Compliance with momentary pain measurement using electronic diaries: A systematic review

MATTIJN MORREN, SANDRA VAN DULMEN, JESSIKA OUWERKERK AND JOZIEN BENSING

NIVEL, Netherlands Institute for Health Services Research, P.O. Box 1568, 3500 BN Utrecht, The Netherlands

ABSTRACT

Electronic diaries are increasingly used to assess daily pain in many different forms and populations. This systematic review aims to survey the characteristics of studies using electronic pain diaries and to examine how these characteristics affect compliance. A literature search of 11 electronic databases was conducted. Studies were evaluated on the basis of predetermined inclusion criteria by two independent reviewers. Study characteristics were grouped into four categories: general, population, electronic diary, and sampling procedure (i.e., response, attrition, and compliance rates) including strategies to enhance compliance. The 62 included publications reported from 43 different datasets. Papers were usually written in English and published as from 2000. Samples mostly consisted of female chronic pain patients aged 19–65 years from western countries. Most diaries held less than 20 items and were completed up to 6 times daily at fixed or prompted times for 1 month at most. Less than 25% of the studies reported both response and attrition rates; however, a majority reported compliance. Compliance was generally high, and positively associated with shorter diaries, age, having a user's manual, financial compensation and using an alarm. It is important that the various study characteristics are catalogued carefully, especially response and attrition rates, because they can affect compliance. Measures of momentary pain are often developed for the purpose of a certain study; standardisation and validation of these measures is recommended. Finally, authors should mention whether they report on data that has also been used in previous studies.

INTRODUCTION

Electronic diaries (e-diaries) are increasingly used to assess daily momentary pain. These e-diary studies vary greatly with respect to the population of interest, the e-diary device applied, the number of daily diaries administered for how long, diary length, and measures to assess momentary pain. In addition, researchers have employed various strategies to encourage participants to comply with their study protocol. However, the characteristics of studies investigating momentary pain using electronic devices, or how these characteristics

affect compliance with filling in the diaries, have not been examined systematically. This review aims to investigate these issues.

  The presence of pain is not constant, but tends to fluctuate depending on environmental and psychological influences. Traditional pain measures do not capture these variations as they require patients to recall their pain over a period of days or weeks (Stone et al., 2003a). Since recall seems to involve an active reconstruction of past experiences ([Eich et al., 1985] and [Stone et al., 2003b]), retrospective assessments may introduce a bias, generally an overestimation of the pain (Sorbi et al., 2006a). To prevent this, new methodologies have been developed which can be used to collect information about patients' every-day pain. Patients are asked to provide multiple reports of their pain over a short period of time (e.g., an hour or a day) or at a particular moment. The latter are called momentary pain reports (Sorbi et al., 2006a), and the method used to collect such data is known as ecological momentary assessment (EMA; [Stone and Shiffman, 1994] and [Shiffman and Stone, 1998]) or the experience sampling method (ESM; Csikszentmihalyi and Larson, 1987).

  In pain research, multiple pain reports are usually measured by means of the pain diary: a short questionnaire that is completed once or several times daily (Stone et al., 2003a). It assesses respondents' momentary health or their health during a short period preceding the assessment (Aaron et al., 2004). This enables examining pain variability over time. Pain diaries involve repeated assessment of immediate/momentary pain, frequently supplemented with questions about pain-related variables such as pain-related fear, mood, and coping (e.g., [Affleck et al., 1992], [Litt et al., 2004] and [Roelofs et al., 2004]). These diaries are increasingly administered on small, portable computers instead of using pencil-and-paper.

  An important advantage of this is that high compliance with e-diary assessment is reported ([Hufford and Shields, 2002], [Stone et al., 2003b] and [Salaffi et al., 2005]), which is crucial for the validity of diary research (Stone et al., 2003a). Modern electronic devices can be programmed with alarms and reminders which can foster compliance, and enable signal-contingent diary assessment schemes. Other advantages of e-diaries reported in the literature include date- and time-stamping of diary entries, no forward- and backward-filling, higher data collection efficiency and accuracy, instant access to data, randomisation of item presentation, item branching on the basis of response, and high user-friendliness and -appreciation ([Conner Christensen et al., 2003], [Stone et al., 2003a], [Aaron et al., 2004], [Gaertner et al., 2004], [Jamison et al., 2006] and [Sorbi et al., 2006a]). Finally, e-diary assessments do not seem to influence pain perception and the available evidence indicates that levels of respondent reactivity are negligible ([Hufford and Shields, 2002], [Stone et al., 2003a] and [Aaron et al., 2005]).

  E-diaries have several important disadvantages as well. Most diary studies require patients to complete multiple daily diaries for several days or weeks. This can be inconvenient at times and may sometimes be a burden, especially with frequent diary assessments and extended assessment periods. Patients may miss questions and rush through or skip diaries, which negatively affects the data quality. Furthermore, the electronic diary device must be carried constantly, which can be uncomfortable due to its weight and size. Fortunately, technological advancements foster miniaturisation, allowing integration of diary technology in small, light-weight mobile phones. Apart from the rare event that devices are lost or stolen (although there are exceptions), technical malfunction may occur due to software/hardware problems or accidental mishandling. This calls for the availability of technical support. Finally, it may be difficult for some, especially older people to use the e-diary device, such that training is required; additional problems in handling the device may include poor eye sight, impaired hand–eye coordination, or computer anxiety ([Peters et al., 2000], [Raymond and Ross, 2000], [Conner Christensen et al., 2003] and [Tennen et al., 2006a]).

  Since the variability among studies using e-diaries to assess momentary pain is high, this review is primarily descriptive. The main purpose was to examine characteristics of studies using e-diaries pertaining to the study in general, its population, the electronic diaries, and the sampling procedure. In addition, we investigated whether these study characteristics,

strategies to enhance compliance, and study quality would be associated with compliance rates.

2. METHODS

2.1. Inclusion and exclusion criteria

Studies assessing momentary pain using portable electronic devices were included in this review. There were no age or disease restrictions: participants could be either adults or children, as long as daily pain reports were provided, and might comprise chronic pain patients, healthy individuals with pain complaints, or individuals experiencing an episode of acute pain (e.g., following surgery or during a headache attack).

Studies were excluded if (1) data collection was achieved via paper diaries only, (2) non-portable computer devices were used, (3) momentary assessments were not ecological (e.g., clinic-based), (4) only a single momentary pain assessment was obtained, or (5) no empirical data were reported, for instance when implementation of an electronic pain measurement system was described.

2.2. Search strategy

We searched for studies published in peer-reviewed journals before and including 2006 in any language meeting the above-mentioned inclusion criteria in the following electronic databases: PubMed, Medline, PsycINFO, EMBASE, CINAHL, SERFILE, ERIC, International Bibliography of the Social Sciences (IBSS), International Pharmaceutical Abstracts (IPA), Global Health, and the Cochrane Library. Appendix A contains details of the electronic search.

Two reviewers (MM and SvD) independently evaluated the citations on the basis of title, abstract, and key words. Inclusion of a study commenced when both reviewers agreed that a study was eligible, and exclusion when both agreed that it was not. When the reviewers disagreed or had doubts, the full-text paper was evaluated. Remaining disagreements were resolved via consensus among the two reviewers. Because consensus could be reached on all papers, consultation of a third reviewer was not required. In addition, the references of included studies were searched for additional pain e-diary studies. The abstracts and/or full-text of candidate studies were subsequently evaluated by both reviewers as described above.

2.3. Multiple publications from the same dataset

When performing a systematic review, it is important to identify multiple publications from the same dataset. When such reports, and with that the data, are counted as separate studies this introduces a bias because this dataset more strongly affects the results of the review (Higgins and Green, 2008). Thus, studies included in this review were examined to establish groups of studies reporting from the same dataset. If this was the case, the information extracted from studies within a group was collapsed. Subsequently, it was examined whether these studies provided citations to previous studies, notably the parent study. More specifically, a distinction was made between studies that provided (1) no citation; (2) an implicit citation, indicating that a citation to the parent study is provided, but it is not mentioned that (part of) the same dataset was used; or (3) an explicit citation: a citation to the parent study is provided, which indicates that (part of) the same dataset was used.

2.4. Extraction of study characteristics

Information was extracted from the included studies by reviewer 1 (MM) and tabulated (see Table 1A and Table 1B). These study characteristics were then grouped into four categories: general, population, e-diary, and sampling characteristics.

[TABLE 1A].

[TABLE 1B].

