



HHS Public Access

Author manuscript

Contemp Clin Trials. Author manuscript; available in PMC 2018 May 04.

Published in final edited form as:

Contemp Clin Trials. 2010 September ; 31(5): 483–490. doi:10.1016/j.cct.2010.06.004.

Validity of Average, Minimum, and Maximum End-of-Day Recall Assessments of Pain and Fatigue

Arthur A. Stone, Ph.D.¹, Joan E. Broderick, Ph.D.¹, and Joseph E. Schwartz, Ph.D.¹

¹Department of Psychiatry and Behavioral Science, Stony Brook University, Stony Brook, New York, USA

Abstract

End-of-day (EOD) diary assessments of symptoms have the potential to reduce recall bias associated with longer recall periods, and therefore be useful for generating accurate patient reported outcomes (PROs). In this report we examine the relative validity of diary questions about the experience of daily pain and fatigue, including several questions about experience for the entire day and questions about minimum and maximum daily levels, with previously collected data. Validity estimates are based on comparisons of EOD reports with momentary recordings of pain and fatigue from the same days. One hundred and six participants with rheumatologic diseases yielded 2,852 days for analysis. Differences in levels as assessed by EOD and momentary reports were small (just a few points), although in many instances were significantly different. Correlational analyses indicated that “how much,” “how intense,” and “on average” EOD questions were more strongly associated with momentary reports ($r_s = .85-.90$ for pain and $.81-.83$ for fatigue) than were minimum and maximum questions ($r_s = .73-.80$ for pain and $.67-.75$ for fatigue). Overall, the pain measures had higher EOD-momentary correspondence than the fatigue measures. Analyses of difference scores between EOD and momentary reports confirmed the better correspondence of the average questions compared with minimum and maximum questions. There was little evidence of individual differences in level and correspondence analyses. The implication of these results is that over-the-day diary measures may yield superior PROs than those based on minimum or maximum daily levels.

Corresponding author: Arthur A. Stone, Ph.D., Department of Psychiatry and Behavioral Science, Putnam Hall, Stony Brook University, Stony Brook, NY 11794-8790 USA, Tel.: 001-631-632-8833, Fax: 001-631-632-3165, Arthur.stone@sunysb.edu.

¹Means and standard deviations in this table are slightly different than those found in our earlier paper, because that paper used only one of the four weeks of data used in these analyses.

²An alternative method of examining correspondence is to compute within-subject pooled correlations, which center each individual's scores around their own mean, eliminating any between-person effects from the correlations. Such correlations were slightly lower than those presented in Tables 2 and 4, but the pattern of the correlations was the same as the raw correlations, therefore, only the raw correlations are presented.

³Theoretically, the likelihood of actually capturing the least or worst pain during the day should increase as a function of the number of assessments per day, with more assessments resulting in more accurate (less error) estimates of minimum and maximum.

⁴We thank an astute reviewer for suggesting this exploration of individual differences.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Patient Reported Outcomes (PROs) are patients' self-reports of their symptoms, the impacts of their symptoms, and their behaviors. PROs have received considerable attention because they provide a unique perspective on patients' health and functioning . One problem with self-report measures is the length of the recall period , , the amount of time to be considered when completing an assessment. Long recall periods may stretch the ability of respondents to accurately recall and summarize information, leading to concerns about accuracy of reports .

By limiting the duration of recall period, daily diaries should reduce recall bias compared to assessments with longer recall periods (e.g., weeks), and they can be aggregated over days to cover reporting periods typically used by retrospective assessments . Diary questions often ask about the entire day's symptoms, but can also include questions about the day's least or lowest level of a symptom or the day's worst or maximum level. Recently these alternatives were explicitly suggested in the FDA's PRO Guidance document . There are two reasons why least/worst levels may be appealing candidates for assessment: 1) they avoid the potentially difficult cognitive process of summarizing experience and 2) least/worst may be the construct of interest as opposed to average experience over the day. For example, one can hypothesize that treatments could reduce maximum pain during a day, yet have only a modest impact on average levels. It is also notable that some weekly recall questionnaires for pain assessment ask about least and worst levels (e.g., BPI;), indicating interest in these constructs.

