Ecological momentary assessment for chronic pain in fibromyalgia using a smartphone: A randomized crossover study

A. García-Palacios¹,², R. Herrero¹, M.A. Belmonte³, D. Castilla¹, J. Guixeres⁴, G. Molinari¹, R.M. Baños²,⁵, C. Botella¹,²

¹ Departamento de Psicología Básica, Clínica y Psicobiología, Universitat Jaume I, Castellón, Spain
² CIBER Fisiopatología Obesidad y Nutrición (CB06/03), Instituto Salud Carlos III, Santiago de Compostela, Spain
³ Servicio de Reumatología, Hospital General de Castellón, Castellón, Spain
⁴ LabHuman, Universidad Politécnica de Valencia, Spain
⁵ Departamento de Personalidad, Evaluación y Tratamientos Psicológicos, Universidad de Valencia, Spain

Abstract

Background: Daily diaries are a useful way of measuring fluctuations in pain-related symptoms. However, traditional diaries do not assure the gathering of data in real time, not solving the problem of retrospective assessment. Ecological momentary assessment (EMA) by means of electronic diaries helps to improve repeated assessment. However, it is important to test its feasibility in specific populations in order to reach a wider number of people who could benefit from these procedures.

Methods: The present study compares the compliance and acceptability of an electronic diary running on a smartphone using a crossover design for a sample with a specific pain condition, fibromyalgia and low familiarity with technology. Forty-seven participants were randomly assigned to one of two conditions: (1) paper diary – smartphone diary and (2) smartphone diary – paper diary, using each assessment method for 1 week.

Results: The findings of this study showed that the smartphone diary made it possible to gather more accurate and complete ratings. Besides, this method was well accepted by a sample of patients with fibromyalgia referred by a public hospital, with an important proportion of participants with low level of education and low familiarity with technology.

Conclusions: The findings of this study support the use of smartphones for EMA even in specific populations with a specific pain condition, fibromyalgia and low familiarity with technology. These methods could help clinicians and researchers to gather more accurate ratings of relevant pain-related variables even in populations with low familiarity with technology.

1. Introduction

Chronic pain requires a multidimensional perspective (Flor and Turk, 2011). A useful procedure to measure symptom fluctuations is daily diaries, which includes fixed-interval assessments. Shiffman et al. (2008) include them as a special case of ecological momentary assessment (EMA), defined as ‘methods using repeated collection of real-time data on subjects’ behaviour and experience in their natural environment’ (p. 3). Daily diaries gather repeated data but not in real time, not solving the problem of retrospective assessment.

Retrospective assessment tends to produce overestimation of the frequency, intensity and duration of the

Correspondence
Guadalupe Molinari
E-mail: molinari@uji.es

Funding sources
The research presented in this paper was funded in part by Fundació La Marató de TV3 (Ajuts de la Marató de TV3 2006), Plan de Promoció de la Investigación Universitat Jaume I, (Fundacio Caixa Castelló-Bancaixa, P11B2009-30) and by Generalitat Valenciana, Redes de Excelencia ISIC (ISIC/2012/012).

Conflicts of interest
None declared.

Accepted for publication
20 October 2013

doi:10.1002/j.1532-2149.2013.00425.x
What's already known about this topic?
- The use of ecological momentary assessment and real-time data capture (RTDC) in the field of chronic pain is a very useful tool for clinicians and researchers because it makes gathering of more accurate and complete ratings of relevant variables possible.

What does this study add?
- The findings of this study contribute with data supporting the use of smartphones for RTDC in a sample of patients with fibromyalgia, with an important proportion of participants with low educational levels and low familiarity with technology.
- It presents a technological tool that has been submitted to a process of careful evaluation in order to improve traditional methods of assessment in chronic pain through RTDC with the aim of reaching a specific population that could be at risk of not benefiting from technological innovations because of the digital divide.

2. Methods

2.1 Participants

Fig. 1 shows a flow chart diagram including the recruitment process. The exclusion criteria for this study were suffering a severe mental illness or severe sensory impairments (visual, motor or hearing).

Seventy-four participants were screened, and a total of 47 women participated in the study; their ages ranged from 37 to 65 [M = 48.05; standard deviation (SD) = 7.95]. All participants were volunteers and were recruited from Castellon General Hospital in Spain. All participants met the criteria for FMS according to the American College of Rheumatology (Wolfe et al., 1990) and were diagnosed by a rheumatologist.

With respect to the educational profile, 10% had not finished elementary education (less than 8 years of education) but they could write and read, 47.5% had elementary education, 25% had high school education and only 17.5% had a university degree.