2.4.1. General characteristics

General characteristics were publication year and language. To describe study design, a distinction was made between primary and secondary investigations. Primary refers to studies using diary assessments to examine momentary pain, and might involve the monitoring of momentary pain in real-life or during treatment, or for testing hypotheses concerning the aetiology and maintenance of pain. Secondary refers to studies using diary assessments for other purposes, and might involve a psychometric evaluation of e-diary methodology or a trial using pain assessed by diary as outcome variable to establish treatment efficacy.

2.4.2. Population characteristics

Population characteristics included sample size, gender, age (mean and standard deviation or median, and range), pain type, and country of origin.

2.4.3. Electronic diary characteristics

E-diary characteristics were frequency (diaries/day) and duration of diary assessment (in days), diary assessment scheme, diary length (number of items, or the mean/diary/day when diaries of different lengths were administered each day), and which pain dimensions and related concepts were measured with the diary.

2.5. Sampling characteristics

Sampling characteristics concerned response, attrition and compliance rates, and strategies to maximise compliance. Response rate was defined as the number of individuals who were willing and able to participate and conformed to the study's inclusion criteria, divided by the total number of individuals identified as potential participants. Attrition was defined as the percentage of respondents who were not included in the final analyses, for instance because they withdrew, made insufficient diary entries, or when technical malfunction occurred and replacement of the diary device was not possible (e.g., [Aaron et al., 2006] and [Kikuchi et al., 2006]). Compliance was reported in different ways, and was defined as the percentage of diary entries, or the percentage of completed items. Compliance rates were either directly taken from the papers or calculated on the basis of the information provided in the papers concerning diary length, and frequency and duration of the diary assessment. Several strategies were employed to maximise compliance with diary completion, such as instruction and training or sounding alarm to prompt diary completion. Both the type and number of strategies employed were examined.

2.6. Quality assessment

The methodological quality of all studies was assessed using a list of 10 criteria, which was formulated for the purpose of this study (Appendix B). Existing quality assessment instruments were unsuitable, because most of them are designed for evaluating clinical trials (Verhagen et al., 2001; cf. Mallen et al., 2006) or for specific study designs (Katrak et al., 2004); the designs of the studies selected for this review were highly variable, ranging from randomised clinical trial to psychometric investigation. Therefore, the focus of our assessment was on the completeness and accurateness of the information provided in each section of the studies (cf. Mallen et al., 2006).

Reviewer 1 (MM) rated each study as either poor (=0), reasonable (=1) or good (=2) on each criterion. Item scores were summed to obtain a total study quality score (range 0–20). Reviewer 2 (SvD) evaluated a sample of thirteen studies to assess agreement with the quality ratings of reviewer 1. For studies from the same dataset that provided an explicit citation,

item ratings of previous reports were imputed if these ratings were higher. For the purpose of the analysis, the quality sum score was dichotomised into studies with above or below average quality.

2.7. Data synthesis and analysis

The extracted information (Table 1A and Table 1B) was used for a descriptive analysis of study, population, e-diary, and sampling characteristics. When the same dataset was used for more than one paper, missing information was imputed. In addition, descriptive information was reported at the highest possible level of aggregation for each variable. In effect, this meant that population, non-response and attrition characteristics were drawn from the study reporting on the largest or most final sample (if this information was missing, information from the second largest sample was used, etc.).

To examine agreement concerning the quality ratings of the two reviewers on a sample of the identified studies, the intra-class correlation coefficient was calculated using the two-way random effects model and absolute agreement definition (ICC). We report the single measure ICC, which must be used when there are two judges. Mean quality was computed excluding one of the two studies by Badr et al., because these were evaluated together.

To examine which study characteristics would predict compliance, and thus might be used to enhance compliance in e-diary research, a linear regression analysis was carried out. Thus, as we did not seek the most parsimonious model of compliance, the backward procedure was used with the critical value for removal of variables set at $p = 0.10$. All study characteristics were entered into the equation with two exceptions. Design was not used because it varied across studies using the same dataset while the same compliance rate was reported. Type of assessment scheme was excluded because this can be captured by the compliance strategy alarm.

3. RESULTS

3.1. Results of the literature search

The search generated 588 unique citations; Fig. 1 shows the number of citations obtained from each database source. The first selection resulted in both reviewers including 39 papers, rejecting 455, disagreeing about 86, and both being uncertain about 8 (kappa = 0.50, $p < 0.001$). Full-text evaluation of the 94 papers for which disagreement or uncertainty existed resulted in the inclusion of an additional 20 papers. As 5 of the 39 originally included papers were excluded during the data extraction, 54 papers were included on the basis of the database search (Fig. 1). Reference tracking yielded 89 new citations and resulted in inclusion of eight additional papers following the same procedure as with the database search. Thus, a total of 62 papers were included in this review, two of which presented two separate studies ([Hermann et al., 1995] and [Badr et al., 2006]).

[FIGURE 1]

3.2. Study characteristics

3.2.1. General characteristics

The publication year of the 62 included papers ranged from 1991 to 2006, three quarters of which were published between 2000 and 2006. Fifty-five papers were written in English (89%), six in German (10%), and one in Italian (2%). The purpose of 32 studies was the primary assessment of momentary pain (seven monitored pain) and 30 were secondary assessments (19 psychometric) (Table 2).

3.2.2. Multiple publications from the same dataset

In 10 cases, it was concluded that more than one publication reported from the same dataset or a growing number of participants (see Table 1A and Table 1B). Nine of the 31 studies involved mentioned explicitly that the paper reported from a previously used dataset, five did so implicitly, and seventeen did not provide a citation. Each group of studies was analysed as a single study, considering population, e-diary, and sampling characteristics, except when different assessment periods were examined with a different sample size (i.e., [Roelofs et al., 2006] and [Turner et al., 2005]). Study quality ratings were not collapsed; however, if an explicit citation was provided information of the cited study was included in the quality rating of the citing study.

3.2.3. Population characteristics

The majority of studies were carried out with populations from the USA ($n = 19$) or Europe ($n = 19$), and only few in Asia ($n = 2$), Oceania ($n = 1$) or in both Europe and North America ($n = 2$). Sample sizes ranged between 10 and 2696 participants, although about two-thirds were between 21 and 100. The samples of almost 70% of studies comprised a majority of females. All but two studies examined adult samples, and almost half sampled individuals with a mean age between 45 and 65 years. Thirty-seven studies examined chronic pain patients, three acute pain patients, one both chronic and acute pain patients, one chronic pain patients and healthy controls, one examined wheelchair users, who often experience low back pain, one individuals with multiple chemical sensitivity, and one patients with Parkinson's disease. Diaries were most often administered to patients with headache ($n = 10$), fibromyalgia ($n = 8$), arthritis ($n = 8$), low back pain ($n = 4$), and cancer pain ($n = 4$) (also see Table 2).

3.2.4. Electronic diary characteristics

Diary assessments covered periods from 4 h to 1 year, but the majority monitored pain 1 week to 1 month (63%). Participants completed between 1 and 10 diaries each day but mostly no more than six (84%). Most diaries comprised 20 items or less (79%; range 1–63). To describe when participants were required to complete their diaries, four assessment schemes were distinguished (based on Conner Christensen et al., 2003). Most studies used time-contingent (i.e., diaries are completed at fixed pre-determined times or time windows) or signal-contingent (i.e., participants are prompted to complete the diary by an alarm, which is set fully random or randomly within a certain time-window) assessment schemes (53% and 42%, respectively). Event-contingent (i.e., diary entry is linked to a certain event, e.g., waking-up or a headache episode) and person-contingent assessments (i.e., diary entries are initiated by the participant) were used less frequently (21% and 16%, respectively). Combinations of two assessment schemes were employed in 14 studies (33%). All studies assessed pain intensity, and pain location or pain quality was examined in about one quarter of the studies. Other concepts that were frequently assessed with the diary were mood (40%), sleep quality and fatigue (37%), limitations of activities or disability (23%), and health care utilisation (14%) (Table 2). Most studies used modern palmtop computers or personal digital assistants (PDAs) to administer the diaries. A few, mostly early diary studies, however, used specifically designed applications (e.g., the Symtrack Logger or Indamos) (Table 1B).

3.2.5. Sampling characteristics

The response rate was reported in less than a third of the studies (mean rate = 53%), almost three of four studies reported the attrition rate (mean rate = 19%), and only nine studies (21%) reported both attrition and response rates. About 75% of the studies reported the compliance rate; the mean rate across studies was 83%. Various strategies to maximise compliance were employed. The most common strategies were use of an alarm on the diary device to notify participants to complete the diary (67%), and training/instruction on how to use the diary (65%). Reminders to notify participants when they missed a diary and the

possibility to postpone the diary for a short time were least utilised (both 16%). The remaining strategies were used in between 18% and 30% of the studies (i.e., financial compensation for participating, providing a diary manual, one or more phone calls by the researcher to discuss difficulties with the diaries, a help-desk which could be phoned or paged in case of diary trouble, and other strategies). On average, almost three compliance strategies were employed (Table 2).

