CARTILAGE
2022, Vol. 13(4) 50–58

DOI: 10.1177/19476035221118417
journals.sagepub.com/home/CAR
SAGE

Recall Bias Affects Pain Assessment in Knee Osteoarthritis: A Pilot Study

Davide Previtali[1*], Angelo Boffa[2*], Alessandro Di Martino[3], Luca Deabate[1], Marco Delcogliano[1,4], and Giuseppe Filardo[1,2,4]

Abstract

Objective. The objective of this study was to evaluate the recall bias of symptoms evaluation in knee osteoarthritis (OA). *Design.* In this multicentric pilot study, 50 patients with knee OA used a mobile App (Ecological Momentary Assessment [EMA]) to collect pain and function on two 0 to 10 numerical rating scales (NRS) 2 times a day for 2 months. At the 1-month and at the 2-month follow-up visits, patients retrospectively evaluated the mean level of pain/function of the last month. Recall bias was computed as the difference between the mean level of pain/function reported using the App and the level reported with the retrospective assessment. The correlation between the recall bias and patients' characteristics, as well as pain/function trajectories, was analyzed. *Results.* A statistically significant recall bias was documented with higher pain reported at 1-month with the retrospective assessment ($P < 0.001$). These results were confirmed also at the 2-month follow-up ($P = 0.002$). For function, no significant recall bias was documented. During the first and second months, 47 and 31 patients showed pain peaks, respectively. The number of pain peaks during the first month was correlated with the magnitude of the recall bias ($P = 0.02$). *Conclusions.* The recall bias influences the retrospective self-assessment of pain at the follow-up visits and the presence of pain peaks, a common event in the patients with OA, increases the magnitude of recall bias. The EMA performed with a mobile App is a useful tool to limit the influence of recall bias in the clinical and research setting evaluation of knee OA.

Keywords

knee, osteoarthritis, pain, ecological momentary assessment (EMA), recall bias

Introduction

Knee osteoarthritis (OA) is the most common joint disease with a prevalence of 10% in the older adults in developed countries, and according to PubMed, more than 3,000 articles are published every year on this topic.[1-3] Although knee OA is a chronic disease and the evolution of its symptoms over the years is slow,[4] most of the patients report unstable symptoms in their daily lives,[5] which makes it complex for clinicians and researchers to correctly track symptoms evolution. The assessment of symptoms in knee OA is commonly based on a questionnaire-based self-assessment,[6] although it has been demonstrated that the self-assessed evaluation of an experience is highly influenced by the fluctuation of symptoms.[7,8] Moreover, it can be affected by the clinical and research settings as well as by the momentary mood of the patients.[9] In this light, the proper self-assessment of pain and function by patients with knee OA at follow-up, as commonly performed for clinical and research purposes, can be affected by the inability of patients to correctly report their symptoms experience due to a recall bias.[10]

[1]Service of Orthopaedics and Traumatology, Department of Surgery, EOC, Lugano, Switzerland
[2]Applied and Translational Research Center, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy
[3]Clinica Ortopedica e Traumatologica 2, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy
[4]Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland

Investigation performed at the Service of Orthopaedics and Traumatology, Department of Surgery, EOC, Lugano, Switzerland and at the Clinica Ortopedica e Traumatologica 2, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy

*Dr. Previtali and Dr. Boffa have contributed equally to this article and share first authorship.

Corresponding Author:
Alessandro Di Martino, Clinica Ortopedica e Traumatologica 2, IRCCS Istituto Ortopedico Rizzoli, Via di Barbiano 1/10, 40136 Bologna, Italy.
Email: alessandro.dimartino@ior.it