Regarding familiarity with technology, 17.5% of participants had no experience with computers, 37.5% had used computers just a few times (less than 10 times) and 45% usually used computers. With regard to Internet use, 27.5% had never used the Internet, 27.5% had used it a few times (less than 10 times) and 45% used it at least twice a week. In reference to the use of mobile phones, 100% of the sample used mobile phones, but 22.5% of the sample did not know how to read short message service (SMS) and 37.5% did not know how to write a SMS. Also, 32.5% of the sample had used a device with a touch screen at least once; the rest had never used a device with a touch screen.
2.2 Measures

2.2.1 Three types of measures were included in this study

(1) EMA measures. We included the recording of three key variables in the study of pain: pain intensity, fatigue intensity and mood. Pain and fatigue were rated on 0–10 Numerical Rating Scales from ‘no pain/fatigue’ to ‘worst pain/fatigue you can imagine’. Mood was assessed with a face-based pictorial 7-point scale. A time-based approach with fixed intervals was chosen for this study. Participants were asked to complete these three ratings three times a day. These measures were analysed following two principles: the presence or lack of presence of the rating and the compliance with the time frame in which the participant had to fill out the records. In this sense and according to Stone et al. (2003), a record completed outside the specified time range was treated as one that was not completed within ±30 min of the exact time. According to this rule, records were classified into four categories: (1) complete record: a record completed in the stipulated time; (2) complete record, out of time: a record completed outside of the time range; (3) incomplete record: a record not totally completed, with at least one piece of data missing; and (4) totally incomplete record: a totally missed record.

(2) Weekly measures of pain and fatigue. Patients completed the Brief Pain Inventory (BPI; Cleeland and Ryan, 1994; Badia et al., 2003) and the Brief Fatigue Inventory (BFI; Mendoza et al., 1999) once a week in order to gather a retrospective rating of average pain/fatigue intensity. The scales from these inventories asking for average pain/fatigue intensity in the last week were the ones chosen for this study.

(3) Self-report inventories to assess the two EMA conditions. These measures were designed for this study. The first was a questionnaire to evaluate each condition separately (acceptability questionnaire), which consisted of 14 items with a range of responses from 1 ‘totally agree’ to 5 ‘totally disagree’. This questionnaire was administered by the assessor, who requested the opinion of the participant about different relevant areas regarding acceptability. The second measure was a comparison questionnaire developed in order to discover the participants’ preferences among the two conditions (preferences questionnaire). This questionnaire was administered by the assessor once the study ended and consisted of nine items with three different options of response ‘traditional’, ‘mobile device’ or ‘indistinct’. The assessor explained the instructions of the questionnaires and then the participants filled the questionnaires by themselves. The items included in these inventories are displayed in Tables 3 and 5.

Additionally, a technological profile questionnaire was administered in order to assess the experience of the sample in the use of technology. The questionnaire was designed for this study, following the guidelines used in other studies developed in our laboratory (Etchemendy, 2011). This questionnaire allows defining the degree of knowledge and frequency of use of technology. The questionnaire has two sections. The first one is composed of two questions that assess the use of computers and the Internet on a scale ranging from ‘1 = never’ to ‘5 = often’. The second section is composed of six questions assessing the frequency of use of mobile devices and the ability of the participant to use them. This section asks if the participant has a mobile phone; if he...
or she knows how to make a phone call, how to read messages and how to write a message; if he or she has ever used Internet on the mobile phone; and if he or she has ever used devices with touch screen (e.g., smartphones or automated teller machines).

2.3 Materials

Two different modalities of the assessment procedures were developed: one using a mobile device and the other using a paper-and-pencil traditional diary. All measures included in the assessment procedures were in Spanish. They have been translated into English for publication in this journal.

Our research team developed a software application (F-EMA) running on a mobile device. The hardware used was a smartphone HTC Diamond 1 (TOUCH Diamond 1, HTC Corporation, New Taipei City, Taiwan) with the following specifications: 51 × 102 × 11.5 mm; read-only memory 4352 MB; random access memory 192 MB; 480 × 640 display resolution; 2.8” display diagonal; 16 bit per pixel display colour depth; audio stereo sound. The smartphones were provided to the participants by the research team and were returned after the study completion. The software was run on Windows Mobile 6.1. In Fig. 2, we offer a picture of the mobile application. The reason for using a mobile phone was that the idea was to develop an application that could be used by patients in their natural environments.

The assessment was carried out three times a day. The default schedule was set at 9:00 a.m., 3:00 p.m. and 9:00 p.m. The system allowed these times to be adjusted according to the particular needs of each participant. An audio signal indicated that the participant should fill out the rating scales. If the user did not complete them, the audio signal sounded again every minute during the first 15 min and then every 15 min during the following hour. After that time, the application considered that the user was not able to answer and the assessment was not performed. The application included the option of not only seeing the images of the scales but also listening to audio-recorded instructions, which were included with the intention of making the system easier to use for a wider number of people (e.g., elderly people or people with some visual impairment).

Usability studies were carried out. Results showed that F-EMA was an easy tool to use and to learn to use (Castilla et al., 2012).

A traditional pencil-and-paper diary was also designed including the same scales as those on the mobile device (see Fig. 3). The only difference between the two conditions was that the mobile device automatically recorded the time at which the participant answered, whereas in the traditional self-record, the participants had to fill in the time of the day at which they completed the rating.