3.3. Quality of the studies

The agreement between both reviewers on the quality of the sample of 13 papers was good (ICC = 0.79, $p < 0.001$), confirming the adequacy of reviewer 1's evaluation. Quality ratings for individual studies are reported in Table 1B. As can be seen in Table 2, mean quality was 11.7 (SD = 3.8); the quality rating of 27 studies was lower than the mean (44%) and that of 35 was higher (56%). Study quality was not associated with publication year, language of the publication, and country of origin (χ^2s < 1.0, ns).

3.4. Predictors of compliance

To identify predictors of compliance, a backward regression analysis was carried out. The resulting model contained six variables and was highly significant ($F(6, 25)$ = 6.04, $p < 0.001$). Shorter diaries, older patients, providing participants with a manual, using alarms to signal when the diary must be completed, financial compensation, and other compliance strategies were associated with higher compliance rates (Table 3). In addition, although a lower percentage of female participants and using less different diary assessment schemes, did not emerge as significant predictors in the regression analysis, the partial correlation of both variables with compliance was higher than 0.2 (ps < 0.21).

[TABLE 3.]

4. DISCUSSION

The present systematic review examined the use of electronic diaries in the assessment of momentary pain. There were two aims. First, the characteristics of the included studies and their populations, e-diary methodology, and sampling procedures were surveyed. Second, it was examined whether these characteristics affected the compliance of participants with completing the diaries.

An important finding of this review is that compliance with the diary protocols was high (83%); only a third of the studies reported a compliance rate below 80%. Higher compliance rates were reported with shorter diaries and older samples. Several strategies were identified that contributed to compliance. Providing participants with a manual, sounding an alarm when a diary must be filled out, financial compensation, and other compliance strategies (e.g., visits by or messages from the study coordinator, compliance feedback, suspend prompting during uninterruptible activities, etc.) were found to be effective. Thus, although the diary assessment protocol is an important source of non-compliance, it is not the only one. Participants themselves have reported that the most common reasons for missing diaries reported by participants are not hearing the alarm, inconvenient time, or being too busy (Aaron et al., 2004). Inquiring reasons for non-compliance with participants may help improve the user-friendliness of the assessment protocol.

Most studies calculated compliance on the basis of those individuals who successfully completed the diary assessment period, after excluding people who refused or were unable to participate. However, several studies made this calculation after excluding individuals who withdrew after starting the diary assessments. It might be argued that these individuals, if included in the diary assessments, would be prone to non-compliance. Thus, their exclusion may have resulted in an overestimation of the compliance rate. It would be interesting to know why these individuals withdrew, whether their participation could have been retained if certain compliance enhancement strategies had been implemented, and to what degree this

affected the study results. Other potential confounders of compliance include (1) self-selection of individuals who are enthusiastic about the diary methodology, (2) user-friendliness of the diary device (e.g., weight, size, font size, item relevance), (3) exclusion of non-compliant individuals from the compliance calculation, (4) the criteria used for considering a diary valid (e.g., whether a certain time boundary is required or how many items must be filled out to consider the diary valid), and (5) which part of the planned assessment period is reported on (also see Tennen et al., 2006a). Thus, to avoid confounding and to ensure that results can be generalised, it is important to report (reasons for) non-response and attrition, and how compliance is calculated. This review found that sampling information was poorly reported in the included studies: less than 25% stated both the response and attrition rate.

Our search strategy identified 62 publications, 31 of which were judged to report from datasets that had been used in previous publications. Even though this amounts to half of the studies included in this review, it is fairly acceptable because e-diary studies typically collect large quantities of data. This permits investigation of distinct research questions, which can be reported in different publications to improve the transparency of the issue under investigation. Nevertheless, when a publication is one of multiple reports from a single dataset, this fact should be reported explicitly and citations to associated studies should be provided. In this review, this was found to be the case for less than a third of these studies.

Another issue is the validity and reliability of the pain measurements. There is a large variety in the items used in e-diaries to measure pain. Each research group seems to develop their own set of questions or extracts items from existing questionnaires. Consequently, there is no information as to whether these measures are valid and reliable. An additional complication is that e-diary questionnaires must be brief, which disqualifies application of traditional psychometric analysis. This may be due to the relatively recent introduction of e-diaries and the considerable increase in their application since then. Therefore, new guidelines for item formulation should be developed (e.g., Van den Heuvel et al., 2006) and more standardised measures should be developed to facilitate comparison of study results (cf. Burton et al., 2007).

This study has several limitations. First, even though we tried to maximise the sensitivity of our search strategy, it is possible that studies were missed. Second, it was difficult to compare the included studies, because of their heterogeneity. In addition, we extracted population, e-diary, and sampling information that was reported in the studies, but cannot be certain that the report is exhaustive. In scientific reports, the information that is reported should be used to answer the research questions, which means that in reports of diary research, for instance, only diary items used for the analyses of the study will be reported. Alternatively, especially in early e-diary studies, certain information that is considered important now may not have been deemed relevant when those studies were published. As it was not enquired with authors whether this information was available, it may have influenced the information extracted as well as our ratings of study quality. Furthermore, as no validated quality criteria were available for rating e-diary studies, a list was developed for the purpose of this review. Even though the inter-rater reliability was good, this may have introduced some subjectivity, so the quality assessment's results should be interpreted with caution.

In conclusion, high compliance is crucial for the ability of e-diary studies to produce useful results. Several efforts can be made to improve compliance with completing daily diaries. Consistent with the literature, this review suggests that non-compliance can be discouraged by limiting the length of diaries to 20–30 items ([Van den Heuvel et al., 2006] and [Burton et al., 2007]). In addition, it was found that several strategies may help to enhance compliance, especially providing a user's manual, offering remuneration, and using alarms to request diary completion. Although the average compliance rate of the included studies was reported to be high, the effectiveness of compliance strategies may depend on characteristics of the sample under investigation. Also, while effective strategies may enhance compliance,

removal of ineffective strategies may increase cost effectiveness and patient burden. Furthermore, compliance rates are subject to confounding, for instance when non-compliant individuals are removed from the analyses. To estimate whether confounding has occurred and to understand how this may have affected compliance, it is important that study methods and materials are reported in detail. This requires that rates and reasons for non-response and attrition, compliance rates, and strategies used to improve compliance are conveyed. Finally, the transparency of the criteria which define a valid diary may be improved if a distinction is made between diary and protocol compliance, i.e., the percentage of items or diaries completed versus the percentage of diary entries made on time (as defined in the assessment protocol).

Potential conflict of interest
None.

REFERENCES

Aaron et al., 2004 L.A. Aaron, L. Mancl, J.A. Turner, C.N. Sawchuk and K.M. Klein, Reasons for missing interviews in the daily electronic assessment of pain, mood, and stress, Pain 109 (2004), pp. 389–398.

Aaron et al., 2005 L.A. Aaron, J.A. Turner, L. Mancl, H. Brister and C.N. Sawchuk, Electronic diary assessment of pain-related variables: is reactivity a problem?, J Pain 6 (2005), pp. 107–115.

Aaron et al., 2006 L.A. Aaron, J.A. Turner, L.A. Mancl, C.N. Sawchuk, K.H. Huggins and E.L. Truelove, Daily pain coping among patients with chronic temporomandibular disorder pain: an electronic diary study, J Orofac Pain 20 (2006), pp. 125–137. Affleck et al., 1992 G. Affleck, S. Urrows, H. Tennen and P. Higgins, Daily coping with pain from rheumatoid arthritis: patterns and correlates, Pain 51 (1992), pp. 221–229.

Affleck et al., 1996 G. Affleck, S. Urrows, H. Tennen, P. Higgins and M. Abeles, Sequential daily relations of sleep, pain intensity, and attention to pain among women with fibromyalgia, Pain 68 (1996), pp. 363–368.

Affleck et al., 2000 G. Affleck, H. Tennen, S. Urrows, P. Higgins and M. Abeles, Downward comparisons in daily life with chronic pain: dynamic relations with pain intensity and mood, J Soc Clin Psychol 19 (2000), pp. 499–518.

Affleck et al., 1998 G. Affleck, H. Tennen, S. Urrows, P. Higgins, M. Abeles and C. Hall et al., Fibromyalgia and women's pursuit of personal goals: a daily process analysis, Health Psychol 17 (1998), pp. 40–47.

Affleck et al., 2001 G. Affleck, H. Tennen, A. Zautra, S. Urrows, M. Abeles and P. Karoly, Women's pursuit of personal goals in daily life with fibromyalgia: a value-expectancy analysis, J Consult Clin Psychol 69 (2001), pp. 587–596.)

