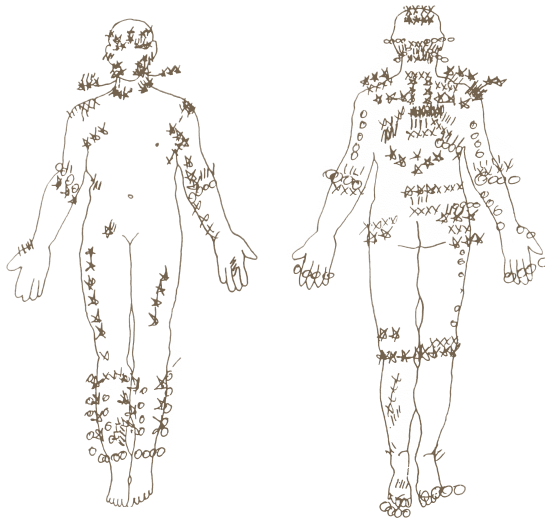


CHAPTER 8

How well do measurements of pain measured
by monthly questionnaires correlate
with data collected weekly?

Submitted

de Vet H.C.W., Ailliet L., Rubinstein S.M., van Tulder M.W., Hoekstra T.



ABSTRACT

Background

Frequent measurements are necessary to obtain detailed information about the clinical course of episodic conditions, such as neck pain and low back pain. In a prospective cohort study, patients visiting a chiropractor for neck pain and/or low back pain completed online questionnaires at baseline, and after 1, 3, 6 and 12 months about the number of days with pain and the number of days limited in activities of daily living (ADL) in the previous month. Patients also responded to weekly text messages (SMS tracking) in which the same questions were posed, but the time frame was the previous week.

Methods

The weekly SMS data were compared with the responses from the questionnaire using paired t-tests, where the standard deviation of the mean difference represents the agreement on the individual level. The mean differences represent agreement on group level. We analysed the data separately for patients with stable and those with fluctuating pain in the weeks prior to completing the questionnaire. Furthermore we examined by comparing Pearson correlation coefficients whether the SMS data of the most recent week corresponded better with the answer in the monthly questionnaires than the SMS data of the other three weeks.

Results

Of the 327 patients with neck pain, 169 (51.8%) agreed to participate in the SMS project, and of the 591 patients with low back pain, 326 (55.2%) participated. On the individual level, we identified substantial differences in days with pain reported by a monthly questionnaire as compared to the SMS data. For month 1, the standard deviation of the mean difference was 9.51 days for number of days with neck pain and 10.26 days for low back pain. For the number of days limited in ADL the standard deviations were 6.41 and 5.54, respectively. For month 3 we found similar results.

Patients with a rather stable pattern of pain throughout the month did not have a better agreement between either method of data collection as compared to patients with fluctuating pain patterns. Finally, the total number of days with pain and days limited in ADL recalled over a month period did not show the highest correlation with the number of days in the most recent week of that month.

Conclusions

This study showed substantial differences between reported number of days with neck pain and low back pain when collected by SMS tracking and by questionnaires. Agreement for both group level and individual level was better for days limited in ADL than

for days with pain. The way data are collected in studies affects the results. We assumed SMS data to be valid. However, more research is needed on the validity of data obtained by SMS tracking.

Neck pain and low back pain are common and costly conditions in Western society and it are major causes of pain and disability¹⁻³. For most patients, neck pain and low back pain represent recurrent conditions with pain free periods varying from weeks to months and even years. Other patients suffer from chronic (persistent) neck or low back pain with or without flare-ups and fluctuations⁴. In order to understand more about the clinical course of neck pain and low back pain, it is necessary to collect data frequently over a long period.

In most studies, the effects of interventions for relieving neck pain or low back pain have been examined by using questionnaires for assessment of outcomes at specific points in time. Typically, pain is measured at specific intervals throughout the course of a year, for example at 1, 3, 6 or 12 months. Pain intensity can be assessed in various ways depending upon the time frame, for example, current pain, pain in the previous 24 hours, average pain in the previous week or the previous month. Questions may also refer to the number of days that patients were in pain in the previous week or month. However, typical questionnaires yield limited information on the course of the disease because it relies either upon the patients' recall of pain over a long period of time, introducing potentially recall, or it covers only short periods. Nowadays, ecological momentary assessments (EMA) methods have become popular⁵. EMA refers to a collection of methods often used in behavioral medicine research by which a research participant repeatedly reports on symptoms, affect, behavior, and cognitions close in time to experience and in the participants' natural environment. EMA provides frequent snapshots of pain experiences during the clinical course, while avoiding of recall bias.