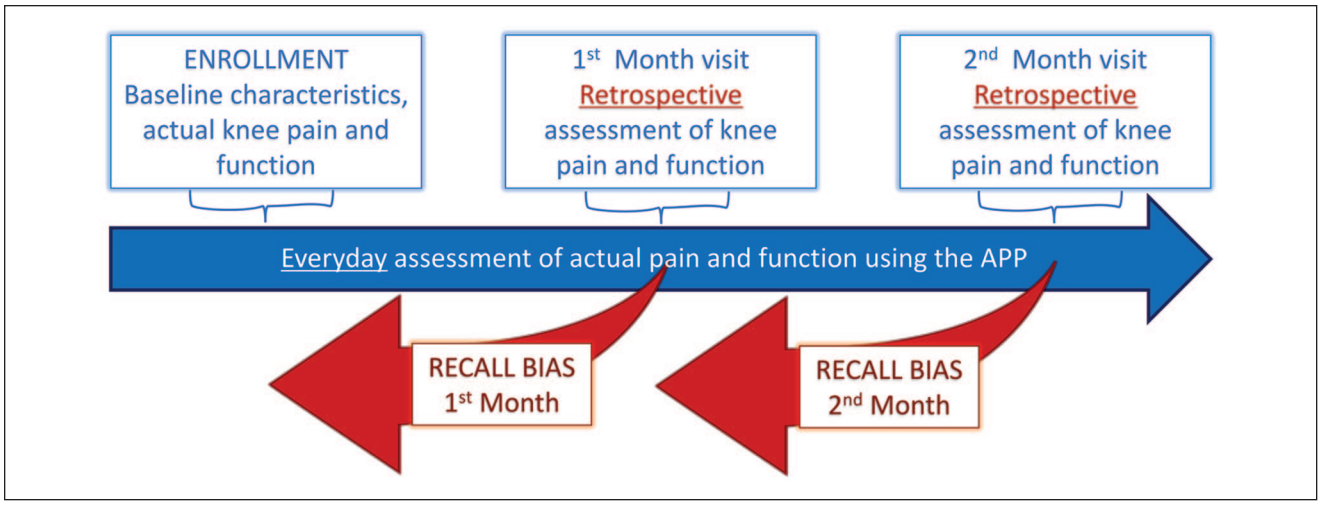

Figure 1. Schematic representation of the design of the study.

Ecological Momentary Assessment (EMA) is a method used to record self-assessed disease-related symptoms that consists of targeted but frequent and repeated reports that are recorded by the patient in a diary.[10,11] By recording symptoms closer to the time at which the patient feels them, the risk of recall bias can be overcome.[12] Furthermore, data are collected in the patient real-world environment providing information that reflects the real-life experience of the patient, limiting the influence of the presence of the physician and of the hospital/research setting.[10] Modern technologies, such as smartphone Apps, have shown to be well accepted by the patient and can help physicians to deal with the big amount of data that have to be recorded and managed when using the EMA approach.[13,14] In this way, EMA can be a useful and feasible tool for the comprehension of a disease pattern and response to treatment, offering to physicians and researchers a more precise assessment of the symptoms suffered by patients with knee OA.[15] However, presence and magnitude of recall bias, as well as the potential of the EMA to assess symptoms at the follow-up visits, have never been documented in the evaluation of knee OA symptoms.

The aim of this study was to investigate the potential of EMA over the traditional recall-based assessment, by documenting and quantifying the recall bias in the study of knee OA patients. Furthermore, punctual recorded data have been used to analyze pain and function trajectories, and to investigate patients' characteristics and aspects of the symptoms experience influencing the recall bias in the evaluation of the knee OA symptoms.

Methods

Study Design and Study Subjects

This study is an observational single-arm prospective multicentric pilot study approved by the Ethic Committees (prot. nr BASEC 2019-00891 and prot. ORL-ORT-010). The selection of the patients was performed at the outpatient facilities of the Ospedale Regionale di Lugano, Lugano, Switzerland (CH) and of the IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy (IT). After the signature of the informed consent, participants affected by knee OA were screened and, if eligible, included in the study. Patients were introduced to the use of the mobile App ("EOC EMApp") for their smartphone to record the EMA: "EOC EMApp" asked (sending notifications to her/his mobile phone) the patients to rate actual pain and function on two 0 to 10 numerical rating scale (NRS) 2 times a day (at 10:00 in the morning and at 18:30 in the evening, to double our data and better document patient status and the changes over time) for 2 months. After 1 month and 2 months, follow-up visits were performed, and patients were asked to retrospectively evaluate the mean pain they suffered and the level of function they had during the last month on a 0 to 10 NRS (**Fig. 1**).