Badr et al., 2006 H. Badr, K. Basen-Engquist, C.L.C. Taylor and C. de Moor, Mood States Associated with Transitory Physical Symptoms Among Breast and Ovarian Cancer Survivors, J Behav Med 29 (2006), pp. 461–475.

Begg et al., 2003 A. Begg, G. Drummond and B. Tiplady, Assessment of postsurgical recovery after discharge using a pen computer diary, Anaesthesia 58 (2003), pp. 1101–1105. F

Bolten et al., 1991 W. Bolten, M. Emmerich, E. Weber and N. Fassmeyer, [Validation of electronic by conventional pain diaries], Z Rheumatol 50 (Suppl. 1) (1991), pp. 55–64.

Brackertz et al., 1991 D. Brackertz, E. Leonhart, H.U. Gerbershagen, C. Panhans, M. Schröder and F.J. Wagenhäuser et al., Real time collection of pain profile in treatment with ibuprofen, Z Rheumatol 50 (Suppl. 1) (1991), pp. 39–47.)

Burton et al., 2007 C. Burton, D. Weller and M. Sharpe, Are electronic diaries useful for symptoms research? A systematic review, J Psychosom Res 62 (2007), pp. 553–561.

Conner Christensen et al., 2003 T. Conner Christensen, L. Feldman Barrett, E. Bliss-Moreau, K. Lebo and C. Kaschub, A practical guide to experience-sampling procedures, J Happiness Stud 4 (2003), pp. 53–78.

Csikszentmihalyi and Larson, 1987 M. Csikszentmihalyi and R. Larson, Validity and reliability of the Experience Sampling Method, J Nerv Mental Dis 175 (1987), pp. 526–536.

Eich et al., 1985 E. Eich, J. Reeves, B. Jaeger and S. Graff-Radford, Memory for pain: Relation between past and present pain intensity, Pain 23 (1985), pp. 375–379.

Feiler et al., 2005 S. Feiler, K.G. Müller, A. Müller, R. Dahlhaus and W. Eich, Using interaction graphs for analysing the therapy process, Psychother Psychosom 74 (2005), pp. 93–99.

Gaertner et al., 2004 J. Gaertner, F. Elsner, K. Pollmann-Dahmen, L. Radbruch and R. Sabatowski, Electronic pain diary: a randomized crossover study, J Pain Symptom Manage 28 (2004), pp. 259–267.

Gendreau et al., 2003 M. Gendreau, M.R. Hufford and A.A. Stone, Measuring clinical pain in chronic widespread pain: selected methodological issues, Best Pract Res Clin Rheumatol 17 (2003), pp. 575–592.

Gerbershagen et al., 1995 H.U. Gerbershagen, P. Frellinger and C. Baedeker, Lysin-Clonixinat in spondylarthrosis deformans – monitoring of the effectivity with an electronical pain-diary, Anasthesiol Intensivmed Notfallmed Schmerzther 30 (1995), pp. 127–128.

Giffin et al., 2003 N.J. Giffin, L. Ruggiero, R.B. Lipton, S.D. Silberstein, J.F. Tvedskov and J. Olesen et al., Premonitory symptoms in migraine: an electronic diary study, Neurology 60 (2003), pp. 935–940.

Han et al., 2002 S.H. Han, J.M. de Klerk, S. Tan, A.D. het Schip, B.H. Derksen and A. van Dijk et al., The PLACORHEN study: a double-blind, placebo-controlled, randomized radionuclide study with (186)Re-etidronate in hormone-resistant prostate cancer patients with painful bone metastases. Placebo Controlled Rhenium Study, J Nucl Med 43 (2002), pp. 1150–1156.

Harris et al., 2005 R.E. Harris, D.A. Williams, S.A. McLean, A. Sen, M. Hufford and R.M. Gendreau et al., Characterization and consequences of pain variability in individuals with fibromyalgia, Arthritis Rheum 52 (2005), pp. 3670–3674.

Hermann et al., 1995 C. Hermann, M.L. Peters and E.B. Blanchard, Use of hand-held computers for symptom-monitoring: the case of chronic headache, Mind/Body Med 1 (1995), pp. 59–69.

Higgins and Green, 2008 Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions Version 5.0.0 [updated February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Höfer and Battig, 1994 I. Höfer and K. Battig, Cardiovascular, behavioral, and subjective effects of caffeine under field conditions, Pharmacol Biochem Behav 48 (1994), pp. 899–908.

Honkoop et al., 1999 P.C. Honkoop, M.J. Sorbi, G.L.R. Godaert and E.L.H. Spierings, High-density assessment of the HIS classification criteria for migraine without aura: a prospective study, Cephalalgia 19 (1999), pp. 201–206. View Record in Scopus | Cited By in Scopus (16)

Hufford and Shields, 2002 M.R. Hufford and A.L. Shields, Electronic diaries. Applications and what works in the field, Appl Clin Trials 11 (2002), pp. 46–56.

Jamison et al., 2001 R.N. Jamison, S.A. Raymond, J.G. Levine, E.A. Slawsby, S.S. Nedeljkovic and N.P. Katz, Electronic diaries for monitoring chronic pain: 1-year validation study, Pain 91 (2001), pp. 277–285.

Jamison et al., 2006 R.N. Jamison, S.A. Raymond, E.A. Slawsby, G.J. McHugho and J.C. Baird, Pain assessment in patients with low back pain: comparison of weekly recall and momentary electronic data, J Pain 7 (2006), pp. 192–199.

Katrak et al., 2004 P. Katrak, A.E. Bialocerkowski, N. Massy-Westropp and K.A. Kumar VSS Grimmer, A systematic review of the content of critical appraisal tools, BMC Med Res Methodol 4 (2004), p. 22.

Kikuchi et al., 2006 H. Kikuchi, K. Yoshiuchi, N. Miyasaka, K. Ohashi, Y. Yamamoto and H. Kumano et al., Reliability of recalled self-report on headache intensity: investigation using ecological momentary assessment technique, Cephalalgia 26 (2006), pp. 1335–1343.

Kop et al., 2005 W.J. Kop, A. Lyden, A.A. Berlin, K. Ambrose, C. Olsen and R.H. Gracely et al., Ambulatory monitoring of physical activity and symptoms in fibromyalgia and chronic fatigue syndrome, Arthritis Rheum 52 (2005), pp. 296–303. Lewis et al., 1995 B. Lewis, D. Lewis and G. Cumming, Frequent measurement of chronic pain: an electronic diary and empirical findings, Pain 60 (1995), pp. 341–347.

Lewis et al., 1994 B. Lewis, D. Lewis and G. Cumming, The comparative analgesic efficacy of transcutaneous electrical nerve stimulation and a non-steroidal anti-inflammatory drug for painful osteoarthritis, Br J Rheumatol 33 (1994), pp. 455–460.

Liszka-Hackzell and Martin, 2004 J.J. Liszka-Hackzell and D.P. Martin, An analysis of the relationship between activity and pain in chronic and acute low back pain, Anesth Analg 99 (2004), pp. 477–481.

Liszka-Hackzell and Martin, 2005 J.J. Liszka-Hackzell and D.P. Martin, Analysis of nighttime activity and daytime pain in patients with chronic back pain using a self-organizing map neural network, J Clin Monit Comput 19 (2005), pp. 411–414.

Litcher-Kelly et al., 2004 L. Litcher-Kelly, A.A. Stone, J.E. Broderick and J.E. Schwartz, Associations among pain intensity, sensory characteristics, affective qualities, and activity limitations in patients with chronic pain: a momentary, within-person perspective, J Pain 5 (2004), pp. 433–439.

Litt et al., 2004 M.D. Litt, D. Shafer and C. Napolitano, Momentary mood and coping processes in TMD pain, Health Psychol 23 (2004), pp. 354–362.

Mallen et al., 2006 C. Mallen, G. Peat and P. Croft, Quality assessment of observational studies is not commonplace in systematic reviews, J Clin Epidemiol 59 (2006), pp. 765–769)

McLean et al., 2005 S.A. McLean, D.A. Williams, R.E. Harris, W.J. Kop, K.H. Groner and K. Ambrose et al., Momentary relationship between cortisol secretion and symptoms in patients with fibromyalgia, Arthritis Rheum 52 (2005), pp. 3660–3669.

Müller et al., 2004 K.G. Müller, C. Bieber, A. Muller, K. Blumenstiel and W. Eich, Psychosocial factors in a fibromyalgia syndrome patient Assessment by means of electron pain diaries. Casuistics and multivariate time series analysis, Schmerz 18 (2004), pp. 45–52.