To study trajectories of neck and low back pain frequent measurements are needed. In a prospective cohort study on low back pain, Dunn et al.⁶ collected information regarding pain intensity over the previous two weeks, using monthly questionnaires during 6 months. They identified four different trajectories described as severe chronic pain, persisting mild pain, fluctuating pain, and recovering. In recent prospective studies in Denmark and Sweden⁷⁻¹⁰, weekly text messages sent to cell phones have been used for the assessment of pain in the previous week. This method made it possible to gather detailed data on the course of low back pain over a longer period.

We recently conducted a prospective cohort study¹¹ examining the course of neck pain and low back pain. We collected data by questionnaires completed at baseline and after 1, 3, 6 and 12 months, in which the number of days with pain and number of days limited in activities of daily living (ADL) in the previous week and month were queried. Additionally, we used SMS tracking on a weekly basis posing the same questions as in the questionnaires referring the past week.

The primary objective of this study was to compare data collected from the questionnaires to data collected by weekly SMS tracking. It has been reported that recall might be more heavily influenced by the most recent experiences^{12,13}. Therefore, as a secondary objective, we examined whether the number of days with pain and number of days limited in ADL as measured by a questionnaire for the previous month more strongly correlated with the SMS data for the number of days in the most recent week. Moreover, we postulated that the agreement between data from either method was likely to be better for patients with a stable course of pain as opposed to those with (widely) fluctuating pain and limitations in ADL.

METHODS

Design

A prospective, multi-center practice-based cohort study was conducted including patients with neck pain and/or low back pain. Participants were recruited by chiropractors working in private practices in Belgium and the Netherlands. All patients received standard chiropractic care, and additional treatments were left to the discretion of the chiropractor¹¹. The study was approved by the medical ethics committees of the University Ghent, Belgium under registration number B67020095664 and reference number 08/232 at VU University Medical Center, Amsterdam, the Netherlands.

Recruitment and inclusion of chiropractors and patients

CHIROPRACTORS

All 72 Dutch speaking Belgian chiropractors and all 189 members of the Netherlands Chiropractors' Association were invited to participate. The chiropractors were asked to recruit 15 patients who presented with neck or back complaints in their practice. Recruitment took place between August 26th and December 30th, 2010.

PATIENTS

Patients with non-specific neck pain and/or low back pain who had not been treated by a chiropractor within the prior six months were eligible for inclusion. In addition, patients had to be 18 to 64 years old, and their chief complaint had to be neck and/or low back pain with or without radiation to an extremity. If patients presented with both neck and low back pain, they were requested to choose which of the two complaints was worse. In addition, patients were required to have a basic understanding of the Dutch language, both in reading and writing.

Data collection: procedures and content of questions

Self-administered patient questionnaires were sent at baseline, and at 1, 3, 6 and 12 months. Patients could choose for paper- or web-based questionnaires. The responses of the web-based questionnaires were collected by NetQ (NetQuestionnaires Nederland BV 2002), which is an online data collection system. Paper questionnaires were returned to and processed at the data collection centers in Belgium and the Netherlands. The SMS part of the study was proposed as additional study with participation being on a voluntary basis. Using a fully automated Web-based system, SMS Track Questionnaire system (<https://sms-track.com>), questions were sent every Friday at 2 PM to the participating patients, who could easily provide their scores with a reply text message.

The weekly text messages contained four questions:

1. "On a scale from 0 to 10 (with 0 = no pain and 10 = worst pain imaginable), how would you rate your pain today"?
2. "On a scale from 0 to 10 (with 0 = not limited in activities of daily living (ADL) at all and 10 = extremely limited in ADL), how much are you limited in your activities of daily living today"?
3. "On a scale from 0 to 7, how many days did you experience neck pain (or low back pain) in the past week"?
4. "On a scale from 0 to 7, how many days were you limited in your activities of daily living in the past week"?