Some validity data are available for EOD diaries with ratings of the over-the-day experience, and those results are encouraging. One study of post-surgical patients compared EOD recall of daily pain with the average, peak, and last-of-day variables based on 5 randomly selected momentary assessments . EOD recall of pain intensity correlated about .70 with the average of momentary reports and only 4% recall bias from peak and end pain was found. Second, previous results from a subset of the current dataset showed good correspondence between EOD dairies and momentary reports for pain and fatigue measures : correlations ranged from .75 to .85.

To our knowledge, this is the first report to compare the validity of EOD recalled over-the-day, least, and worst pain and fatigue diary questions with multiple momentary assessments from the same day. For EOD questions of average and “how much” pain/fatigue, we use the average of moments for the same day as the validity criterion; for the EOD measures of least pain/fatigue, we use the minimum value of the day's momentary reports; and, for EOD measures of worst pain/fatigue, we use the maximum value of the day's momentary reports. Evidence for the validity of EOD measures would be 1) that their levels are similar to the corresponding momentary measure and 2) that the correspondence over days between EOD and the momentary measure was high.

Methods

Participants

Patients were recruited from two offices of a community rheumatology practice. Participants were required to be available for 30 consecutive days and to meet the following eligibility

criteria: 18 years of age; physician-confirmed diagnosis of a chronic rheumatological illness; experienced symptoms of pain or fatigue during the last week; no significant sight, hearing, or writing impairment; fluency in English; normal sleep-wake schedule; ability to come to the research office twice within a month; had not participated in another electronic diary study in the last 5 years. A total of 279 patients were telephone screened, and 86 (31%) were excluded due to one or more of the above eligibility criteria. Of the 193 eligible patients, 76 (39%) declined participation, and 117 (61%) participated. We examined the demographic characteristics of those who were eligible and participated versus those who were eligible and declined participation. Age, sex, educational achievement, marital status, race, and reported pain and fatigue at screening were examined by participation status. A near-significant difference was found for age where those who participated (56.3 years) were older than those who declined participation (52.8 years; $t(191)=1.94$, $p=.053$); none of the other comparisons were significant. Over the course of the study eleven participants dropped out, and 106 completed the study. The final sample was middle-aged ($X=55.5$ years), predominantly female (91%), white (92%), married (65%), and well-educated (63% had at least some college).

Procedure

The study protocol was approved by the Stony Brook University Institutional Review Board. Participants provided informed consent and were compensated \$100. Data were collected from September 2005 through June 2006. Eligible patients came to the research office to complete demographic and questionnaire measures and to be trained in the use of an electronic diary (ED). Momentary and daily recall ratings of pain and fatigue intensity were collected for 29-31 days on a hand-held computer (Palm Zire 31). The ED utilized a software program provided by invivodata, inc. (Pittsburgh, PA) that featured auditory tones to signal the participant to complete a set of momentary ratings. It was programmed to generate an average of 7 randomly-scheduled (within intervals) prompts spread across the participant's waking hours (an average of one every 2 hours and 20 minutes, constrained to ensure a minimum of 30 minutes between prompts) determined by when the participant informed the ED that she was going to bed at night and set the wake up alarm the next morning. In addition to the random signals, the ED prompted the participant to complete a daily recall assessment at the time the ED was put to sleep at night, the "End of Day" assessment. A research assistant telephoned the patient 24 hours after the initial research office visit to answer any questions and troubleshoot potential problems with using the ED. A follow-up call was made once per week for the following three weeks to ensure the ED was working properly and to answer any questions. At the end of the month, patients returned the ED to the research office.

Measures

Items for this study were drawn from the Brief Pain Inventory (BPI) and the Brief Fatigue Inventory (BFI), with wordings modified to correspond to the different reporting periods. Zero to 100-point Visual Analog Scales were used, but scale endpoints varied according to question content. For the "how much" bodily pain question the anchors were "none" (0) and "very severe" (100), whereas for all other questions the anchors were "not at all" (0) and "extremely" (100). The EOD questionnaire contained several questions that were used to

address the aims of this paper. Three asked about over-the-day levels of pain: How much bodily pain did you have?, How intense was your bodily pain?, and What was your average level of pain today? Another two questions asked about the lowest (What was the lowest level of your pain today?) and highest (What was the worst level of your pain today?) levels of pain for the day. A parallel set of questions was available for the construct of fatigue/tiredness: How fatigued (weary, tired) did you feel? and How tired did you feel? There were also questions about the lowest (What was the lowest level of your fatigue today?) and highest (What was the worst level of your fatigue today?) levels of fatigue for the day. Each of these EOD questions began with the stem “DURING THE DAY.” These questions were also asked on a momentary basis. Each of these included the stem “BEFORE PROMPT.” From each of these four momentary questions, the average, the minimum, and the maximum were derived.