Nived et al., 1994 O. Nived, G. Sturfelt, S.A. Eckernas and P. Singer, A comparison of 6 months' compliance of patients with rheumatoid arthritis treated with tenoxicam and naproxen. Use of patient computer data to assess response to treatment, J Rheumatol 21 (1994), pp. 1537–1541.

Nyholm et al., 2004 D. Nyholm, J. Kowalski and S.M. Aquilonius, Wireless real-time electronic data capture for self-assessment of motor function and quality of life in Parkinson's disease, Mov Disord 19 (2004), pp. 446–451.

Ott and Scholz, 1999 R. Ott and O.B. Scholz, [The Electronic Bonn Pain Diary: Report of a computer-aided procedure for pain diagnosis], Z Klin Psychol Psychiatr Psychother 47 (1999), pp. 191–206.

Palermo et al., 2004 T.M. Palermo, D. Valenzuela and P.P. Stork, A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability, Pain 107 (2004), pp. 213–219.

Peters et al., 2000 M.L. Peters, M.J. Sorbi, D.A. Kruise, J.J. Kerssens, P.F.M. Verhaak and J.M. Bensing, Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain, Pain 84 (2000), pp. 181–192.)

Raymond and Ross, 2000 S.A. Raymond and R.N. Ross, Electronic subject diaries in clinical trials, Appl Clin Trials 9 (2000), pp. 48–58.

Roelofs et al., 2004 J. Roelofs, M.L. Peters, J. Patijn, E.G. Schouten and J.W. Vlaeyen, Electronic diary assessment of pain-related fear, attention to pain, and pain intensity in chronic low back pain patients, Pain 112 (2004), pp. 335–342.

Roelofs et al., 2006 J. Roelofs, M.L. Peters, J. Patijn, E.G. Schouten and J.W. Vlaeyen, An electronic diary assessment of the effects of distraction and attentional focusing on pain intensity in chronic low back pain patients, Br J Health Psychol 11 (2006), pp. 595–606.)

Saito et al., 2005 M. Saito, H. Kumano, K. Yoshiuchi, N. Kokubo, K. Ohashi and Y. Yamamoto et al., Symptom profile of Multiple Chemical Sensitivity in actual life, Psychosom Med 67 (2005), pp. 318–325.

Salaffi et al., 2005 F. Salaffi, A. Stancati, R. Procaccini, F. Cioni and W. Grassi, Assessment of circadian rhythm in pain and stiffness in rheumatic diseases according the EMA (Ecologic Momentary Assessment) method: patient compliance with an electronic diary, Reumatismo 238 (2005), pp. 238–249.

Serif and Ghinea, 2005 T. Serif and G. Ghinea, Recording of time-varying back-pain data: a wireless solution, IEEE Trans Inf Technol Biomed 9 (2005), pp. 447–458.

Sheftell et al., 2005 F.D. Sheftell, C.G. Dahlof, J.L. Brandes, R. Agosti, M.W. Jones and P.S. Barrett, Two replicate randomized, double-blind, placebo-controlled trials of the time to onset of pain relief in the acute treatment of migraine with a fast-disintegrating/rapid-release formulation of sumatriptan tablets, Clin Ther 27 (2005), pp. 407–417.)

Shiffman and Stone, 1998 S.S. Shiffman and A.A. Stone, Ecological Momentary Assessment: a new tool for behavioral medicine research. In: D. Krantz and A. Baum, Editors, Technology and methods in behavioral medicine, Erlbaum, Mahwah (NJ) (1998), pp. 117–131.

Smith and Safer, 1993 W.B. Smith and M.A. Safer, Effects of present pain level on recall of chronic pain and medication use, Pain 55 (1993), pp. 355–361.

Sorbi et al., 2006a M.J. Sorbi, M.L. Peters, D.A. Kruise, C.J. Maas, J.J. Kerssens and P.F. Verhaak et al., Electronic momentary assessment in chronic pain I: psychological pain responses as predictors of pain intensity, Clin J Pain 22 (2006), pp. 55–66.

Sorbi et al., 2006b M.J. Sorbi, M.L. Peters, D.A. Kruise, C.J.M. Maas, J.J. Kerssens and P.F.M. Verhaak et al., Electronic momentary assessment in chronic pain II: pain and psychological pain responses as predictors of pain disability, Clin J Pain 22 (2006), pp. 67–81.

Sprott and Muller, 1998 H. Sprott and W. Muller, Functional symptoms in fibromyalgia: monitored by an electronic diary, Clin Bull Myofascial Ther 3 (1998), pp. 61–67.

Sprott et al., 1993 H. Sprott, P. Mennet, T. Stratz and W. Müller, Course of pain in fibromyalgia syndrome – recorded by an electronic diary, Aktuelle Rheumatol 18 (1993), pp. 38–43.

Stone and Shiffman, 1994 A.A. Stone and S.S. Shiffman, Ecological Momentary Assessment (EMA) in behavioural medicine, Ann Behav Med 16 (1994), pp. 199–202.

Stone et al., 2004 A.A. Stone, J.E. Broderick, S.S. Shiffman and J.E. Schwartz, Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain, Pain 107 (2004), pp. 61–69.

Stone et al., 2005 A.A. Stone, J.E. Schwartz, J.E. Broderick and S.S. Shiffman, Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic, Pers Soc Psychol Bull 31 (2005), pp. 1340–1346.

Stone et al., 2003b A.A. Stone, S. Shiffman, J.E. Schwartz, J.E. Broderick and M.R. Hufford, Patient compliance with paper and electronic diaries, Control Clin Trials 24 (2003), pp. 182–199.)

Stone et al., 2003a A.A. Stone, J.E. Broderick, J.E. Schwartz, S. Shiffman, L. Litcher-Kelly and P. Calvanese, Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction, Pain 104 (2003), pp. 343–351.

Tennen et al., 2006b H. Tennen, G. Affleck and A. Zautra, Depression history and coping with chronic pain: a daily process analysis, Health Psychol 25 (2006), pp. 370–379.

Turner et al., 2005 J.A. Turner, L. Mancl and L.A. Aaron, Brief cognitive-behavioral therapy for temporomandibular disorder pain: effects on daily electronic outcome and process measures, Pain 117 (2005), pp. 377–387.

Turner et al., 2004 J.A. Turner, L. Mancl and L.A. Aaron, Pain-related catastrophizing: a daily process study, Pain 110 (2004), pp. 103–111.

Tennen et al., 2006a H. Tennen, G. Affleck, J.C. Coyne, R.J. Larsen and A. DeLongis, Paper and plastic in daily diary research: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006), Psychol Methods 11 (2006), pp. 112–118.

Van den Heuvel et al., 2006 E.H.M. Van den Heuvel, M. Sonnenschein and M. Sorbi, Manual for electronic diary research: example in a somatoform population, Department of Health Psychology, University of Utrecht, Utrecht, The Netherlands (2006).

Van Gerven et al., 1996 J.M. Van Gerven, R.C. Schoemaker, L.D. Jacobs, A. Reints, M.J. Ouwersloot-van der Meij and H.G. Hoedemaker et al., Self-medication of a single headache episode with ketoprofen, ibuprofen or placebo, home-monitored with an electronic patient diary, Br J Clin Pharmacol 42 (1996), pp. 475–481.

Verhagen et al., 2001 A. Verhagen, H. de Vit, Bie de, M. Boers and P. van den Brandt, The art of quality assessment of RCTs included in systematic reviews, J Clin Epidemiol 54 (2001), pp. 651–654.

Viane et al., 2004 I. Viane, G. Crombez, C. Eccleston, J. Devulder and W. De Corte, Acceptance of the unpleasant reality of chronic pain: effects upon attention to pain and engagement with daily activities, Pain 112 (2004), pp. 282–288.

Walker and Sorrells, 2002 L.S. Walker and S.C. Sorrells, Brief report: assessment of children's gastrointestinal symptoms for clinical trials, J Pediatr Psychol 27 (2002), pp. 303–307.

Wig et al., 2004 A.D. Wig, L.A. Aaron, J.A. Turner, K.H. Huggins and E. Truelove, Short-term clinical outcomes and patient compliance with temporomandibular disorder treatment recommendations, J Orofac Pain 18 (2004), pp. 203–213.

Williams et al., 2004 D.A. Williams, M. Gendreau, M.R. Hufford, K. Groner, R.H. Gracely and D.J. Clauw, Pain assessment in patients with fibromyalgia syndrome: a consideration of methods for clinical trials, Clin J Pain 20 (2004), pp. 348–356.

Zautra et al., 2001 A. Zautra, B. Smith, G. Affleck and H. Tennen, Examinations of chronic pain and affect relationships: applications of a dynamic model of affect, J Consult Clin Psychol 69 (2001), pp. 786–795.
Appendix I. Electronic search

The search term "pain" was combined with a collection of descriptors of momentary assessments using portable electronic devices. The strategy used to find relevant studies in PubMed was adapted accordingly for searching the other databases.