The answers on the questions were automatically transferred to a SMS-Track® data file. (SMS-Track Questionnaire 1.1.3. New Agenda Solutions; 2007. <https://www.sms-track.com>)

The self-administered questionnaires contained, among other things, the following questions: "How many days in the previous month did you have neck pain (or low back pain) (none = 0, all days = 31)"? and "How many days in the previous month were you limited in your activities of daily living (none = 0, all days = 31)"? We also inquired about the average pain intensity in the previous week, but refrained from asking about average pain intensity in the previous month as it is almost impossible to reconstruct the average intensity of pain over such a long period¹⁴. Data were compared from the questionnaire to the SMS tracking system for the number of days with neck pain (or low back pain) during the previous month and for the number of days limited in ADL.

Statistical analyses

To assess the agreement between data collected by questionnaire and by SMS tracking, we performed a paired t-test. For that purpose, we added the total number of days with pain reported by SMS tracking over the previous four weeks. In order to make it possible to compare data between months and 4 weeks, we applied a correction factor of 31/28, as the questionnaire covered a month of 31 days, and the SMS data 28 days.

TABLE 1A: CHARACTERISTICS OF SAMPLES OF NECK PAIN PATIENTS

CHARACTERISTIC	TOTAL SAMPLE (N=326)	QUESTIONNAIRE AT 1 MONTH (N=259)	SMS PARTICIPANTS (N=169)	TEXT MESSAGE AND QUESTIONNAIRE AT 1 MONTH (N=130)	p-VALUE ²
Gender (% female)	66.3%	65.6%	66.7%	64.6%	p=0.959
Age ¹	42.17 [11.80]	42.22 [11.81]	40.9 [11.21]	41.52 [11.28]	p=0.248
Duration (years) ¹	2.37 [0.84]	2.33 [0.86]	2.38 [0.84]	2.32 [0.86]	p=0.858
Pain intensity today (0-10) ¹	4.94 [2.21]	4.96 [2.12]	5.14 [2.14]	5.15 [2.13]	p=0.233
How much were you limited in ADL (0-10) ¹	3.28 [2.62]	3.22 [2.61]	3.25 [2.15]	3.16 [2.55]	p=0.365
Neck Disability Index (0-50) ¹	12.84 [6.18]	12.62 [5.91]	12.87 [6.30]	12.37 [5.70]	p=0.216

¹These characteristics are described as means [standard deviations].

²p-values correspond to Chi-squared test for categorical data or t-tests for continuous data comparing the subgroup of patients participating in the text message and questionnaire study versus those who did not

TABLE 1B: CHARACTERISTICS OF SAMPLES OF LOW BACK PAIN PATIENTS

CHARACTERISTIC	TOTAL SAMPLE (N=591)	QUESTIONNAIRE AT 1 MONTH (N=498)	SMS PARTICIPANTS (N=326)	TEXT MESSAGE AND QUESTIONNAIRE AT 1 MONTH (N=238)	p-VALUE ²
Gender (% female)	51.3	50.5	49.1	48.7	p=0.312
Age ¹	41.86 [11.75]	41.96 [11.72]	41.34 [11.21]	41.82 [10.89]	p=0.530
Duration (years) ¹	2.19 [0.89]	2.07 [0.92]	2.04 [0.91]	2.18 [0.91]	p=0.054
Pain intensity today (0-10) ¹	4.78 [2.21]	4.80 [2.22]	4.90 [2.17]	4.75 [2.16]	p=0.439
How much were you limited in ADL (0-10) ¹	3.77 [2.74]	3.78 [2.75]	3.72 [2.69]	3.42 [2.64]	p=0.888
Oswestry Disability Index (0-50) ¹	10.63 [7.67]	10.81 [7.8]	10.69 [7.48]	10.17 [7.35]	p=0.607

¹These characteristics are described as means [standard deviations].

²p-values correspond to Chi-squared test for categorical data or t-tests for continuous data comparing the subgroup of patients participating in the text message and questionnaire study versus those who did not

In order to examine whether the agreement between methods was better in patients with a stable pattern of pain as opposed to patients with a fluctuating pattern, we first calculated the average number of days with pain over four weeks and their corresponding standard deviations. We then compared the agreement between number of days with pain from text messages and questionnaires for that quartile of patients with the highest standard deviation (i.e. those with maximal fluctuation) and for that quartile with the lowest standard deviation (i.e. those with minimal fluctuation).