2.4 Procedure

This is a single-centre, randomized, crossover study with FMS sufferers. Participants were recruited from the rheumatology unit of Castellon General Hospital in Spain. Approval was granted by the institutional review board. All participants attended voluntarily and signed an informed consent form. The participants completed a brief interview, where information about demographics and their clinical status was gathered, as well as a technological profile.

After the initial assessment, which comprised a 1-h session, participants were randomly assigned to one of two conditions: (1) P: paper-and-pencil diary – smartphone diary; (2) S: smartphone diary – paper-and-pencil diary (see Fig. 4). At the start of the first week, participants attended an individual information session (S1) in which the experimenter provided the corresponding self-record (paper vs. smartphone) and provided verbal instructions about the use of the self-record. Experimenters explained to the participants that they were going to be asked to assess three important aspects in the field of chronic pain: pain intensity, fatigue intensity and mood. The experimenter explained the meaning of the scales and the way in which they should be
rated. Participants had to fill in the self-record three times a day every day, so the experimenters asked for three different hours during the day (morning, afternoon and evening) that were convenient in the daily schedule of each participant. The participants practised rating the scales with the experimenter, and finally an information sheet with the meaning of each scale and instructions for filling in the record was given to each participant. After the practice, all participants were able to use the paper diary and the smartphone diary. Participants recorded the assessments daily in their natural environments for 7 days.

At the end of the first week, a second session was held (S2). In this session, experimenters received the self-record data from the participants, administered the acceptability questionnaire regarding the self-record procedure used during the week (paper or smartphone) and performed a weekly rating of pain and fatigue (average pain intensity and fatigue measured by the BPI and BFI). Then each participant received the other self-record (paper or smartphone). The experimenters explained the procedure, and the participants practised the method of rating the scales and were given an information sheet with the instructions. They recorded the assessments daily in their natural environments for 7 more days.

At the end of the second week, a third session was held (S3). In this session, experimenters received the self-record diaries and administered the acceptability questionnaire regarding the self-record procedure used that week and the preference questionnaire in order to compare both conditions. Participants also gave weekly ratings of average fatigue and pain intensity using the scales included in the BPI and BFI.

After the completion of the study, participants were offered to attend a free psychological treatment for FMS during 3 weeks (six group 2-h sessions).

### 3. Results

#### 3.1 Compliance with two EMA methods: smartphone versus paper and pencil

The first objective of this study was to compare the compliance with a traditional paper-and-pencil diary with a smartphone diary. \(t\)-tests for related samples
were conducted in order to compare the adherence of participants with the instructions in both conditions, using the four categories set out.

The results showed significant differences in three of the four categories (see Table 1). Significant differences were obtained in complete records, showing a higher number of complete records in the smartphone condition than in the traditional condition with a large effect size. A significant difference was found between the two conditions regarding incomplete records, showing a higher number of incomplete records in the traditional condition with a large effect size. Significant differences were also obtained regarding the totally incomplete records, showing a higher number of totally incomplete records in the smartphone condition, although the effect size here was medium. Finally, no significant differences were found regarding records completed out of time.

We would like to highlight that, taking into account that the total number of records was 21 per week (three per day for 7 days), the rate of complete records was much higher with the use of the mobile device (18.2: 86.66%) compared with the use of the traditional diary (11.12: 52.95%).

We also conducted correlations between aggregated data of the paper diary and of the smartphone, regarding pain intensity, \( r = 0.79, p < 0.001 \) and regarding fatigue, \( r = 0.88, p < 0.001 \). In both cases, the correlations were positive and statistically significant.

### 3.2 Relationship between aggregated EMA data and retrospective data

Another of our objectives was to compare recall-based and real-time data. To achieve this goal, we compared aggregated EMA data with the retrospective rating of pain intensity and pain fatigue that the participants reported once a week during the study. The mean ratings and SD are reported in Table 2. \( t \)-tests revealed significant differences between the recall-based ratings and the EMA data in both pain and fatigue intensity. Aggregated EMA data using the traditional diary were lower than the retrospective ratings regarding pain intensity and fatigue intensity. Aggregated EMA data using the smartphone were also lower than the recall-based ratings regarding pain intensity and fatigue intensity. That is, participants tended to describe their symptoms as more intense when they gave retrospective weekly ratings.

We also examined the correlation between aggregated EMA and recall-based data for pain and fatigue intensity. Positive and significant correlations were found between aggregated EMA data and recall-based data regarding pain rated with the paper diary and the weekly average pain intensity \( (r = 0.59; \ p < 0.001) \) and with the smartphone diary and the weekly average pain intensity \( (r = 0.39; \ p < 0.02) \). With regard to fatigue, correlations were calculated for fatigue intensity rated with the paper diary and the weekly measure \( (r = 0.67; \ p < 0.001) \), and fatigue intensity rated with the smartphone and the weekly measure of fatigue intensity \( (r = 0.47; \ p < 0.01) \). In all cases, correlations were positive and statistically significant.

