue measurement.

- Children strongly overestimated their headache intensity and duration when answering a retrospective headache questionnaire compared to a diary completed four times a day (van den Brink et al, 2000). Another study found low agreement between momentary and recalled headache intensity, especially in those participants whose headaches varied widely (Kikuchi et al, 2006)
- Recall of fatigue for the previous week by participants with chronic fatigue syndrome was consistently higher in comparison to real-time momentary ratings of fatigue (Friedberg and Sohl, 2008)
- There can be substantial difference between recalled pain and fatigue, and momentary assessments (Broderick et al, 2008)
- Recalled assessments may be influenced by peak and recency (Broderick et al, 2008)
- Cognitive debriefing of rheumatology patients (Broderick et al, 2006) who had completed a pain visual analogue scale to rate their pain over the last 7 days revealed that 31% did not consider the entire previous week - when completing the scale. The cognitive debriefing also revealed that participants employed a number of different strategies when answering the recall question:
 - Only 40% attempted to calculate an average of their pain experiences over the week
 - Only half took into account pain-free periods when making their rating
 - Some participants paid particular attention to pain associated with events or activities, when pain interfered with the enjoyment of the event

Results (continued)

- Some used their usual pain level as a reference point to evaluate the previous week
- Cognitive debriefing of the daily diary version of the Gastroparesis Cardinal Symptom Index (GCSI) found that the change to a 24 hour recall (from 2 weeks) was well accepted by the patients. It was also felt that the symptom variability, in terms of frequency and severity would be better captured with the questionnaire being completed on a daily basis (Revicki et al, 2009). These suggest that more direction should be given when a participant is asked to consider their pain over a certain time period, for example, ask them to think about pain free periods, or ask them to think about the worst pain experienced in the previous week. Cognitive debriefing also highlights that even when asked to rate their pain over the previous week, many participants do not attempt to do this.

Recall of Health-related quality of life

Less research has focused on recall bias in HRQL studies; however theoretically, shorter recall periods should be more sensitive to changes in health status than longer recall periods.

- A longitudinal study of men who had undergone radical prostatectomy found that patients do not accurately recall their pretreatment HRQL when asked several months or weeks later. Patients recalled their baseline HRQL as higher than the actual baseline assessment (Litwin and McGuigan, 1999)

Recall of adherence

Research on self-reported adherence to medication, in contrast to much of the pain and symptom research, has found that 1 month recall periods may be more accurate than 3- or 7- day recall periods. The research found that the patients were less likely to over-estimate their adherence when the recall period was longer (Lu et al, 2008).

Conclusions:

Empirical research suggests a lack of correlation between actual experienced symptoms and recalled symptoms, with variability in patient attention to the recall period instruction. Recall is significantly influenced by the concept being measured and attributes of the patient at the time of assessment. Research on recall of pain and symptoms suggests a shorter recall period is more appropriate, whereas research on treatment adherence suggests a longer recall period provides more accurate results. The findings from the research are in line with the FDA concerns and their preference for shorter recall periods for symptoms. The final FDA PRO guidance takes a considered approach to PRO recall periods in light of available research. Recommendations are presented on how best to select and justify the most appropriate recall period for a PRO measure in order to support regulatory review of drug approval label claims.

Recommendations:

- When selecting a PRO, or developing a new PRO, attention should be paid to the recall period
 - It is important to consider whether the recall period is appropriate to the concept being measured, the patient group, the treatment being evaluated and the study design
- In order to assess the appropriateness of a recall period
 - Literature reviews, qualitative research and clinical input should be used
 - Cognitive debriefing can be conducted to identify any specific issues with the recall period
- Recall bias may be reduced by having participants complete daily diaries rather than recall over a longer time period
- Recall bias may also be reduced by being more specific in the questions, for example asking participants to consider their 'worst pain' over a certain time period, or asking them to consider 'pain free periods' when rating their pain
- Recall period assessment, selection, and justification requires full documentation
- Recall periods of existing PRO's should be changed if necessary and the measure should be revalidated.

References:

- Broderick JE, Schwartz JE, Vikingstad G, Probbornow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain*. 2008; 139:146-157
- Broderick JE, Stone AA, Calvanese P, Schwartz JE, Turk DC. Recalled pain ratings: A complex and poorly defined task. *The Journal of Pain*. 2006; 7(2): 142-149
- Coughlin SS. Recall bias in epidemiological studies. *Journal of Clinical Epidemiology*. 1990;43:87-91
- DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates. The PROMIS qualitative item review. *Medical Care*, 2007; 45 (5: Suppl1): S12-S21
- Friedberg F, Sohl SJ. Memory fatigue in chronic fatigue syndrome: The relation between weekly recall and momentary ratings. *International Journal of Behavioural Medicine*. 2008; 15:29-33
- Kikuchi H, Yoshiuchi K, Miyasaka N, Ohashi K, Yamamoto Y, Kumano H, Kuboki T, Akabayashi A. Reliability of recalled self-report on headache intensity: investigation using ecological assessment technique. *Cephalalgia*, 2006; 26:1335-1343
- Litwin MS, McGuigan A. Accuracy of recall in health-related quality-of-life assessment among men treated or prostate cancer. *Journal of Clinical Oncology*, 1999; 17(9): 2822-2888
- Lu M, Safren SA, Skolnik PR, Rogers WH, Coady W, Hardy H, Wilson IB. Optimal recall period and response task for self-reported HIV medication adherence. *AIDS Behaviour*, 2008; 12:86-94
- Revicki DA, Camilleri M, Kuo B, Norton NJ, Murray L, Palsgrove A, Parkman HP. Development and content validity of a gastroparesis cardinal symptom index daily diary. *Alimentary Pharmacology & Therapeutics*, 2009; 30:670-680
- Van den Brink M, Bandell-Hoekstra ENG, Abu-Saad H. The occurrence of recall bias in pediatric headache: A comparison of questionnaire and diary data. *Headache*, 2001; 41:11-20