The following baseline characteristics of the patients were recorded: age, sex, body mass index (BMI), affected knee, Kellgren-Lawrence (K-L) grade, and symptoms duration. Female and male patients with a clinical and radiological diagnosis of knee OA, suffering from knee OA symptoms (knee pain, stiffness, and loss of function), older than 18 years, with a BMI between 18.5 and 35 kg/m^2, owning a smartphone with Android or iOS as operative systems, were included. Patients unable to follow the study protocol, asymptomatic, without a radiological confirmation of knee OA, with a planned surgery in the next 60 days, with a surgical treatment at the knee in the last year, with a knee injection in the last 6 months, or with another disease causing knee symptoms were excluded to be sure as much as possible that the reported symptoms were related to the knee OA disease.

Outcome Measures

Knee pain and function were evaluated with a 0 to 10 NRS both for the EMA (twice a day) and for the retrospective

assessment (1 month, 2 months). The 0 to 10 NRS is a valid and reliable way used for the self-assessment of symptoms: it consists of a single 11-point numeric scale, with 0 indicating no pain and 10 reflecting the worst possible pain or 0 indicating no function and 10 indicating optimal function.[16] The primary outcome of the study was the difference between the level of pain reported with the retrospective assessment and the mean level of pain reported during the previous month with the EMA, which is representative of the magnitude of the recall bias, at 1 month. Similarly, the magnitude of the recall bias was computed for the level of pain reported during the second month, as well as the level of function reported during both the first and the second months.

Pain and function were analyzed considering

- the mean level of pain (or function) of the whole month,
- the mean level of pain (or function) of every single week,
- the evolution of pain (or function) in the last 2 weeks (the difference between the mean level of pain of the last week and the mean level of pain of the third week),
- the evolution of pain (or function) during the whole month according to the retrospective assessment (the difference between the level of pain/function reported with the retrospective assessment at the 1-month follow-up and at baseline and the difference between the level of pain/function reported at the 2-month and at the 1-month follow-ups; to compute of patients with a change a clinically significant change of 1 point on the 0 to 10 NRS was considered),[17]
- the presence of pain (or function impairment) peaks,
- the number of pain (or function impairment) peaks. Peaks of pain or function were defined by transitory (for less than 3 consecutive days) increases of pain intensity (or decrease of function) of more than 1 point above the mean of pain intensity over the 1-month follow-up period.[18]

These factors were considered as possible determinants of the recall bias. Moreover, age, sex, BMI, K-L grade, and symptoms duration were baseline patients' characteristics that were considered as possible determinants of the recall bias as well.

Sample Size Calculation and Statistical Analysis

No previous studies evaluated the recall bias with an EMA in patients with knee OA. Data on the various musculoskeletal pain conditions and follow-up length are heterogeneous with a computed effect size ranging from 0.14 to 1.06.[14,15] In this light, a medium effect size of 0.4, according to Cohen, was used to determine the sample size.[19] Setting the α-error at 0.05 a sample size of 41 patients is needed to have a statistical power of power of 0.80 (β-error = 0.20). Considering a 20% dropout rate, a total of 50 patients were planned.

The mean level of pain (or function) reported with the EMA and the level of pain reported with the retrospective assessment were expressed as means with standard deviations. The paired *t* test was used to assess their difference. For each patient, the "recall bias" was computed as the difference between the mean level of pain (or function) reported with the EMA and the level of pain reported with the retrospective assessment and the unpaired *t* test was used to assess the influence of possible predictors. All the continuous baseline variables were expressed as means with standard deviations and compared computing mean differences whose significance was tested using a Student *t* test for independent means. Non-continuous baseline variables were expressed as frequencies, compared computing risk ratios whose significance was tested using a chi-squared test. The correlation between possible determinants of the recall bias and the magnitude of the recall bias were calculated using the Spearman correlation methods. Significance level was 2-sided with α error set at 0.05.