Our strategy for determining the validity of the EOD assessments is to compare data collected at several random points during each day with the EOD assessments. Momentary reports are thought to be relatively free from distortion due to recall and, because they are sampled at random points from the day they provide an unbiased view of average daily pain and fatigue. Two ways of comparing EOD and momentary reports are computed. *Level* differences are defined as the average difference between EOD and momentary data from the same day; *correspondence* differences are defined as the covariation between EOD and momentary data over days. Level and correspondence are both relevant to understanding the validity of recall. These analyses were conducted with data collected in a diary study that examined several recall periods (1-, 3-, 7- and 28-days). We have previously reported on the level and correspondence of EOD average pain/fatigue (based on only 1 week of data per participant), and those analyses did not examine EOD reports about worst and least pain/fatigue and did not use the full 4 weeks of data.

Results

For the analyses to yield good estimates of level differences and correspondence, there had to be an adequate number of momentary assessments each day. The design of the study specified 7 momentary samples per day, although in practice this number varied from day to day. It could be greater than 7 if a person was awake for more than 14 hours, or it could be less than 7 if compliance for the day was poor. To balance the goals of including as many days in the analysis as possible, yet keeping a reasonable number of assessments, we decided that at least 4 assessments per day would be required. This yielded a sample of 106 participants with 2,852 days of data (with an average of 5.6 reports per day and between 8 and 34 days in the study). In secondary analyses described below, we tested the possibility that our estimates of EOD and momentary differences were affected by the number of assessments per day by conducting the analyses with participants who had 4-5 momentary assessments per day (1406) and those with 6 or more momentary assessments per day (1446).

For pain, there were 3 EOD variables that were intended to capture the overall experience of the day: how much bodily pain, how intense was bodily pain, and what was the average level of pain intensity. For EOD comparisons with momentary reports, which had 2 questions

(how much pain and how intense was the pain), we compared 1) EOD “how much” pain with the average of momentary “how much” pain, 2) EOD “how intense” was pain with the average of momentary “how intense” was pain, and 3) EOD “what was your average level of pain” with the average of momentary “how much” pain (see Table 1). For the EOD lowest and worst, we examined both momentary variables, how much pain and pain intensity. The same strategy for comparing EOD to momentary variables was employed for the fatigue comparisons.

Level Differences

These analyses are intended to compare the level of pain/fatigue ratings from EOD measures with the level of the corresponding momentary measures; for instance, the level of EOD average pain with the average of momentary pain assessments, or the level of EOD ratings of least pain with the minimum of the momentary ratings of pain. Results of these analyses are shown in Table 1, where the EOD questions are presented on the left side of the table and the momentary variables to the right. Statistical testing of level differences was done with multilevel modeling in order to control for differences in the number of momentary reports per participant and to model the nesting of days within person so as not to inflate the degrees of freedom used for statistical testing.

EOD reports of average pain and fatigue over the day are rated significantly higher than the average of momentary reports of pain and fatigue. On the other hand, EOD reports of least pain and fatigue are significantly *lower* than the momentary minimums and EOD ratings of worst pain are *higher* than the momentary maximums. The exception is that EOD ratings of worst fatigue were *lower* than the momentary maximums. Across the comparisons, the level differences (averaged across days and people) ranged from 1.6 to 5.4 points on a 101-point scale.