TABLE 2A: COMPARISON BETWEEN DAYS WITH PAIN PER MONTHS REPORTED IN THE WEEKLY TEXT MESSAGES AND BY QUESTIONNAIRE FOR NECK PAIN AND LOW BACK PAIN

	DAYS PER MONTH IN QUESTIONNAIRE	DAYS IN 4 WEEKLY TEXT MESSAGES	DIFFERENCE IN DAYS BETWEEN THE METHODS MEAN DIFFERENCE [SD]	PEARSON CORRELATION COEFFICIENT
Neck pain				
1 month	15.97	12.61	3.02 [9.51]	0.51
3 months	12.84	6.81	5.76 [10.06]	0.50
Low back pain				
1 month	15.97	12.35	3.62 [10.26]	0.41
3 months	11.81	6.07	5.74 [11.43]	0.26

TABLE 2B: COMPARISON BETWEEN NUMBER OF DAYS LIMITED IN ADL PER MONTH REPORTED IN THE WEEKLY TEXT MESSAGES AND BY QUESTIONNAIRE FOR NECK PAIN AND LOW BACK PAIN

	DAYS PER MONTH IN QUESTIONNAIRE	DAYS IN 4 WEEKLY TEXT MESSAGES	DIFFERENCE IN DAYS BETWEEN THE METHODS MEAN DIFFERENCE [SD]	PEARSON CORRELATION COEFFICIENT
Neck pain				
1 month	7.49	8.67	-1.08 [6.41]	0.75
3 months	5.77	6.21	-0.42 [3.55]	0.70
Low back pain				
1 month	8.46	8.95	-0.60 [5.54]	0.61
3 months	6.46	6.37	0.21 [3.10]	0.45

Finally, we examined whether the number of days recalled by patients with neck pain and low back pain obtained from the questionnaires over a month was influenced by the most recent pain experience (i.e. the number of days in the most recent week of that month). Therefore, we checked whether the recalled number of days with pain per month showed a higher Pearson correlation coefficient with the most recent week than with the earlier weeks of that month.

We encountered a large number of missing values due to technical problems with the SMS tracking system. There was a nine week period wherein which no text messages were sent. Therefore almost all patients had some missing data during the 12 month data collection period. These data were considered to be ‘missing completely at random’¹⁵. We used multiple imputation by predictive mean matching with 15 imputed datasets to impute missing values for patients who had at least SMS data for one week within the relevant month. Available SMS data for that patient, age, gender and the data provided in the questionnaires were used as additional predictors in the imputation model. Pooled results are presented¹⁵.

TABLE 3: DIFFERENCE IN DAYS WITH PAIN PER MONTH REPORTED IN THE WEEKLY TEXT MESSAGES AND IN THE QUESTIONNAIRE FOR NECK PAIN AND LOW BACK PAIN ACCORDING TO STABLE AND FLUCTUATING PAIN PATTERNS

	AVERAGE STANDARD DEVIATION	DIFFERENCE IN DAYS BETWEEN THE TWO METHODS MEAN [SD]
Neck pain		
Highest SD: 1 month	2.59	1.79 [9.66]
Lowest SD: 1 month	0.71	3.88 [8.87]
Highest SD: 3 months	2.34	4.35 [8.67]
Lowest SD: 3 months	0.25	5.05 [7.43]
Low back pain		
Highest SD: 1 month	2.51	0.75 [8.79]
Lowest SD: 1 month	0.57	7.12 [9.71]
Highest SD: 3 months	1.75	2.06 [11.89]
Lowest SD: 3 months	0.00	7.52 [10.91]

TABLE 4: CORRELATION OF NUMBER OF DAYS REPORTED IN THE WEEKLY TEXT MESSAGES AND BY QUESTIONNAIRE FOR NUMBER OF DAYS LIMITED IN ADL

	WEEK 1	WEEK 2	WEEK 3	WEEK 4
Neck pain				
Month 1	0.38	0.62	0.72	0.68
Month 3	0.60	0.68	0.67	0.63
Low back pain				
Month 1	0.62	0.75	0.72	0.71
Month 3	0.92	0.90	0.86	0.92

We followed the same procedure for all analyses for number of days limited in ADL. We present the data for the assessments after one month and three months.