Results

Characteristics of the Included Patients

A total of 59 patients were asked to participate in this study from July 2019 to August 2020 at the study investigation sites (24 at the Ospedale Regionale di Lugano and 35 at the Istituto Ortopedico Rizzoli). Among these, 9 patients were excluded after enrolment: 5 used the mobile application only for few days, 2 did not start using the application, 1 was not able to use the application properly (did not understand how to answer the questions), and 1, after reporting pain and function impairment at the baseline visit, reported no pain and a perfect function during the follow-up period. Moreover, 2 additional patients did not complete the second month assessment since they decided to stop using the mobile application and 1 did not attend the 2-month follow-up visit (**Fig. 2**). There was no significant difference between the included and the dropout patients in terms of sex, age, BMI, length of symptoms, and baseline K-L grade.

Twenty-six out of the 50 patients who used the application properly were women, 24 were men. The mean age was 58.4 $\pm$ 12.2 years, the mean BMI was 26.9 $\pm$ 5.0 kg/m^2, and the mean length of symptoms was 5.2 $\pm$ 4.7 years. Regarding OA severity, 8 patients had grade 1 knee OA, 22 had grade 2 knee OA, 14 had grade 3 knee OA, and 6 had grade 4 knee OA. The mean level of pain at baseline was 3.4 $\pm$ 2.6. There was no significant difference between the patients enrolled in the 2 centers in terms of sex, age, BMI,