Correspondence Differences

These analyses compare the correspondence or correlation between EOD and momentary measurements. To estimate the association between EOD and momentary scores, Pearson correlation coefficients were computed. It is plausible that there could be error, especially for the “least” and “most” measures, in cases where there are a limited number of daily assessments. To test this, the correlations for days that had 4 or 5 momentary assessments (average= 4.6) are compared to those with 6 or more momentary assessments per day (average= 6.6). There were only small differences in the values between the full sample and the selected samples, thus indicating no clear advantage for the sample with 6 or more daily assessments.

The second column of Table 2 presents the correlations with all days, the third column the correlations for days with 4-5 moments, and the fourth column for days with 6 or more moments. For pain, the range of correlations for EOD over-the-day measures with their momentary counterparts is .85 to .90; for the EOD minimum with its momentary counterparts the range is .73-.74; and, for EOD maximum with its momentary counterparts the range is .78-.80. For fatigue, the range for EOD over-the-day measures with their

momentary counterparts is .81-.83; for the EOD minimum the range is .67-.68; and, the EOD maximum it is .72-.75. All of these correlations are statistically significant.

Discrepancies between EOD and Momentary Measures—What are *a priori* acceptable levels of difference for EOD reports versus their momentary counterparts? Because there are no standards for guiding the answer to this question, we evaluate two levels of error for the reader to consider. Accepting a seemingly small degree of error, we chose a difference score acceptability range of plus or minus 10 points on the 101-point VAS scale, and for the wider threshold we used plus or minus 20 points. The difference was computed by subtracting the momentary values from the EOD value. For example, to compute the proportion of 10-point discrepancies for “Worst” pain versus the maximum momentary value for the day, a difference was said to exist if the absolute difference of (“Worst” Pain – Maximum Momentary Pain) was greater than or equal to 10. The percentages of days that had these levels of difference are shown in Table 2. There is more error for EOD “least” and “worst” pain and fatigue relative to error rates for the EOD average and usual scores. Yet another way to examine this data is by histograms of the differences between the EOD variables and momentary variables, as shown in Figure 1 for EOD “average” pain (upper panel) and for “how” fatigued (lower panel) (other pain measures and fatigue measures not shown). A much narrower distribution of errors for the over-the-day variables compared with minimum and maximum variables is evident.

Individual Differences

It is plausible that there are individual differences in level and correspondence between EOD variables and their corresponding momentary-based variables. Given the potentially large number of comparisons and the secondary nature of these analyses, we examined individual differences in a subset of the data. One pain variable (Pain Intensity) and one fatigue variable (Fatigue Intensity) were chosen to represent the two content domains. The individual difference variables selected were also a subset of all possible variables; they were age (less than 57 years [n=52] vs. greater than or equal to 57 years [n=54]), educational level (Some college or less [n=60] vs. College graduate or more [n=45]), overall health status based on SF-36 global health question (Poor or Fair health [n=38] vs. Good, Very Good or Excellent health [n=68]), average level of EMA pain (for the pain variables only, Low [n=53] vs. High [n=53], based on median split), and average level of EMA fatigue (for the fatigue variables only, Low [n=53] vs. High [n=53], based on median split). Gender was not considered because there were so few men in the sample.

To examine individual differences in level, new person-level variables were computed to represent the difference between EOD and momentary variables; this was done for daily average, daily minimum, and daily maximum for both pain and fatigue intensity, yielding 6 variables. To test individual differences in level, a comparison of the mean difference scores between EOD and EMA between the two groups defined by each individual attribute was computed. To examine individual differences in correspondence, the correlation between EOD variables and their corresponding momentary-based variables were computed separately for each group and the difference tested; for example, the correlations between EOD pain and momentary pain were compared for older and younger persons (the

individual difference variable). To be consistent with the set of correspondence analyses shown earlier, the unit of analysis for these comparisons was a person-day.

Individual Differences in Level—Table 3 presents the EOD—EMA difference scores for each level of the individual difference variables (column headings) with significance level of t-tests. A total of 24 t-tests were computed and 4 reached a significance level of .01 or greater. All significant differences at this alpha level were found for the test of EOD least pain or fatigue versus EMA minimum pain or fatigue. However, these tests were spread across the five individual difference variables, suggesting that none of the individual difference variables was consistently associated with level differences.