RESULTS

Participation and response rates

In total, 97 chiropractors recruited 917 patients. Of the 326 participants with neck pain, 169 (51.8%) agreed to participate in the SMS project, and of the 591 with low back pain, 326 (55.2%) agreed to respond to the weekly text messages. After one month, these numbers diminished to 130 (77.0%) for neck pain patients and 238 (73.0%) for low back pain. The number of returned questionnaires after one month was 259 (79.4%) for neck pain patients and 498 (84.3%) for low back pain.

The characteristics of the total sample of neck pain and low back pain patients were compared with the samples of patients who completed the self-administered questionnaire at one month and the sample providing questionnaire data and SMS data at one-month follow-up. Tables 1A and 1B show no large differences in demographic and clinical characteristics between the total sample and subsamples.

Correspondence between number of days with pain and days limited in ADL obtained by questionnaire and text messages

Table 2A shows the agreement between the questionnaires and SMS in the corresponding four weeks period for number of days with neck pain and low back pain, and Table 2B for number of days limited in ADL. The number of days with pain as reported by the questionnaires exceeded the number of days reported by the text messages. This was true for all time points (6 and 12 months data not shown). The standard deviations of the differences were around 10 indicating that, on the individual level, large differences were observed between number of days with pain as reported by both methods. None of the differences were statistically significant.

The mean difference in number of days limited in ADL by both methods was negligible. This means that the agreement on group level between the questionnaires and SMS tracking was better for number of days limited in ADL than for number of days with pain. The standard deviations of the differences were somewhat smaller but still substantial, meaning that on the individual level we found substantial differences in reported number of days limited in ADL.

Table 3 shows the influence of stable (lowest SD) or fluctuating patterns (highest SD) on agreement between the number of days with pain obtained by both methods. We hypothesized that the patients with a rather stable pattern would be better able to estimate the number of days with pain in the previous month by questionnaire than those with fluctuating levels of pain, leading to smaller standard deviations of the differences between both methods. Results show that this was not the case. Remarkably, for both neck pain and low back pain patients, those with a fluctuating pattern obtained the best agreement (one month data). The data on number of days limited in ADL showed similar results: no better agreement for the stable patients (data not shown).

To explore the influence of each week separately, we examined the correlation of number of days limited in ADL as reported by the questionnaire with the weekly number of days obtained by SMS tracking. Table 4 shows that for number of days limited in ADL the correlations for week 4 (i.e., the most recent week) were not substantially higher than for the other three weeks. We found a similar pattern for the number of days with pain, although the strength of the correlation was lower, with correlations between 0.79 and 0.86 for the first month and between 0.45 and 0.64 for the third month. So, we found no convincing evidence for a clearly higher correlation with the most recent week.

DISCUSSION

Summary of results

In short, we identified substantial differences on the individual level in days with pain as reported by a monthly questionnaire as compared to data collected by SMS tracking. The agreement was better for number of days limited in ADL than for number of days with pain. Moreover, patients with a rather stable pain pattern throughout the month did not show better agreement between both methods compared to patients with fluctuating pain patterns. Finally, recall of total number of days with pain and days limited in ADL over the previous month appeared not to be most influenced by the reported SMS data in the most recent week of that month.

Comparison with other studies

Previous studies comparing different methods of data collection have focused on agreement in pain intensity and not number of days with pain. Moreover, pain and other symptoms have been subject of study many times but, as far as we know, studies on limitations in ADL are lacking. Furthermore, many studies compared EMA methods (measuring real time experiences) with methods based on pain recall¹⁶. Broderick et al.¹⁷ examined the influence of different periods of recall, compared with aggregated momentary ratings.