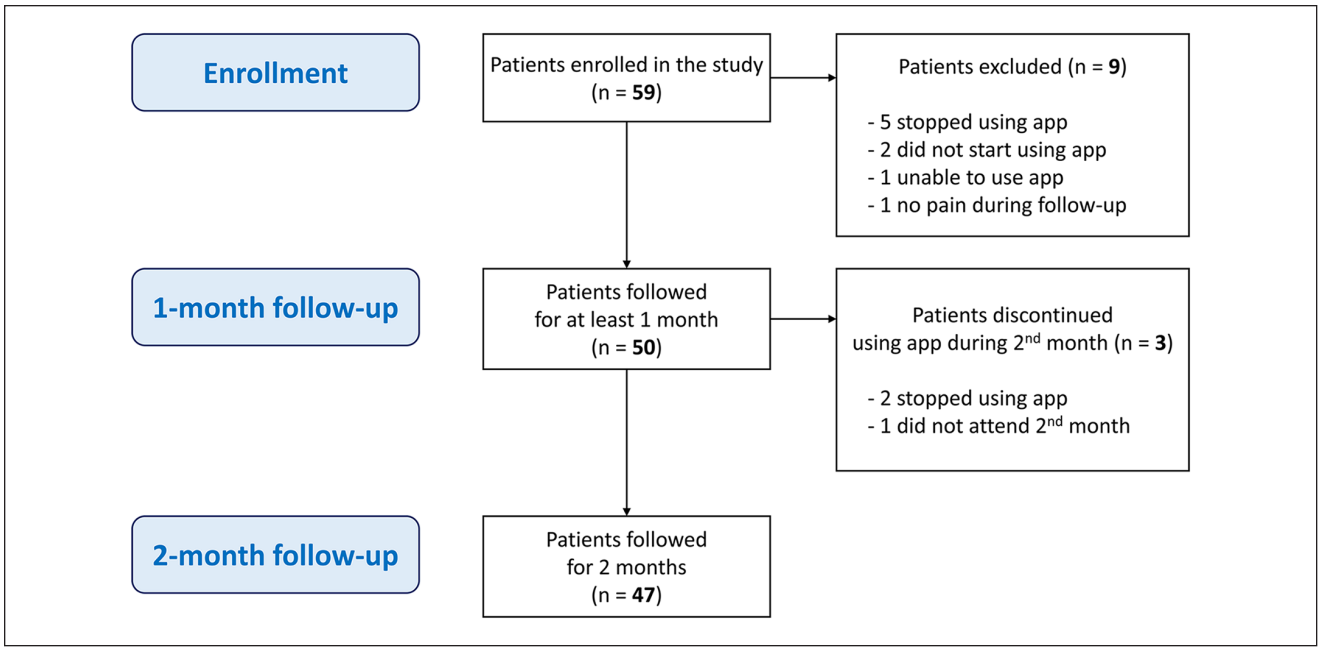

Figure 2. Strengthening the reporting of observational studies in epidemiology (STROBE) flowchart showing patient enrolment and evaluation.

baseline K-L grade, and baseline pain, whereas a significant difference was detected in terms of length of symptoms (greater time since the onset of symptoms for the patients enrolled in Bologna). Details are reported in **Table 1**.

Symptom Trajectories

Pain during the first month, according to the retrospective assessment, increased in 26 patients, decreased in 12 patients, and was stable in 12 patients with a mean increase from baseline to 1 month of 0.69 (P = n.s.). During the second month, pain increased in 19 patients according to the retrospective assessment, decreased in 10 patients, and was stable in 18 patients (3 patients did not complete the 2-month assessment), with a mean increase from the first month to the second month of 0.32 (P = n.s.).

From baseline to the second month follow-up, pain increased in 32 patients according to the retrospective assessment, decreased in 8 patients, and was stable in 7 patients (3 patients did not complete the 2-month assessment), with a mean increase from the baseline to the second month of 1.01 ($P = 0.004$).

Function during the first month, according to the retrospective assessment, improved in 15 patients, worsened in 13 patients, and was stable in 22 patients with a mean change from baseline to 1 month of 0.09 (P = n.s.). During the second month, according to the retrospective assessment, function improved in 9 patients, worsened in 13 patients, and was stable in 25 patients (3 patients did not attend the 2-month follow-up) with a mean change from the first month to the second month of −0.11 (P = n.s.).

From baseline to the second month follow-up, according to the retrospective assessment, function improved in 11 patients, worsened in 14 patients, and was stable in 22 patients (3 patients did not attend the 2-month follow-up) with a mean change from baseline to the second month of 0.02 (P = n.s.).

During the first month, only 3 out of 50 patients reported no pain peaks, whereas during the second month, 17 out of 48 patients reported no pain peaks. Regarding function, during the first month, 15 out of 50 patients reported no function impairment peaks, whereas during the second month, 21 out of 48 patients reported no function impairment peaks. The mean number of pain peaks was 4.0 during the first month and 2.7 during the second month. The mean number of function impairment peaks was 2.9 during the first months and 2.0 during the second month.

EMA Versus Traditional Assessment

A statistically significant difference was identified between the level of pain documented with the EMA (mean of the self-assessments made during the last month) and the one documented retrospectively (mean pain suffered during the last month as reported at the follow-up visit) with a higher pain reported at 1 month with the retrospective assessment (mean difference [MD] = 0.41, 95% confidence interval [CI] = 0.63-0.20, $P < 0.001$, **Fig. 3**) with no difference between the two study centers (P = n.s.). These results were confirmed also at the 2-month follow-up, with a statistically significant higher pain reported with the retrospective assessment (MD = 0.50, 95% CI = 0.81-0.19, $P = 0.002$) with no difference between the two study centers (P = n.s.).

Table 1. Characteristics of the Included Patients.