Individual Differences in Correspondence—Table 4 presents correlations between EOD and EMA measures separately for the mean, minimum, and maximum variables for each subgroup defined by the individual difference variables. We examined the table for major differences in correlations for the low and high level of the individual difference variables; no difference or a small difference would indicate no evidence for individual differences whereas major differences would indicate individual differences. There were 24 comparisons and most differences between correlations were .05 or smaller. The largest difference was .12 for the Age variable with the Minimum variables and only a total of three correlation comparisons were discrepant by .10 or more. This suggests that individual differences play only a minor role in EOD—EMA correspondence.

Discussion

The goal of this report was to provide empirical evidence pertaining to the validity of end-of-day reports of pain and fatigue (over-the-day, least, and worst) that are used by outcomes researchers. EOD diaries are an expedient means of assessing overall or average levels of an outcome while trying to minimize recall bias. The researcher has the option of averaging as many daily assessments as needed for their research purposes (e.g., over 7 days to characterize the outcome for a week). The FDA's PRO Guidance encourages the use of diaries and brief recall periods, and it mentions using diaries to collect information about worst and least daily outcomes. This paper reports the first comparison of the validity of EOD “worst” and “least” assessments based on comparisons with momentary reports from the same day.

Over all comparisons for both pain and fatigue, EOD *levels* were within a few points of their momentary counterparts. The largest mean differences were about 5 points; and although most were statistically different, they will probably be viewed as minor from a clinical point of view. These discrepancies are less than we have found when comparing 7-day recall of average pain and fatigue to aggregated momentary reports, where the mean discrepancies were as large as 15 points. Thus, on average, EOD reports closely reflect the average, least, and worst levels of pain and fatigue as measured by momentary reports for a reporting period of a single day. Importantly, they appear to introduce less recall bias than recall ratings using a 7-day reporting period. However, on an individual-day basis, we observed many discrepancies of 10, 20, and more points.

Results for correspondence yielded distinct patterns between the over-the-day versus “least” and “worst” measures; there was also an overall pattern observed for pain measures versus fatigue measures. Correspondence with their momentary counterparts were higher for the EOD over-the-day measures (“How much,” “How intense,” and “Average” questions) than for the “least” and “worst” measures: from 10% to 20% more variance was shared between EOD and momentary variables in the over-the-day versus the least or worst measures. We view these as significant differences with the results suggesting that EOD ratings of “How much” or “How intense” or “On average” for daily pain or fatigue are more valid EOD measures than EOD ratings of “least” or “worst” pain or fatigue. This result was confirmed with the analyses of discrepancy scores that were based on two thresholds for defining differences between EOD and momentary measures. Considerably lower rates of “large” discrepancies were found for EOD over-the-day measures compared with least and worst EOD measures. Another general observation was that pain measures were more closely associated with momentary assessments than were fatigue measures, which was also the case in our previous comparison of weekly recall to average momentary experience. These correlations are higher than we have found when comparing 7-day recall of average pain and fatigue and aggregated momentary reports, which is likely the result of the shorter recall period in this study. We do not have a satisfactory explanation for this finding, but it may be that pain is a relatively more distinct state than fatigue and, hence, is recalled more accurately even within a period as short as a day.

There was the possibility that there were individual differences in discrepancies in end-of-day versus momentary measures of mean, minimum, and maximum pain and fatigue. We examined this possibility in a subset of the pain and fatigue measures and with a set of five factors that could moderate level or correspondence differences. Only a few significant effects were observed for the individual differences variables and they were spread over the individual differences variables, suggesting either none or only a minor effect of the variables. Of course, variables that were not examined could have individual difference effects, but this seems less likely given the current findings.

There are several limitations that must be considered in the interpretation of these results. One is that the comparisons of EOD minimums and maximums with the momentary assessments were based on a limited number of momentary assessments. It is possible that the momentary assessments missed the highest and lowest symptom levels, thus not capturing the true minimum and maximum levels. Two aspects of the results suggest this was not a major factor in the analysis. The first is that the means of *recalled* maximums or minimums were not very different from their *momentary* assessment counterparts, though they were mostly in the predicted direction. If momentary assessments had regularly missed “true” maximums or minimums, then we would have expected a larger difference between recall and momentary averages, with much larger recalled maximums and much lower recalled minimums. But this was not the case. One possible explanation is that periods of low or high pain/fatigue typically last for several hours and not just for brief intervals, thus being captured by at least one momentary assessment. The second finding that suggests that biasing due to daily assessments was not large were the analyses showing that increasing the number of daily momentary assessments from 4-5 per day to 6 or more per day had very little impact on the results, reducing but not eliminating our concern about the coverage of

momentary assessments. Finally, there is also the possibility that our estimates of level and correspondence differences are biased upward because participants' EOD evaluations may have been more accurate due to the momentary recording they did throughout the day. That is, the EMA component of the design may have enhanced their memories of pain and fatigue at the end of the day.