Pain AND ("computer diary" OR "computer diaries" OR momentary OR "mobile pc" OR "mobile computer" OR "computerised data collection" OR "computerized data collection" OR "handheld electronic diary" OR "handheld electronic diaries" OR "pocket PC" OR "pocket computer" OR "pocket computers" OR "pda" OR "electronic diary" OR "electronic diaries" OR "personal digital assistant" OR "personal digital assistants" OR "handheld computer" OR "handheld computers" OR "handheld device" OR "handheld devices" OR "electronic assessment" OR "electronic assessments" OR "electronic interview" OR "electronic interviews") AND ("1" [PDat]: "2006/12/31" [PDat]).

APPENDIX II. QUALITY ASSESSMENT TOOL

1. The formulation of the research question.
2. Specification of in- and exclusion criteria.
3. Sampling (random, sufficient size).
4. Description of non-response (intention-to-treat, non-response, attrition).
5. Design (control group, matching).
6. Sample description (sample size and composition, demographics).
7. Reliability and validity of measures.
8. Description of study procedure.
9. Statistical analyses.
10. Research questions are answered.

TABLES AND FIGURES

Table 1A Study and population characteristics

First author (year)	Design[b]	N[c]	Age in years ($M \pm SD, r$)[c]	% female[c]	Country[d]	Diagnosis[e]	Type[f]	Duration (year) ($M \pm SD, r$)
Aaron et al. (2006)	M	137	38 ± 11, $r = 18$–68	88	USA	TMD	C	Md = 5
Turner et al. (2004)	H	(100)		(87)				Md = 0.9, $r = 0.3$–48
Wig et al. (2004)	O	(81)		(81)				6.7 ± 7.8, $r = 0.3$–13
Aaron et al. (2005)	M	(71)		(86)				2, $r = 0.1$–3
Turner et al. (2005)	O	126		88				–
Aaron et al. (2004)	P	(62)		(52)				$r = 1$–11
Affleck et al. (2001)	H	89	44.4 ± 8.8, $r = 25$–66	100	USA	F	C	–
Zautra et al. (2001, study 2)								
Affleck et al. (2000)								
Tennen et al. (2006a)								
Affleck et al. (1998)		(50)	(43.9 ± 8.2)					10.2 ± 8.4
Affleck et al. (1996)								
Badr et al. (2006, study 1)	H	66	56.7 ± 10.2, $r = 38$–73	100	USA	C	C	3.3 ± 1.4
Badr et al. (2006, study 2)	H	33	58.3 ± 11.0, $r = 27$–81	100	USA	C	C	–
Begg et al. (2003)	M	27	48.9, $r = 27$–77	100	UK	Post-surgery	A	n/a
Bolten et al. (1991)	P	16	53.9 ± 8.9	15	Germany	A	C	7.3 ± 4.8

| Study[a] | | | | | | Pain | | |
Population								
First author (year)	Design[b]	N[c]	Age in years $(M \pm SD, r)$[c]	% female[c]	Country[d]	Diagnosis[e]	Type[f]	Duration (year) $(M \pm SD, r)$
Brackertz et al. (1991)	P	16	59 ± 8[g]	59[g]	Switzerland	A	C	7.6 ± 8.7
Feiler et al. (2005)	H	43	51.4 ± 9.5	100	*Germany*	F	C	5.6 ± 3.7
Müller et al. (2004)	H	1	61		Germany			>30
Gaertner et al. (2004)	P	24	49.9 ± 15.1, $r = 16–75$	54	*Germany*	Various	C	2.9, $r = 0.2–10$
Gerbershagen et al. (1995)	O	25	–	–	Germany	A	C	–
Giffin et al. (2003)	H	97	42, $r = 24–69$	95	USA, UK, *Denmark*	H	C	21
Han et al. (2002)	O	79	69.6 ± 7.8	0	*Netherlands*	C	C	–
Harris et al. (2005)	H	125	47.1 ± 11.2	98	USA	F	C	4.1 ± 4.2
Hermann et al. (1995, study 1)	P	49	38.5 ± 11.7	82	USA	H	C	14.1 ± 10.8
Hermann et al. (1995, study 1)	P	33	38.8 ± 12.5	88	USA	H	C	18.4 ± 14.5
Höfer and Battig (1994)	O	120	31.7 ± 7	50	*Swiss*	H	A	n/a
Honkoop et al. (1999)	M	56	41, $r = 18–65$	100	*Netherlands*	H	C	–
Jamison et al. (2001)	P	20	42.1 ± 5.0, $r = 34–54$	45	USA	LBP	C	6.6 ± 5.8, $r = 0.8–26$
Jamison et al. (2006)		(21)	(42.0 ± 4.9)	(43)				
Kikuchi et al. (2006)	P	40	39.0 ± 10.8, $r = 22–60$	78	*Japan*	H	C	–
Kop et al. (2005)	H	38	41.5 ± 8.2	74	USA	F/CF	C	–
Lewis et al. (1994)	O	26	66 ± 10.2[g], $r = 31–83$	58[g]	Australia	A	C	–
Lewis et al. (1995, experiment 3)	P							

Study[a]			Population				Pain		
First author (year)	Design[b]	N[c]	Age in years $(M \pm SD, r)$[c]	% female[c]	Country[d]	Diagnosis[e]	Type[f]	Duration (year) $(M \pm SD, r)$
Liszka-Hackzell and Martin (2004)	H	A:15	A:46 ± 10.6	A:40	USA	LBP	A + C	–
		C:15	C:51 ± 10.2, $r = 18–75$	C:47				
Liszka-Hackzell and Martin (2005)	H	18	52 ± 10.5	44	USA	LBP	C	–
Litt et al. (2004)	H	30	35.8 ± 11.5	87	USA	TMD	C	6.7 ± 5.6, $r = 0.5–26.7$
McLean et al. (2005)	H	P:20	P:43 ± 9	P:80	USA	F	C	–
		H:16	H:39 ± 9	H: 25				
Nived et al. (1994)	O	197	54.5[g]	71[g]	*Sweden*	A	C	–
Nyholm et al. (2004)	P	10	64.7 ± 9.9, $r = 40–75$	40	*Sweden*	P	n/a	13.0 ± 7.4, $r = 5–27$
Ott and Scholz (1999)	P	13	44.1, $r = 28–59$	62	Germany	H/other	C	–
Palermo et al. (2004)	P	30	12.3 ± 2.4	73	USA	A/H	C	–
Peters et al. (2000)	H	80	40.6 ± 6.7, $r = 18–60$	78	*Netherlands*	Various	C	5.5 ± 7.2
Sorbi et al. (2006a)	H							
Sorbi et al. (2006b)	H							
Roelofs et al. (2004)	H	40	46.4 ± 9.9, $r = 30–64$	57	*Netherlands*	LBP	C	12.3 ± 9.8, $r = 0.5–37.0$
Roelofs et al. (2006)	H	38	46.4 ± 9.8, $r = 30–64$	58				12.1 ± 9.8, $r = 0.5–37.0$
Saito et al. (2005)	M	13	37.8 ± 7.6, $r = 23–47$	50	*Japan*	MCS	n/a	–
Salaffi et al. (2005)	P	49	61.2 ± 12, $r = 39–77$	76	Italy	A/F	C	–
Serif and Ghinea (2005)	P	25	Md = 44[g], $r = 27–65$	–	UK	Wheelchair users	n/a	–

Study[a]			Population				Pain		
First author (year)	Design[b]	N[c]	Age in years ($M \pm SD$, r)[c]	% female[c]	Country[d]	Diagnosis[e]	Type[f]	Duration (year) ($M \pm SD$, r)
Sheftell et al. (2005)	O	2696	40.6 ± 10.8	85	USA, Europe, Canada	H	C	–
Smith and Safer (1993)	H	30	$r = 28–53$	77	USA	Various	C	–
Sprott et al. (1993)	M	23	54.5, $r = 36–70$	48	Germany	F	C	–
Sprott and Muller (1998)	M				*Germany*			
Stone et al. (2003a)	P	68	50.9 ± 10.5, $r = 25–75$	85	USA	A/F	C	–
Stone et al. (2004)	P							
Stone et al. (2005)	P							
Litcher-Kelly et al. (2004)	H	(66)	$(51.0 \pm 10.6$, $r = 25–75)$					
Stone et al. (2003b)	P	40	43.0 ± 9.0	47	USA	Various	C	–
Van Gerven et al. (1996)	O	159	48.7, $r = 21–73$	69	*Netherlands*	H	A	n/a
Viane et al. (2004, study 2)	H	62	46.1, $r = 22–65$	81	*Belgium*	F/other	C	–
Walker and Sorrells (2002)	P	11	Md = 8, $r = 6–10$	–	USA	GI	C	–
Williams et al. (2004)	P	14	46.5 ± 13.2, $r = 23–65$	93	USA	F	C	5.5 ± 4.7, $r = 0.3–17$
Gendreau et al. (2003)	O							

Notes: (1) Information is only reported for participants who completed e-diaries, except when compliance information was reported for patients and healthy controls together. (2) For studies reporting from the same dataset, empty cells indicate that the information of the study that is described on the previous line applies (i.e., data was pooled across studies). (3) Study characteristics in parenthesis are not used in the statistical analyses.