Characteristic	Sub-Group	Value	Significance
Age	All patients	58	n.s.
	Dropouts	62	
	Lugano	62	n.s.
	Bologna	57	
Sex	All patients	24 M 26 F	n.s.
	Dropouts	4 M 5 F	
	Lugano	7 M 11 F	n.s.
	Bologna	17 M 15 F	
BMI	All patients	27 kg/m^2	n.s.
	Dropouts	28 kg/m^2	
	Lugano	26 kg/m^2	n.s.
	Bologna	27 kg/m^2	
Length of symptoms	All patients	5 years	n.s.
	Dropouts	7 years	
	Lugano	3 years	**0.047**
	Bologna	6 years	
Kellgren-Lawrence score	All patients	8 KL1 22 KL2 14 KL3 6 KL4	n.s.
	Dropouts	3 KL1 3 KL2 3 KL3 0 KL4	
	Lugano	3 KL1 9 KL2 5 KL3 1 KL4	n.s.
	Bologna	5 KL1 13 KL2 9 KL3 5 KL4	
Baseline pain	All patients	3.4	n.s.
	Dropouts	3.3	
	Lugano	3.2	n.s.
	Bologna	3.5	

n.s. = not significant. Bold value is statistically significant.

In the evaluation of function, no differences were identified between the level documented with the EMA and the level documented retrospectively during both the first (MD = −0.06, 95% CI = 0.25 to −0.36, P = n.s.) and the second (MD = 0.01, 95% CI = 0.19 to −0.16, P = n.s.) month of follow-up with no differences between the two study centers (P = n.s.).

Factors Influencing the Recall Bias

Since the evaluation of pain assessment during the first month showed the presence of a significant recall bias, possible influencing factors were tested on this outcome. The number of pain peaks showed to be correlated with the magnitude of the recall bias in pain assessment during the first month, with more peaks correlating with a greater recall bias (ρ = 0.32, P = 0.02) (**Figs. 4** and **5**). No significant correlation with the magnitude of the recall bias was documented for age (P = n.s.), sex (P = n.s.), BMI (P = n.s.), length of symptoms (P = n.s.), baseline reported pain (P = n.s.), evolution of pain in the last 2 weeks (P = n.s.), and K-L grade (P = n.s.). No correlations were documented during the second month (**Table 2**).

When the correlation between the mean level of pain reported in every single week and the level of pain reported retrospectively was tested, no difference was reported among the 4 weeks.

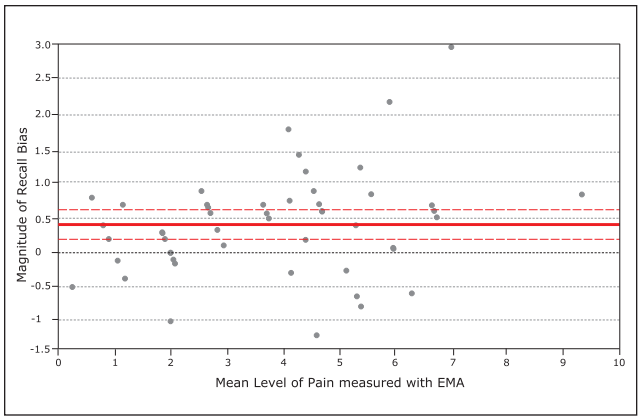

Figure 3. Bland-Altman plot for pain during the first month: patients (black dots) are distributed based on the mean level of pain measured with EMA (x-axis) and the magnitude of the recall bias (y-axis). Continuous red line represents the overall recall bias, dotted red lines represent the 95% confidence interval, dotted black line represents the 0. EMA = Ecological Momentary Assessment.

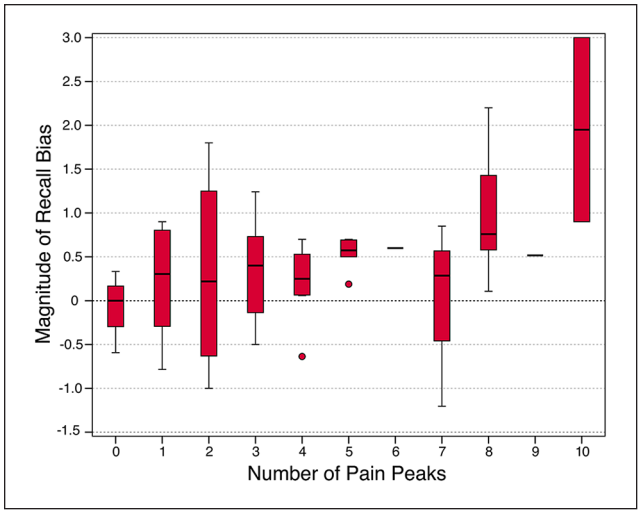

Figure 4. Box plot showing the correlation between the number of pain peaks during the first month and the magnitude of recall bias.