If the intention of a study is to detect changes in least pain or worst pain experienced during the day, then it would make little sense to use the more valid over-the-day measure since they assess a different construct. However, if there was flexibility in the choice of outcome measures for a trial or if it was unclear which outcomes would be affected by treatment, then we believe the results reported here indicate that over-the-day measures have better psychometric properties than measures of least and worst pain and fatigue.

In summary, this study suggests that end-of-day reports of over-the-day pain and fatigue were strongly associated with momentary assessments of the same, justifying their use as PROs. We are less sanguine about the use of “Worst” and “Least” given the weaker associations with momentary assessments, but they did have substantial correlations with momentary reports, which may also justify their use.

Acknowledgments

This work was supported by a grant from the National Institutes of Health (1 U01-AR052170-01; Arthur A. Stone, principal investigator) and by GCRC Grant no. M01-RR10710 from the National Center for Research Resources. A.A.S. is a Senior Consultant for invivodata, inc., and a Senior Scientist with the Gallup Organization.

References

1. Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain*. Sep 30; 2008 139(1):146–157. [PubMed: 18455312]
2. Burke LB, Kennedy DL, Miskala PH, Papadopoulos EJ, Trentacosti AM. The use of patient-reported outcome measures in the evaluation of medical products for regulatory approval. *Clinical Pharmacology Therapeutics*. 2008; 84:281–283.
3. Bradburn NM, Rips LJ, Shevell SK. Answering autobiographical questions: The impact of memory and inference on surveys. *Science*. 1987; 236:151–167.
4. Gorin, AA., Stone, AA. Recall biases and cognitive errors in retrospective self-reports: A call for momentary assessments. In: Baum, A., Revenson, T., Singer, J., editors. *Handbook of Health Psychology*. Mahwah, N.J.: Erlbaum; 2001. p. 405-414.
5. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. *Annu Rev Psychol*. 2003; 54:579–616. [PubMed: 12499517]
6. Guidance for Industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009 <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>
7. Cleeland C. Pain assessment: global use of the Brief Pain Inventory. *Annals of Academic Medicine Singapore*. 1994; 23:129–138.
8. Jensen MP, Mardekian J, Lakshminarayanan M, Boye ME. Validity of 24-h recall ratings of pain severity: Biasing effects of “Peak” and “End” pain. *Pain*. 2008; 137:422–427. [PubMed: 18035495]
9. Daut RL, Cleeland CS. The prevalence and severity of pain in cancer. *Cancer*. Nov 1; 1982 50(9): 1913–1918. [PubMed: 7116316]

10. Mendoza TR, Wang XS, Cleeland CS, et al. The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. *Cancer*. Mar 1; 1999 85(5):1186–1196. [PubMed: 10091805]
11. Stone, AA. *The science of real-time data capture : self-reports in health research* . Oxford; New York: Oxford University Press; 2007.
12. Stone AA, Broderick JE, Schwartz JE, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. *Pain*. Jul; 2003 104(1-2):343–351. [PubMed: 12855344]
13. Schwartz JE, Stone AA. Data analysis for EMA studies. *Health Psychology*. 1998; 17:6–16. [PubMed: 9459065]