### Table 1: \( t \)-test for related samples to compare the adequacy with the instruction in both conditions.

|                      | Paper | Smartphone |
|----------------------|-------|------------|
|                      | \( n \) | \( M \) | \( SD \) | \( M \) | \( SD \) | \( t \) | \( d^a \) |
| Complete record      | 21    | 11.12      | 9.13      | 18.2   | 3.49   | -4.446** | 1.02    |
| Complete record out of time | 21    | 0.98       | 2.35      | 1.36   | 3.01   | -0.654   | 0.14    |
| Incomplete record    | 21    | 8.57       | 9.61      | 0.00   | 0.00   | 5.642**  | 1.26    |
| Totally incomplete record | 21    | 0.33       | 0.79      | 1.44   | 2.13   | -3.093** | 0.69    |

M, mean; SD, standard deviation.

**\( p < 0.01 \).

*\( d = 0.2 \) are regarded as ‘small’ effect size, \( d = 0.5 \) as ‘medium’ and \( d = 0.8 \) as ‘large’.

### Table 2: Comparisons between aggregated EMA data and recall-based data.

|                      | \( M \) | \( SD \) | \( t \) | \( d^a \) |
|----------------------|---------|---------|-------|---------|
| Pain                 |         |         |       |         |
| Aggregated pain paper diary | 5.98    | 1.36    |       |         |
| Average pain in the last week (BPI) | 7.51     | 1.54    | -7.273** | 1.05    |
| Aggregated pain smartphone diary | 5.57    | 1.44    |       |         |
| Average pain in the last week (BPI) | 7.13     | 1.49    | -6.082** | 1.06    |
| Fatigue              |         |         |       |         |
| Aggregated fatigue paper diary | 5.17    | 1.78    |       |         |
| Average fatigue in the last week (BFI) | 7.13     | 2.05    | -7.843** | 1.02    |
| Aggregated fatigue smartphone diary | 5.24    | 1.73    |       |         |
| Average fatigue in the last week (BFI) | 6.80     | 2.12    | -4.917** | 0.81    |

BFI, Brief Fatigue Inventory; BPI, Brief Pain Inventory; EMA, ecological momentary assessment; M, mean; SD, standard deviation.

**\( p < 0.01 \).

*\( d = 0.2 \) are regarded as ‘small’ effect size, \( d = 0.5 \) as ‘medium’ and \( d = 0.8 \) as ‘large’.
3.3 Acceptability of the two EMA procedures

Our final aim was to explore the acceptability of the EMA methods (paper-and-pencil vs. smartphone). In order to analyse the acceptability, satisfaction and preference between the two conditions, t-tests for related samples were conducted, comparing the answers to each item of the acceptability questionnaire and the preference questionnaire. Regarding the acceptability questionnaire, the results showed significant differences in six items (see Table 3): The smartphone condition was perceived as easier to use and faster to answer. The participants perceived that the instructions of the smartphone condition were significantly easier to follow than those of the paper condition, even when in both conditions the instructions were exactly the same. Regarding the general opinion of the participants about the two conditions, the smartphone method was evaluated as significantly easier and more useful than the paper method. Finally, significant differences in opinions were found regarding whether other people with the same condition should use the assessment procedures, showing that the smartphone was more highly recommended by the participants than the paper-and-pencil diary.

We also examined whether acceptability was different between participants familiar versus less familiar with the use of technology. Participants were divided into two groups based on their answers to the technology profile questionnaire. The first group was composed of those participants with some experience in the use of smartphones, computers and Internet (n = 20), whereas the second group was composed of those participants without any experience or with a low familiarity or no familiarity with the use of smartphones, computers and Internet (n = 20).

In Table 4, we show the results of this comparison. Results showed no differences between the groups in most acceptability questions. There were only significant differences in two items. Participants with more familiarity with technology considered that the smartphone was easier to use than participants with less familiarity with technology (Cohen’s d = 1.01). Also, participants with less familiarity reported that the use of smartphone in front of people was more annoying than the participants more familiar with technology (Cohen’s d = 0.69).

In Table 5, we show the percentage of participants reporting a preference for one of the two methods (paper vs. smartphone) in several domains. The results showed that participants preferred the smartphone method in general (65% vs. 15%); a higher percentage of them also thought it was possible to answer faster with the smartphone (70% vs. 17.5%), and that the smartphone method was more useful (50% vs. 5%). The smartphone was also considered easier to remember to fill in (90% vs. 2.5%), more comfortable to complete (55% vs. 2.5%) and more comfortable to carry (85% vs. 2.5%). A higher number of participants also reported that the paper diary bothered them more (45% vs. 10%). When dividing the sample in participants with familiarity versus less familiarity with technology, the percentages were similar in all the domains assessed but in two of them. There were a

Table 3 t-test for related samples: Acceptability questionnaire.