Discussion

The main finding of this study is that pain assessment at follow-up, as currently performed in knee OA studies, is affected by a recall bias. The level of pain documented through a daily monitoring and the level of pain asked retrospectively at the monthly follow-up visit are different, and both at the first month and at the second month follow-up, patients report an overestimation of pain with the traditional assessment.

Patients tend to report a higher level of pain when asked to recall the level of pain suffered during the last month at the follow-up visit. These results in knee OA patients are in line with the ones documented in several other fields, such as hemicrania or nausea and vomit intensity,[20,21] and in different types of patients in terms of age, severity of symptoms, and so on.[22-25] Thus, recall bias should be taken into account and efforts should be done to limit its influence on the results of research and clinical assessment of knee OA. In fact, the difference between the level of pain documented in this pilot study with the EMA and the one documented retrospectively could affect the results of many studies on knee OA. Accordingly, the biased results obtained by several studies could lead to misleading conclusions, with potential important consequences on the conclusions of previous clinical trials. Even more crucial may be the role of recall bias in the clinical setting: the tendency of the patients to overestimate their symptoms at the follow-up visit can influence the treatment approach of the physician who may be more aggressive than needed. Moreover, these data also underlined the common presence of pain peaks, which could have a relevant impact on the evaluation of the physician, as well. With the aging of the population and the increasing number of total knee replacement performed for patients presenting with knee OA related pain, EMA could be extremely useful preventing recall bias and thus avoiding unnecessary procedures for patients overestimating their pain due to the high number of pain peaks.[26]

EMA should be considered not only a good option to help the researchers and the physicians in the evaluation of the symptoms suffered by their knee-OA patients, but rather a necessary tool to properly interpret the findings of more classic questionnaire-based evaluation methods. Despite the potential of this approach, the big amount of data that need to be recorded and managed properly was a problem which limited the application of the EMA. Traditionally, this was problematic, as it was achieved using paper diaries with a lot of missing data, the possibility for the patient to retrospectively fulfill diaries in case of forgotten reports and a poor acceptance of the method by the patients.[27] However, modern technologies, such as smartphone applications, can help the patients and physician and have shown to be well accepted by patients.[13] In this regard, the present study showed that only 8 out of 59 patients stopped using the mobile application during the first month, thus confirming the good acceptance of this tool. Moreover, the average age of this patient group was 58 and it is possible that the acceptance rate would be even higher when studying different diseases and patient categories. Also, the use of mobile applications and smartphones will likely increase in the next years. Overall, both patient acceptance and documented results support that the collection of EMA with a mobile application is a suitable solution to avoid recall bias and increase our knowledge of knee OA and of the pain experience of the affected patients.