We identified one study that used a similar design as ours. In a study of 25 patients, Johansen and Wedderkopp⁷ compared in patients with low back pain the number of days with problems due to low back pain reported by weekly SMS tracking with the number of days obtained by telephone interview. The telephone interview was held in week 53 and queried patients regarding the number of days in the previous week, the previous month and previous year. For number of days in the previous month reported in the interview and calculated on 4 weeks SMS data, a mean difference of 0.7 days was found with limits of agreement¹⁸ between -4 days and 5 days. This is a remarkably good agreement, and much better than the rather poor agreement that we found. When we express our results as limits of agreement ($1.96 \times SD$ ¹⁸) these would be more than 20 days for the number of days with pain and between 6 and 12 days for the number of days limited in ADL. These poor results are supported by similar findings and explanations in the literature. First, there is evidence that different strategies are used to answer frequency questions covering short and long periods^{12,19}. For recall over short periods people tend to count individual incidents. This strategy is called episode enumeration. The other strategy, applied to longer periods, is called rate based estimations, and then people reconstruct the total by estimating it^{12,19}. In our study for the weekly SMS patients might tend to count the days, while for the questionnaires they probably estimate the number of days within the previous month. Secondly, the overestimation by questionnaire compared to momentary findings or short term recall has

been reported in many studies^{17,20}. People tend to overlook pain free periods²¹ and the recall of serious incidents or symptoms is better than more trivial ones²². Whether the latter observation might explain the better agreement for number of days limited in ADL than for number of days with pain is unclear. Rationally we would assume that that patients who are limited in ADL probably have more pain than patients who are less limited in ADL. From clinical experience we know that patients may be limited in ADL without reporting much pain, among other things because they avoid certain activities (kinesiophobia). This might explain the finding (comparing Tables 2A and 2B) that at the three months evaluation, for patients with low back pain the mean number of days limited in ADL (6.37) is even slightly higher than the number of days with pain (6.07).

Strengths and limitations

A major strength of this study was the large number of patients with neck pain and low back pain. This made it possible to conduct these analyses with great precision. Moreover, it provided the opportunity to examine the consistency of the results by presenting the comparisons for patients with neck pain and low back pain separately. The consistency of the results supports the robustness of our findings.

Selection bias might represent an important limitation of this study, that is, approximately only half of the patients participating in the cohort study were willing to volunteer for the SMS data collection. It is important to emphasize that the SMS tracking was explicitly suggested as an extra study in which they could participate. We intended to make it easy for participants of the main study to decline participation, as replying to the weekly text messages required a substantial effort and we considered a high compliance rate more important than a high participation rate. About half of the patients who agreed to participate did not differ much from those in the total sample (Tables 1A and 1B); therefore, selection bias is not likely to be an important issue at this phase. Of those patients who agreed to participate in the SMS project, approximately 25% did not start or dropped out within the first month. There were no differences in patient characteristics between drop outs and participants.

Due to technical problems the sending of text messages was interrupted for a period of nine weeks. For some patients this interfered with the weekly assessments during the first month, and for others with the three months or later assessments. As this was an incident, we assumed that these missing responses were missing completely at random. We used multiple imputation to handle these missing data. This also adjusted slight differences in characteristics of patients who had data reported by methods versus those who had missings after one or three months on questionnaire and/or SMS data.

A limitation of comparison of recall of data with data obtained in another way is that the patients answering the weekly text messages are more aware of the number of days with pain than patients who only receive a questionnaire. However, when filling

in the number of days with pain in the previous month in the questionnaire, the patients apparently did not calculate the sum of the previous four weekly text messages, otherwise the difference between the two methods could not have been that large.

Our assumption was that text messages are a more valid method of measuring days with pain than standard questionnaires because the recall time is shorter. However, patients may become complacent and therefore, do not correctly fill-in the text messages. On the other hand, if SMS tracking is not the 'gold standard' for collecting these type of data than what is? Future studies should compare data collected by SMS with other data collection methods to explore the validity of the method, preferably on a topic where the data obtained by SMS tracking can be compared with verifiable data, such the number of days of sick leave registered by a company.

CONCLUSIONS

This study demonstrated poor agreement on the individual level between reported number of days with pain when collected by text messages and by questionnaire; for number of days limited in ADL, the agreement was better but on the individual level we found substantial differences. Patients with a stable course of low back pain did not show a better agreement than those with a fluctuating course. We did not find convincing evidence that the most recent week was disproportionate weighted when estimating the number of days in the previous month.