| It was easy to use.* | Paper | 40 | 2.18 | 0.844 | 1.65 | 0.770 | 3.557** | 0.656 |
|----------------------|-------|----|------|-------|------|-------|--------|-------|
| People could learn to use it. | 40 | 1.58 | 0.931 | 1.35 | 0.622 | 1.597 | 0.290 |
| I felt safe responding. | 40 | 1.67 | 0.737 | 1.46 | 0.854 | 1.750 | 0.263 |
| I could answer fast. | 40 | 1.50 | 0.847 | 1.18 | 0.446 | 2.242* | 0.472 |
| It was comfortable to use. | 40 | 3.38 | 0.747 | 3.44 | 0.821 | -0.374 | -0.076 |
| It hindered the daily activities. | 40 | 4.33 | 0.898 | 4.54 | 0.854 | -1.071 | -0.240 |
| The fact that people could see me using it annoyed me. | 40 | 4.38 | 0.963 | 4.46 | 0.913 | -0.595 | -0.085 |
| It can be used in any context. | 40 | 1.97 | 1.174 | 1.71 | 1.063 | 1.326 | 0.000 |
| Instructions were easy. | 40 | 1.13 | 0.478 | 1.13 | 0.339 | 2.731* | 0.352 |
| The font size was sufficient. | 40 | 1.32 | 0.525 | 1.16 | 0.570 | 1.639 | 0.000 |
| In general, using the self-record was easy. | 40 | 1.90 | 0.778 | 1.28 | 0.506 | 5.922** | 0.944 |
| In general, using the self-record was useful. | 40 | 2.41 | 1.141 | 2.08 | 0.870 | 2.179* | 0.325 |
| People like me should use the self-record. | 40 | 2.53 | 1.219 | 2.10 | 0.955 | 2.429* | 0.392 |

M, mean; SD, standard deviation.
* p < 0.05.
** p < 0.01.
*Each item was rated from 1 to 5, where 1 was ‘totally agree’ and 5 ‘totally disagree’.
lower percentage of participants more familiar with technology who reported that the smartphone was more comfortable to fill (45% vs. 70% in the group with low familiarity with technology). Also, more participants with higher familiarity reported that the smartphone bothered them more (15% vs. 5%).

### 4. Discussion

RTDC methods can improve the accuracy and validity of traditional assessment methods. However, it is important to test its feasibility in specific populations in order to reach a wider number of people who could benefit from them. This work is a contribution to this field.

Using a randomized crossover design, participants used a smartphone diary versus a paper-and-pencil diary. The correlations between the two diaries were high, meaning that the use of a smartphone did not produce important changes in the measure. However, when comparing the frequency of complete and incomplete records, the smartphone condition showed higher levels of compliance than the paper condition. Daily diaries were introduced to obtain ratings of key variables in real-time and in natural environments, thus reducing the bias of retrospective data. The fact that the person does not complete the record on time could have a negative influence on the validity of the data. Therefore, we consider on-time completion of the ratings to be an important variable. The results confirm that the smartphone diary obtained a higher compliance. We believe the features of the smartphone application (use of alarm, audio, etc.) contributed to this finding. We would like to point out that, although we were able to have control over the actual compliance with the smartphone, we were not able to have control over the actual compliance of the paper diary, having to rely on reported compliance. When participants used the paper diary, we did not have a way of knowing whether they filled in the ratings at the specified times or whether they filled the diary forward or backward. Stone et al. (2003) included a

| Table 4 Acceptability questionnaire: participants familiar versus less familiar with technology. |
|-------------------------------------------------|---------------------------------|-----------------|-----------------|-----------------|
|                                                  | LFT                             | FT               |                  |                  |
|                                                  | n                  | M (SD)            | n                  | M (SD)            | t               |
| It was easy to use.*                             | Paper               | 20                | 2.25 (0.85)        | 20                | 2.10 (0.85)       | 0.56 |
|                                                  | Smartphone          | 20                | 2.00 (0.80)        | 20                | 1.30 (0.57)       | 3.20** |
| People could learn to use it.                    | Paper               | 20                | 1.50 (0.83)        | 20                | 1.65 (1.04)       | –0.51 |
|                                                  | Smartphone          | 20                | 1.45 (0.76)        | 20                | 1.25 (0.44)       | 1.02 |
| I felt safe responding.                          | Paper               | 20                | 1.70 (0.73)        | 20                | 1.63 (0.76)       | 0.29 |
|                                                  | Smartphone          | 20                | 1.35 (0.59)        | 20                | 1.55 (1.05)       | –0.74 |
| I could answer fast.                             | Paper               | 20                | 1.35 (0.59)        | 20                | 1.65 (1.04)       | –1.12 |
|                                                  | Smartphone          | 20                | 1.20 (0.41)        | 20                | 1.15 (0.49)       | 0.35 |
| It was comfortable to use.                       | Paper               | 20                | 3.47 (0.70)        | 20                | 3.30 (0.80)       | 0.72 |
|                                                  | Smartphone          | 20                | 3.40 (0.82)        | 20                | 3.50 (0.83)       | –0.38 |
| It hindered the daily activities.                | Paper               | 20                | 4.21 (1.03)        | 20                | 4.45 (0.76)       | –0.83 |
|                                                  | Smartphone          | 20                | 4.45 (0.76)        | 20                | 4.60 (0.94)       | –0.56 |
| The fact that people could see me using it annoyed me. | Paper               | 20                | 4.63 (0.68)        | 20                | 4.15 (1.14)       | 1.59 |
|                                                  | Smartphone          | 20                | 4.75 (0.44)        | 20                | 4.15 (1.14)       | 2.20* |
| It can be used in any context.                   | Paper               | 20                | 2.00 (1.11)        | 20                | 1.95 (1.27)       | 0.14 |
|                                                  | Smartphone          | 20                | 1.55 (1.00)        | 20                | 1.95 (1.10)       | –1.21 |
| Instructions were easy.                          | Paper               | 20                | 1.42 (0.51)        | 20                | 1.25 (0.44)       | 1.12 |
|                                                  | Smartphone          | 20                | 1.20 (0.41)        | 20                | 1.10 (0.31)       | 0.87 |
| The font size was sufficient.                    | Paper               | 20                | 1.39 (0.61)        | 20                | 1.25 (0.44)       | 0.81 |
|                                                  | Smartphone          | 20                | 1.20 (0.41)        | 20                | 1.15 (0.37)       | 0.41 |
| In general, using the self-record was easy.     | Paper               | 20                | 2.00 (0.86)        | 20                | 1.80 (0.70)       | 0.81 |
|                                                  | Smartphone          | 20                | 1.25 (0.44)        | 20                | 1.30 (0.57)       | –0.31 |
| In general, using the self-record was useful.   | Paper               | 20                | 2.42 (1.26)        | 20                | 2.40 (1.05)       | 0.06 |
|                                                  | Smartphone          | 20                | 1.90 (0.91)        | 20                | 2.20 (0.83)       | –1.09 |
| People like me should use the self-record.      | Paper               | 20                | 2.55 (1.23)        | 20                | 2.50 (1.24)       | 0.13 |
|                                                  | Smartphone          | 20                | 1.95 (0.89)        | 20                | 2.25 (1.02)       | –0.99 |