[a] Studies reporting from the same dataset are grouped together.

[b] M, monitoring; H, hypothesis testing; P, psychometric; O, e-diary as outcome variable.

[c] A, patients with acute pain; C, patients with chronic pain; H, healthy controls; P, patients; N, sample size; M, mean; SD, standard deviation; r, range; Md, median.

[d] Papers were written in the language of the country where the study was carried out, except where country is printed in italics the paper was in English language.

[e] A, arthritis (ankylosing spondylitis, juvenile idiopathic arthritis, osteoarthritis, rheumatoid arthritis); C, cancer pain; CF, chronic fatigue; F, fibromyalgia; GI, gastrointestinal symptoms; H, headache (migraine, tension headache); LBP, low back pain; P, Parkinson's disease; TMD, temporomandibular disorder; MCS, multiple chemical sensitivity.

[f] C, chronic; A, acute; H, healthy control.

[g] Calculated for individuals who started the diary assessment, including dropouts.

Table 1B. Electronic diary and sampling characteristics, and study quality

Study[a]	Sampling		Compliance e-diary						Quality[b]
First author (year)	Response (n/N)	Attrition (n/N)	%	Strategy[c]	Device	Period in days, scheme [d]	No. items[e]	Assessment[f]	
Aaron et al. (2006)	158/366	21/158	90	abcdefgh	Palm	14, 3T	35	abfghio	H
Turner et al. (2004)	(110/244)	(10/110)							H
Wig et al. (2004)		(29/110)							H
Aaron et al. (2005)	(97/226)	(26/97)							H
Turner et al. (2005)	158/366	32/158	90			56, 3T			H
Aaron et al. (2004)	(76/164)	(14/76)							H
Affleck et al. (2001)	89/203	–	99	abcefgho	Psion	30, 3S + 1E	33.3	abdhjo	H
Zautra et al. (2001, study 2)									H
Affleck et al. (2000)									L
Tennen et al. (2006b)									L
Affleck et al. (1998)	50/71		98						H
Affleck et al. (1996)									L

Study[a]	Sampling		Compliance e-diary						Quality[b]
First author (year)	Response (*n/N*)	Attrition (*n/N*)	%	Strategy[c]	Device	Period in days, scheme[d]	No. items[e]	Assessment[f]	
Badr et al. (2006, study 1)	–	–	69	bce	Palm	7, 4S	18	ahj	L
Badr et al. (2006, study 2)	–	33/42	86	bo	Palm	±21, 4S	18	ahjo	L
Begg et al. (2003)	–	5/32	75	bh	Psion/Apple	28, 2P	16	afhj	L
Bolten et al. (1991)	–	4/20	98	h	Indamos	14, 2T	8	acjo	H
Brackertz et al. (1991)	–	1/17	63	e	Indamos	21, 6T	50	acjk	L
Feiler et al. (2005)	–	14/57	–	e	Psion	77, 1T	16	afhij	L
Müller et al. (2004)	n/a	n/a	95			105, 1T			H
Gaertner et al. (2004)	28/52	4/28	92	bdeh	Palm	14, 1P	8	ao	L
Gerbershagen et al. (1995)	–	8/33	–	–	–	28, 4E	28	acjo	L
Giffin et al. (2003)	–	23/120	60	ef	Philips	30, 1S + *x*P	3.8	afko	H
Han et al. (2002)	111/130	32/111	–	–	HP 95LX	98, 2T	–	ac	H
Harris et al. (2005)	–	–	–	e	Palm	98, *M* = 3.4S	1	a	L
Hermann et al. (1995, study 1)	–	–	72	b	Atari	14, 3T + 1E	3	aho	H
Hermann et al. (1995, study 2)	–	–	86	b	Atari	14, 3T + 1E	3	aho	H
Höfer and Battig (1994)	–	29/149	–	abcdho	Psion	14, 6T	17.2	ahjo	H
Honkoop et al. (1999)	73/687	17/73	80	e	–	70, 6S	8	abco	H

Study[a]	Sampling		Compliance e-diary						Quality[b]
First author (year)	**Response (n/N)**	**Attrition (n/N)**	**%**	**Strategy[c]**	**Device**	**Period in days, scheme [d]**	**No. items[e]**	**Assessment[f]**	
Jamison et al. (2001)	21/48–16[g]	1/21	72	bcdo	Apple	365, 1P	17	af	H
Jamison et al. (2006)		(0/21)	(–)						H
Kikuchi et al. (2006)	44/66	4/44	96	begho	Seiko watch	7, 2P + 4S	1	a	H
Kop et al. (2005)	–	–	–	e	Actigraph	5, 3S + 2T	3	ajo	H
Lewis et al. (1994);	–	10/36	88	e	PIPER	70, 4T	1	a	H
Lewis et al. (1995, experiment 3)									L
Liszka-Hackzell and Martin (2004)	–	11/41	81	bce	Mayo Clinic diary	21, 10T + xP	1	a	L
Liszka-Hackzell and Martin (2005)	–	1/19	–	e	Mayo Clinic diary	6, 10T + xP	1	a	L
Litt et al. (2004)	–	0/30	81	be	Psion	7, 4S	21	abghi	H
McLean et al. (2005)	–	8/28	98	e	Actigraph	2, 3S + 2T	3	ajo	L
Nived et al. (1994)	–	110/307	–	ef	MiniDoc	28, 1T	3	ao	H
Nyholm et al. (2004)	–	–	88	o	Ericsson	8 of 28, 7T	10.1	afhjo	L
Ott and Scholz (1999)	–	–	61	b	Atari	14, 2T	13	abcjk	L
Palermo et al. (2004)	–	5/35	83	abceo	HP	7, 1T	19	abcfh	L
Peters et al. (2000)	–	7/87	85	abcdefho	–	28, 2T + 4S	63	aefhijo	H
Sorbi et al. (2006a)									L

Study[a]	Sampling		Compliance e-diary						Quality[b]
First author (year)	Response (*n/N*)	Attrition (*n/N*)	%	Strategy[c]	Device	Period in days, scheme [d]	No. items[e]	Assessment[f]	
Sorbi et al. (2006b)									L
Roelofs et al. (2004)	–	–	76	abdeg	Palm	7, 8S	37	adeo	H
Roelofs et al. (2006)			73			21, 8S			H
Saito et al. (2005)	–	5/18	62	be	Seiko watch	7, 2S + xE	28	aho	H
Salaffi et al. (2005)	49/66	–	94	be	Data Logger	7, 7T	2	ao	L
Serif and Ghinea (2005)	–	–	–	bh	HP	5, 4T	13	abc	L
Sheftell et al. (2005)	–	635/3331	–	b	–	1 post-event, 3E + 2S	19	acko	H
Smith and Safer (1993)	–	1/31	n/a	bdho	Symtrack Logger	±7, M = 42.6E	4	ajko	L
Sprott and Muller (1998)	–	9/23	61	e	Kanoldt	18, 2T	42	abhjo	L
Sprott and Muller (1998)									L
Stone et al. (2003a)	95/322	4/95-23[g]	94	abefgo	Sony	14, 3/6/12S	19	acefh	H
Stone et al. (2004)									H
Stone et al. (2005)									H
Litcher-Kelly et al. (2004)									H
Stone et al. (2003b)	84/230	2/84-42[g]	94	abefgo	Sony	21, 3T + xE	20	acefh	H
Van Gerven et al. (1996)	166/303	7/166	99	be	MiniDoc	4h post-event, 8S	2.9	ako	H
Viane et al. (2004, study 2)	–	11/73	88	abdeh	Palm	14, 8S + 2E	8	ado	H

Study[a]	Sampling		Compliance e-diary							Quality[b]
First author (year)	**Response (n/N)**	**Attrition (n/N)**	**%**	**Strategy[c]**	**Device**	**Period in days, scheme [d]**	**No. items[e]**	**Assessment[f]**		
Walker and Sorrells (2002)	11/12	–	100	bde	Compaq Aero	7, 1T	9	ao		L
Williams et al. (2004)	–	5/14	85	beg	Palm	84, 4–6S	1	a		L
Gendreau et al. (2003)										L

Notes: (1) Information is only reported for participants who completed e-diaries, except when compliance information was reported for patients and healthy controls together. (2) Empty cells indicate that the information described for the above study is applicable (i.e., data was pooled across studies). (3) Study characteristics in parenthesis are not used for the statistical analyses.