REFERENCES

1. US Burden of Disease Collaborators. The state of US health, 1990e2010. Burden of disease, injuries and risk factors. *JAMA* 2013;310(6):591e610.
2. Hoy D, March L, Woolf A, Blyth F, Brooks P, Smith E, et al. The global burden of neck pain: estimates form the global burden of disease 2010 study. *Ann Rheum Dis* Jul 2014;73(7):1309e15.
3. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J* 2008;8:8e20.
4. Von Korff M, Saunders KW. The course of back pain in primary care. *Spine* 1996;21:2833–7.
5. Schiffman S. Real-time self-report of momentary states in the natural environment: computerized ecological momentary assessment. In: *The science of self-report. Imple-*

- cations for research and practice. Eds. Stone AA et al. Lawrence Erlbaum Associates, Inc. Publishers, Mahwah, NJ. U.S.A. 2000.
6. Dunn KM, Jordan K, Croft PR. Characterizing the Course of Low Back Pain: A Latent Class Analysis. *Am J Epidemiol* 2006;163:754–761.
 7. See comment in PubMed Commons below Johansen B, Wedderkopp N. Comparison between data obtained through real time data capture by SMS and a retrospective telephone interview. *Chiropractic and Manual Therapies* 2010; 18: 10.
 8. Axén I, Bodin L, Bergstrom G, Halasz L., Lange F, Lovgren PW, Rosenbaum A, Leboeuf-Yde C, Jensen I Clustering patients on the basis of their individual course of low back pain over a six month period. *BMC Musculoskeletal Disorders* 2011, 12:99.
 9. Kongsted A, Leboeuf-Yde C. The Nordic back pain subpopulation program-individual patterns of low back pain established by means of text messaging: a longitudinal pilot study. *Chiropr Osteopat.* 2009 Nov 17;17:11. doi: 10.1186/1746-1340-17-11.
 10. Axén I, Leboeuf-Yde C. Trajectories of low back pain. *Best Practice & Research Clinical Rheumatology* 27 (2013) 601–612.
 11. Ailliet L, Rubinstein SM, Knol D, van Tulder MW, de Vet HC. Somatization is associated with worse outcome in a chiropractic patient population with neck pain and low back pain. *ManTher.* 2015 Jul 20. pii: S1356-689X(15)00146-0. doi:10.1016/j.math.2015.07.007.
 12. Tourangeau R. Remembering what happened: Memory errors and survey responses. In: *The science of self-report. Implications for research and practice.* Eds. Stone AA et al. Lawrence Erlbaum Associates, Inc. Publishers, Mahwah, NJ. U.S.A. 2000 (page 43).
 13. Guyatt GH, Norman GR, Elizabeth F. Juniper EF, Griffith LE. A critical look at transition ratings. *Journal of Clinical Epidemiology* 2002; 55: 900–908.
 14. Stone AA, Broderick JE. Real-time data collection for pain: appraisal and current status. *Pain Med* 2007; 8 Suppl3: S85–93.
 15. Little RJ, Rubin DB. *Statistical analysis with Missing Data (Probability and statistics).* 2nd ed. Hoboken, NJ: John Wiley and Sons; 2002.
 16. Shiffman S. Real-time self-report of momentary states in the natural environment: computerized ecological momentary assessment . In: *The science of self-report. Implications for research and practice.* Eds. Stone AA et al. Lawrence Erlbaum Associates, Inc. Publishers, Mahwah, NJ. U.S.A. 2000 (pages 277–96).
 17. Broderick JE, Schwartz JE, Vikingstad GE , Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain* 2008; 139: 146–157.
 18. Bland JM and Altman DG. Measuring agreement in methods comparison studies. *Statistical Methods in Medical Research*, 8, 135–60.
 19. Burton S and Blair S. Task conditions, response formulation processes, and response accuracy for behavioural frequency questions. *Public Opinion Quarterly* 1991; 55: 50–79.

20. Van den brink M, Bandell-Hoekstra EN, Abu-Saad HH. The occurrence of recall bias in paediatric headache: a comparison of questionnaire and diary data. *Headache* 2001; 14: 11-20.
21. Stone AA, Broderick JE, Shiffman S, Schwartz JE. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. *Pain* 2004; 107: 61-9.
22. Schneider S, Stone AA, Schwartz JE, Broderick JE. Peak and end effects in patients' daily recall of pain and fatigue: a within-subjects analysis. *The Journal of Pain* 2011; 12: 228-235.