FT, some familiarity with technology; LFT, low familiarity with technology; M, mean; SD, standard deviation.

* p < 0.05

** p < 0.01.

*Each item was rated from 1 to 5, where 1 was ‘totally agree’ and 5 ‘totally disagree’.
Comparing EMA using smartphone versus paper

Table 5 Frequency and percentage for preferences between paper and smartphone: total sample and participant familiar versus less familiar with technology.

|                                | Paper | Indistinct | Smartphone |
|--------------------------------|-------|------------|------------|
| I prefer using                 | Total | 6          | 8          | 26         | 65 |
|                                | LFT   | 3          | 5          | 12         | 60 |
|                                | FT    | 3          | 3          | 14         | 70 |
| It is easier to use            | Total | 4          | 10         | 15         | 37.5 |
|                                | LFT   | 4          | 20         | 8          | 40 |
|                                | FT    | 0          | 0          | 13         | 35 |
| It is most useful              | Total | 2          | 5          | 18         | 45 |
|                                | LFT   | 1          | 5          | 8          | 40 |
|                                | FT    | 1          | 5          | 10         | 50 |
| It is easier to remember that I have to fill it. | Total | 1          | 2.5        | 2          | 5 |
|                                | LFT   | 0          | 0          | 1          | 5 |
|                                | FT    | 1          | 5          | 1          | 5 |
| It is more comfortable to fill | Total | 1          | 2.5        | 16         | 40 |
|                                | LFT   | 0          | 0          | 6          | 30 |
|                                | FT    | 1          | 5          | 10         | 50 |
| It is more comfortable to carry with me. | Total | 1          | 2.5        | 5          | 12.5 |
|                                | LFT   | 0          | 0          | 2          | 10 |
|                                | FT    | 1          | 5          | 3          | 15 |
| It is easier to learn how to handle. | Total | 1          | 2.5        | 24         | 60 |
|                                | LFT   | 1          | 5          | 10         | 50 |
|                                | FT    | 0          | 0          | 14         | 70 |
| It is faster.                  | Total | 7          | 17.5       | 5          | 12.5 |
|                                | LFT   | 2          | 10         | 3          | 15 |
|                                | FT    | 5          | 25         | 2          | 10 |
| It bothers me more            | Total | 18         | 45         | 18         | 45 |
|                                | LFT   | 10         | 50         | 9          | 45 |
|                                | FT    | 8          | 40         | 9          | 45 |

FT, some familiarity with technology; LFT, low familiarity with technology.