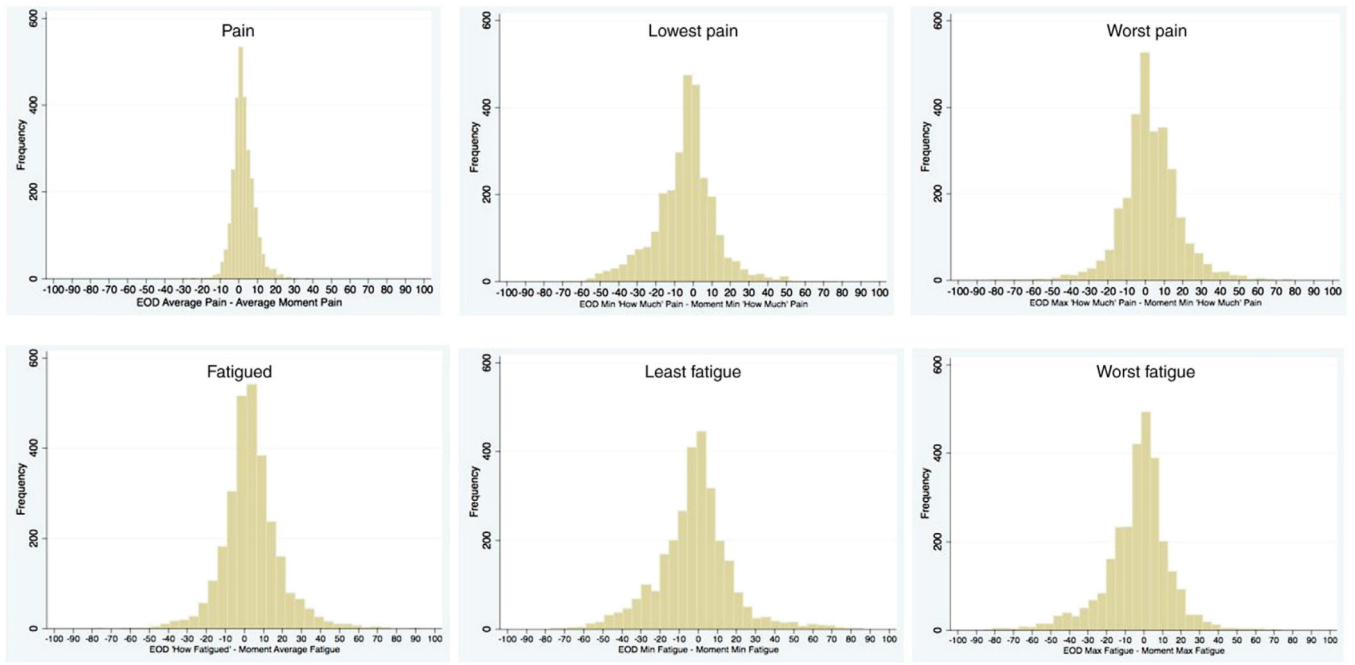


Figure 1. Histograms of difference scores between EOD and momentary variables for “how much” pain, lowest pain, and worst pain (upper panels) and “how” fatigued, least fatigued, and worst fatigue (lower panels).

Table 1

Levels of momentary assessments and end-of-day ratings and significance of their differences.

PAIN				
END-OF-DAY		MOMENTARY		Z-test (N=2825 days)
How much bodily pain did you have?	50.1 (24.5)	Average of "How much" bodily pain did you have?	45.0 (22.7)	-9.61 ***
How intense was your bodily pain?	47.1 (26.9)	Average of "How intense" was bodily pain?	42.2 (24.2)	-11.01 ***
What was your average level of pain today?	46.1 (22.1)	Average of "How intense" was bodily pain?	42.2 (24.2)	-4.66 ***
What was the lowest level of your pain today?	27.6 (22.5)	Computed: Minimum momentary "How much" bodily pain	32.1 (23.5)	4.23 ***
		Computed: Minimum momentary "How intense" was bodily pain	29.4 (24.3)	1.45
What was the worst level of your pain today?	61.4 (24.4)	Computed: Maximum momentary "How much" bodily pain	58.8 (23.4)	-3.71 ***
		Computed: Maximum momentary "How intense" was bodily pain	56.1 (25.6)	-5.97 ***
FATIGUE				
END-OF-DAY		MOMENTARY		Z-Test
How fatigued (weary, tired) did you feel?	51.2 (26.7)	How fatigued (weary, tired) did you feel?	47.5 (23.1)	-6.80 ***
How tired did you feel?	52.4 (25.5)	How tired did you feel?	47.0 (22.1)	-9.52 ***
What was the least level of your fatigue today?	28.6 (23.0)	Computed: Minimum momentary "How fatigued"	30.9 (24.6)	2.10 *
		Computed: Minimum momentary "How tired"	30.2 (23.3)	1.66
What was the worst level of your fatigue today?	61.6 (25.0)	Computed: Maximum momentary "How fatigued"	64.9 (23.7)	4.73 ***
		Computed: Maximum momentary "How tired"	65.0 (22.9)	4.67 ***

Note. Testing difference score with multilevel modeling.