[a] Studies reporting from the same dataset are grouped together.

[b] H, above average quality; L, below average quality.

[c] a, financial compensation; b, instruction/training; c, phoned by researcher; d, help-desk; e, alarm; f, repeat alarm; g, postpone diary; h, manual; o, other.

[d] T, time contingent (e-diary on fixed times); E, event-contingent (dependent on occurrence of event, e.g., waking-up, pain episode); S, signal-contingent (in response to *random* signal, may occur within predetermined time-windows); P, person-contingent (i.e., participant initiates the e-diary entry, minimum number of entries required is mentioned); *x*, number of entries depended on occurrence of an event or on the participant; h, hours; *M*, mean.

[e] Mean number of items per daily diary.

[f] a, pain intensity; b, pain location; c, pain quality; d, attention to pain; e, pain-related fear; f, interference with activities/limitations/disability; g, coping; h, mood; i, cognitive aspects of pain (beliefs, attributions, catastrophising, self-efficacy); j, sleep quality, fatigue; k, health care utilisation (medication, visit health care provider); o, other (e.g., nausea, stress, well-being).

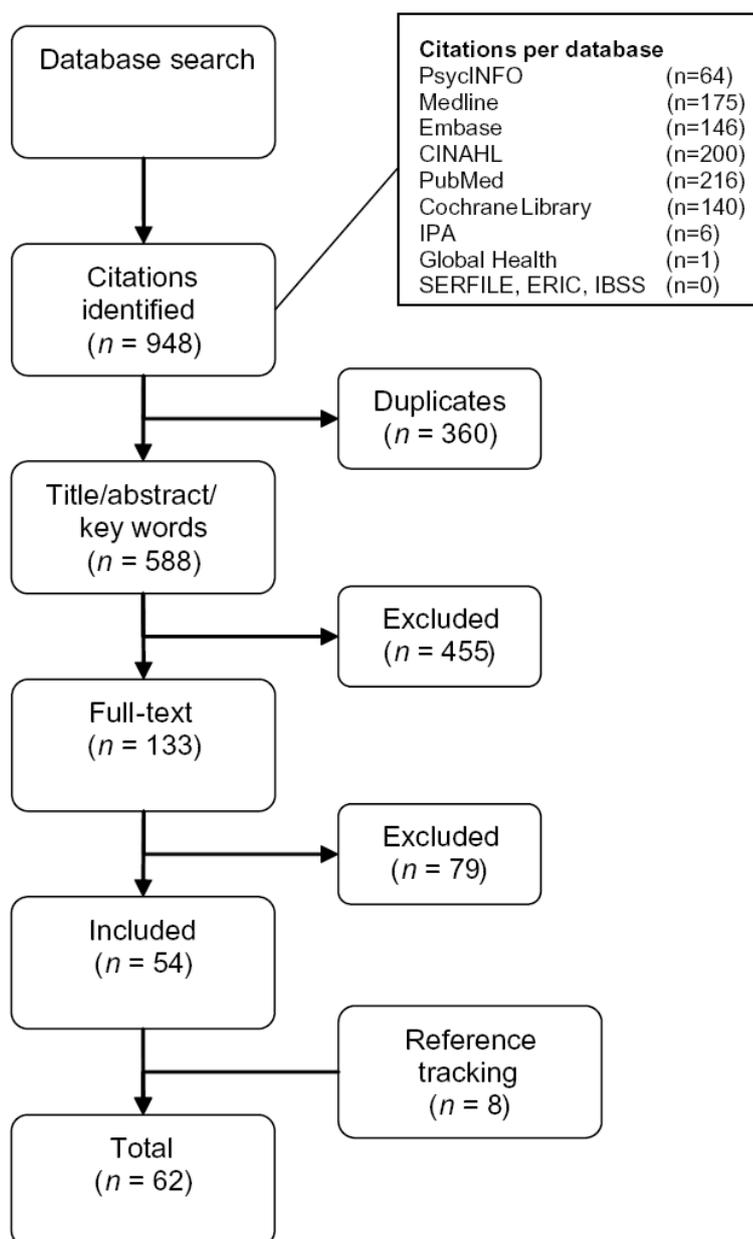**Fig. 1.** Flow diagram of identification and inclusion of papers.

Table 2. Study, population, diary, and sampling characteristics

	N	**%**
General characteristics		
Year of publication (*n* = 62)		
1991–1999	16	25.8
2000–2006	46	74.2
Language of paper (*n* = 62)		

	N	%
English	55	88.7
Non-English	7	11.3
Study quality (*n* = 62)		
Below average	27	43.5
Above average	35	56.5
Mean (SD)	11.7 (3.8)	
Study design (*n* = 53)		
Primary, hypothesis testing	25	40.3
Primary, monitoring	7	11.3
Secondary, psychometric	19	30.6
Secondary, outcome variable in trial	11	17.7
Population characteristics (n = 43)		
Sample size		
1–20	10	23.3
21–100	27	62.8
>100	6	14.0
Gender		
>50% female	30	69.8
>50% male	8	18.6
50% female, 50% male	2	4.7
Mean % female (SD)	69.9% (24.2)	
Age (years)		
18	2	4.7
19–44	18	41.9
45–65	19	44.2
65	2	4.7
Pain		

	N	%
Chronic	37	86.0
Acute	3	7.0
Other	3	7.0
Country of origin		
USA	19	44.2
Europe	19	44.2
Oceania	1	2.3
Asia	2	4.7
USA, Canada, and Europe	2	4.7
Diary characteristics		
Assessment period (*n* = 46)		
<1 week	6	13.0
1–2 weeks	19	41.3
3–4 weeks	10	21.7
>4 weeks	11	23.9
Assessment scheme (*n* = 43)		
Time-contingent	23	53.5
Signal-contingent	18	41.9
Event-contingent	9	20.9
Person-contingent	7	16.3
Combines 2 schemes	14	32.6
Number of diaries per day (*n* = 43)		
1–2	12	27.9
3–6	24	55.8
7–10	7	16.3
Mean (SD)	4.3 (2.6)	
Diary length (*n* = 42, 1 missing)		
1 item	6	14.3

	N	%
2–10 items	14	33.3
11–20 items	13	31.0
>20 items	9	21.4
Mean (SD)	14.9 (14.5)	
Pain aspects assessed (*n* = 43)		
Pain intensity	43	100.0
Pain location	8	18.6
Pain quality	11	25.6
Attention to pain	3	7.0
Pain-related fear	3	7.0
Interference with activities/limitations/disability	10	23.3
Coping	2	4.7
Mood	17	39.5
Cognitive aspects of pain	4	9.3
Sleep quality, fatigue	16	37.2
Health care utilisation (consultations, medication)	6	14.0
Other	25	58.1
Sampling characteristics		
Response rate (*n* = 43)		
Reported	12	27.9
Mean (*SD*) if reported	52.8% (23.5)	
Attrition rate (*n* = 43)		
Reported	30	69.8
Mean (SD) if reported	19.0% (15.6)	
Response and attrition reported	9	20.9
Compliance rate (*n* = 41)[a]		
Reported	33	76.7
Mean (SD) if reported	83.4% (12.6)	

	N	%
Compliance strategies ($n = 43$)		
Notification/alarm	29	67.4
Training/instruction	28	65.1
Manual	11	25.6
Financial compensation	9	20.9
Help-desk	9	20.9
Phone call by researcher	8	18.6
Postpone option	7	16.3
Reminder	7	16.3
Other	11	25.6
Mean number of strategies (SD)	2.8 (2.2)	

[a] Müller et al. (2004) was used instead of Feiler et al., 2005

Table 3. Final model of the regression analysis to identify predictors of compliance

Study characteristic	**B**	**(95% CI)**	**p**	**F**	**R²**
Constant	58.93	(39.69; 78.18)	<0.001		
Diary length (items)	−0.48	(−0.73; −0.24)	<0.001		
Manual	11.53	(3.72; 19.34)	0.005		
Other	9.02	(0.52; 17.51)	0.038		
Mean age (years)	0.35	(0.01; 0.70)	0.047		
Alarm	9.30	(0.02; 18.59)	0.050		
Financial compensation	9.10	(−1.25; 19.44)	0.082		
Model			*<0.001*	*6.04*	*0.59*

$N = 32$.