A second objective was to explore the relationship between aggregated EMA data and retrospective data. We found positive and significant correlations between them, indicating that both procedures were measuring a similar variable. However, retrospective reports of pain and fatigue were higher than aggregated EMA data. Our results are in line with those found in the RTDC literature reviewed by Stone and Broderick (2007) who found a correlation between recalled pain and average momentary pain for a 1-week period of around 0.75 and higher pain retrospective reports when compared with the average of RTDC data for the same period. In our case we explored not only pain but also another important pain-related symptom, fatigue. However, regarding the correlations, our data showed significant correlations between aggregated EMA data and retrospective data but we would like to notice that they were lower than the ones found by Stone and Broderick, and when comparing the correlations between paper and retrospective data (0.59 for pain and 0.67 for fatigue) and smartphone and retrospective data (0.39 for pain and 0.47 for fatigue), the correlations regarding the smartphone were lower. This could be due to the fact that paper diaries and the weekly reports used shared a common format, paper and pencil, and the smartphone included different features like the use of sound or touch screen (without the need of discovering the actual compliance with the paper diary by using a mechanism that could detect when the diary was opened. They found that while reported compliance was 90%, actual compliance was only 11%. Our study was not so focused on compliance, we were more interested in exploring the utility of a smartphone for RTDC and because of that we compared it to one of the most common procedures in regular practice in our context: a paper diary. It is important to notice that, even considering reported compliance, it was significantly lower with the paper diary. Our study contributes to the literature indicating that even in conditions where it was possible to backfill and forward-fill the paper diary, the smartphone diary obtained a better compliance. We believe this finding strengthens the merits of the use of smartphones versus paper diaries.
of writing). This result may open research questions regarding the role of specific features of RTDC in the assessment procedure.

Regarding the overestimation of pain and fatigue, our results confirm that retrospective assessment tends to produce higher estimations of events (Houtveen and Oei, 2007; Broderick et al., 2008). This could be due to the rules of ‘peak’ and ‘closest’ pain, whereby more weight is given to the peak of pain experienced and to the most recently experienced pain (Redelmeier and Kahneman, 1996). We believe this has important implications, meaning that patients remember feeling more pain than they actually felt when looking at the daily ratings. Thus, there is a distorted view of the intensity of the pain. On one hand, this supports the role of cognitive factors in the experience of pain such as catastrophizing (Keefe et al., 2004). On the other hand, the comparison between real-time data and recall data could be used as a therapeutic tool to promote decatastrophizing.

The last objective was to explore the acceptability and preferences regarding the two assessment methods. Several studies indicate that e-diaries have higher acceptability than traditional assessment methods (Cranford et al., 2006; Wilhelm and Schoebi 2007). Other studies report that traditional diaries present the advantage of being more familiar and maybe easier to use (Stone et al., 2003). Another variable to take into account is the specific population to be assessed. In our case, we know that FMS is more prevalent in people between 40 and 60 years old (Baldry, 2001) and some studies in our context, and also in other populations, indicate that patients with FMS are related to lower educational and socioeconomic levels (White et al., 1999; Mas et al., 2008). These features are usually related with less familiarity with technology. Because of that, we consider it important to explore the acceptability of the smartphone diary in this specific population. The results obtained show that the smartphone diary was rated as easier, more useful and more highly recommended than the paper diary. Regarding preferences, most participants preferred the smartphone over the paper diary. Moreover, when comparing participants familiar versus less familiar with technology, acceptability was very similar. Participants with more familiarity with technology only reported higher acceptability of the smartphone regarding ease of use. Participants less familiar with technology reported higher acceptability only regarding the item related with annoyance to carry the smartphone. Also, participants with less familiarity still preferred the smartphone over the paper diary. These results could be influenced by the fact that a lot of care was put into the design of the application, being sensitive to the characteristics of the target population. In summary, the smartphone presented a high acceptability, even in a sample with an important proportion of participants with low familiarity with technology. The evolution of technology is one of the biggest achievements of our recent history; however, it is important to make this technology available to everybody who can benefit from it. Specific populations could be at risk of not benefiting from technological innovations because of the digital divide, and we believe researchers and clinicians have a responsibility to offer technological tools that have been submitted to a process of careful evaluation in order to reach most people who could benefit from them. This has been our goal in the development of the tool we present in this study.

This study presents some limitations. We already indicated that we could not report actual compliance with the paper diary. Another limitation is the small sample size. The strength of the sample was that it was representative of the population that suffers from FMS in our country (referred by the rheumatology unit of a public hospital). A third limitation is that the comparisons between aggregated EMA data and recall data for mood were not presented because the weekly measures used (BPI and BFI) did not include a measure of mood that could be compared with the daily measure (face-based pictorial scale). A final limitation is that the application runs only on Windows Mobile software, leaving out other important mobile platforms. We are now developing an updated version that could run on the main platforms.

This study belongs to a promising line of research that is providing interesting data for the field of pain. Another future issue in this line is the use of these procedures in longitudinal studies to analyse the relationship between different variables and individual differences in the experience of pain. Finally, fast advances in technology will reduce costs and improve the feasibility of using patients’ mobile phones for EMA and for giving therapeutic feedback during the administration of an intervention.

The findings of this study contribute with data supporting the use of smartphones for RTDC in the field of chronic pain. This is a very useful tool for clinicians and researchers because it makes it possible to gather more complete ratings of relevant variables. Besides the good results regarding data collection, this method was accepted well by a sample of patients with FMS referred by a public hospital with an important proportion of participants with low familiarity with technology.
Comparing EMA using smartphone versus paper

Author contributions

All authors have contributed to the manuscript. A.G.-P. and C.B. contributed to the conception and design of the study. D.C. and J.G. conducted the design and development of the computer software. The statistical analysis and the interpretation of data were conducted by R.H. and A.G.-P. An initial draft of the manuscript was written by R.H. and A.G.-P. The manuscript was written by R.H. and A.G.-P. under the guidance of A.G.-P. R.M.B. and M.A.B. provided helpful advice revising it critically. All authors contributed to and approved the final draft.