* p<.05,

*** p<.001

Table 2

Correlations of EOD assessments with the corresponding momentary measure. Unit of analysis is a person-day.

	Correlation between EOD and Momentary EOD N=2825 days	Correlation between EOD and EMA (4-5 moments/day; mean=4.6) N=1384 days	Correlation between EOD and EMA (Minimum of 6 moments/day; mean=6.6) N=1441 days	Proportion of All Days with a Difference of greater than 10 points (on 0-100 point scale)	Proportion of All Days with a Difference of greater than 20 points (on 0-100 point scale)
PAIN					
EOD: "How much" MOMENT: Mean "How much"	.88	.88	.88	36%	11%
EOD: "How intense" MOMENT: Mean "How intense"	.90	.90	.90	33%	12%
EOD: "Average" MOMENT: Mean "How much"	.85	.84	.85	36%	11%
EOD: "Lowest level" MOMENT: Min "How much"	.74	.76	.73	45%	21%
EOD: "Lowest level" MOMENT: Min "How intense"	.73	.75	.72	46%	22%
EOD: "Worst" MOMENT: Max "How much"	.80	.80	.80	44%	17%
EOD: "Worst" MOMENT: Max "How intense"	.78	.78	.78	47%	21%
FATIGUE					
EOD: "How much" MOMENT: Mean Fatigued	.83	.84	.82	41%	16%
EOD: "How much" MOMENT: Mean Tired	.81	.83	.78	43%	18%
EOD: "Lowest" Fatigued MOMENT: Min Fatigued	.67	.69	.66	50%	25%
EOD: "Lowest" Fatigued MOMENT: Min Tired	.68	.70	.67	50%	24%
EOD: "Worst" Fatigued MOMENT: Max Fatigued	.75	.76	.73	46%	21%
EOD: "Worst" Fatigued MOMENT: Max Tired	.72	.75	.69	48%	22%

Table 3

Individual differences in Level (Person is unit of analysis; n=106).

	Age		Education		Health		Pain Level		Fatigue Level	
	Low	High	Low	High	Low	High	Low	High	Low	High
Pain										
Mean difference	5.3	5.0	5.2	4.8	5.2	5.1	5.9	4.4	4.9	5.4
Minimum difference	-3.7	-5.0	-6	-9.1***	-3.8	-4.7	.3	-9.1***	-4.9	-3.8
Maximum difference	2.7	2.8	2.5	3.0	2.8	2.7	4.5	.9*		
Fatigue										
Mean difference	2.9	4.3	3.6	3.3	3.2	3.8	3.0	4.2	3.4	3.8
Minimum difference	-5.5	1.1**	-1.0	-3.4	-2.0	-2.3	-2	-4.1	2.5	-6.8***
Maximum difference	-3.5	-3.5	-4.8	-1.9	-2.8	-3.9			-3.3	-3.7

* p<.05
 ** p<.01
 *** p<.001

Correspondence differences by Individual Differences (Day is unit of analysis; n=2825).

Table 4

	Age		Education		Health		Pain Level		Fatigue Level	
	Low	High	Low	High	Low	High	Low	High	Low	High
Pain										
Mean	.88	.88	.89	.87	.83	.89	.78	.80		
Minimum	.79	.67	.80	.71	.68	.75	.63	.61		
Maximum	.80	.79	.81	.77	.70	.81	.71	.68		
Fatigue										
Mean	.82	.82	.83	.84	.80	.83	.71	.71	.71	.71
Minimum	.69	.64	.68	.66	.62	.68	.49	.53		
Maximum	.73	.74	.76	.73	.70	.76	.68	.58		