References

Badia, X., Murriel, C., Gracia, A., Nunez-Olarte, J.M., Perulero, N., Galvez, R., Carulla, J., Cleeland, C.S. (2003). Validation of the Spanish version of the Brief Pain Inventory in patients with oncological pain [in Spanish]. Med Clin (Barc) 120, 52–59.

Baldry, P.E. (2001). Clinical characteristics and biopathophysiological mechanisms of fibromyalgia syndrome. In Myofascial Pain and Fibromyalgia Syndromes, P.E. Baldry, ed. (Edinburgh: Churchill Livingstone) pp. 351–377.

Bolten, W., Emmerich, M., Weber, E., Fassmeyer, N. (1991). Validation of electronic by conventional pain diaries. Z Rheumatol 1, 55–64.

Broderick, J.E., Schwartz, J.E., Vikingstad, G., Pribbernow, M., Grossman, S., Stone, A.A. (2008). The accuracy of pain and fatigue items across different reporting periods. Pain 139, 146–157.

Castilla, D., Botella, C., Garcia-Palacios, A., Farfallini, L., Miralles, I. (2012). Experience-sampling methodology with a mobile device in fibromyalgia. Int J Telemed Appl 2012, 162673. doi:10.1155/2012/162673.

Cleeland, C.S., Ryan, K.M. (1994). Pain assessment: Global use of the Brief Pain Inventory. Ann Acad Med Singapore 23, 129–138.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, 2nd edn (Lawrence Erlbaum Associates: Hillsdale).

Cranford, N.J., Shrout, P., Iida, M., Rafaeli, E., Yip, T., Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliability. Pers Soc Psychol Bull 32, 917–929.

Etchemendy, E. (2011). Butler System: Positive Technology for the Elderly (Spain: Doctoral Dissertation (Universitat Jaume I).

Fietta, P., Fietta, P., Manganelli, P. (2007). Fibromyalgia and psychiatric disorders. Acta Biomed 78, 88–95.

Flor, H., Turk, D.C. (2011). Assessment of characteristics of pain and pain behaviors: Laboratory and clinical methods. In Chronic Pain. An Integrated Biobehavioral Approach, F.I. Flor, D.C. Turk, eds. (Seattle: IASP Press) pp. 217–252.

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

Gвалтней, C.J., Shields, A.L., Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. Val Health 11, 322–333.

Houtveen, J.H., Oei, N.Y.L. (2007). Recall bias in reporting medically unexplained symptoms comes from semantic memory. J Psychosom Res 62, 277–282.

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

Keefe, F.J., Rumble, M.E., Scripio, C.D., Giodano, I.A., Perri, L.M. (2004). Psychological aspects of persistent pain: Current state of the science. J Pain 5, 195–211.

Marceau, L.D., Link, C., Jamison, R.N., Carolan, S. (2007). Electronic diaries as a tool to improve pain management: Is there any evidence? Pain Med 8, 101–109.

Mas, A.J., Carmona, L., Valverde, M., Rivas, B., EPISER Study group. (2008). Prevalence and impact of fibromyalgia on function and quality of life in individuals from the general population: Results from a nationwide study in Spain. Clin Exp Rheumatol 26, 519–526.

Mendoza, T.R., Wang, X.S., Cleeland, C.S., Morrissey, M., Johnson, B.A., Wendt, J.K., Huber, S.L. (1999). The rapid assessment of fatigue severity in cancer patients: Use of the Brief Fatigue Inventory. Cancer 85, 1186–1196.

Palermo, T.M., Stork, P.P., Valenzuela, D. (2003). A randomized controlled trial of electronic versus paper pain diaries in children: Impact on compliance, accuracy, and acceptability. Pain 107, 213–219.

Redelmeier, D., Kahneman, D. (1996). Patients’ memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. Pain 66, 3–8.

Shiffman, S., Stone, A.A., Hufford, M. (2008). Ecological momentary assessment. Ann Rev Clin Psychol 4, 1–32.

Stone, A.A., Broderick, J.E. (2007). Real-time data collection for pain: Appraisal and current status. Pain Med 8, 85–93.

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

White, K.P., Speechley, M., Harth, M., Ostbye, T. (1999). The London fibromyalgia epidemiology study: The prevalence of fibromyalgia syndrome in London, Ontario. J Rheumatol 26, 1570–1576.

Wilhelm, P., Schoebi, D. (2007). Assessing mood in daily life: Structural validity, sensitivity to change, and reliability of a short-scale to measure three basic dimensions of mood. Eur J Psychol Assess 23, 258–267.

Wolfe, F., Smythe, H.A., Yunus, M.B., Bennett, R.M., Bombardier, C., Goldenberg, D.L. (1990). The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia: Report of the Multicenter Criteria Committee. Arthritis Rheum 33, 160–172.