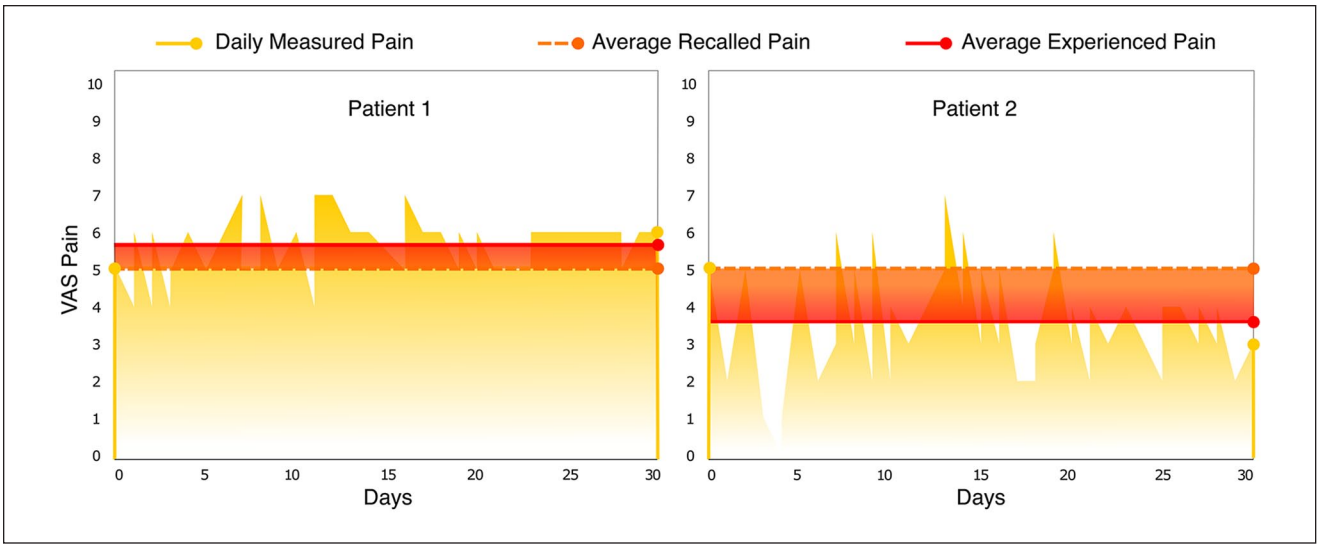

Figure 5. Pain representation of 2 different patients during the first month of assessment. Patient 1 (graph on the left) reported retrospectively a level of pain (5.0) similar to that evaluated with ecological momentary assessment (EMA) (5.6). On the other hand, patient 2 (graph on the right) reported retrospectively a level of pain of 5.0 but the mean level of pain collected with EMA during the month was 3.6, documenting a great recall bias (Mean Difference: 1.4). This could be due to the great number of pain peaks suffered during the month. While the pain level recalled retrospectively was similar, the real pain experience of these 2 patients was different.

Table 2. Results of the Correlation Analysis Between Possible Influencing Factors and the Magnitude of Recall Bias.

	First Month		Second Month	
Influencing Factor	Correlation Coefficient (ρ)	Significance	Correlation Coefficient	Significance
Number of pain peaks	0.32	**0.02**	0.12	n.s.
Age	0.18	n.s.	−0.001	n.s.
Sex	N/A	n.s.	N/A	n.s.
Body mass index	0.13	n.s.	−0.17	n.s.
Length of symptoms	−0.20	n.s.	0.12	n.s.
Mean reported pain	0.09	n.s.	−0.18	n.s.
Kellgren-Lawrence grade	0.18	n.s.	−0.13	n.s.
Evolution of pain in the last 2 weeks	0.20	n.s.	−0.19	n.s.

N/A = not applicable; n.s. = not significant. Bold value is statistically significant.

The significant findings of recall bias in the pain assessment were not confirmed for knee function, where the presence of the recall bias was not documented in this series. The reason behind this divergency may be manifold. First, to limit the everyday commitment of the patients, a single 0 to 10 NRS was used to evaluate knee function and this outcome measure. Despite being validated for its use in knee OA,[28] this method may be too simple to provide a whole caption of a complex symptom such as knee function impairment. In particular, KOOS (Knee Injury and Osteoarthritis Outcome Score) and WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index) are 2 validated and probably the most used scales to evaluate the symptoms of knee OA[29,30]; however, due to their complexity (patients should answer 39 questions for KOOS and 24 for WOMAC) they require a mean of 10 and 5 minutes for their completion and this could have limited the compliance of patients in this proof-of-concept trial. Besides these, several other complex scales are used in the clinical and research setting of knee OA to quantify its symptoms and their implementation with an EMA approach should be investigated in further studies to better characterize the recall bias.[6] Second, based on the data of the present study, function tends to be more stable with 15 patients reporting no function worsening peaks compared with 3 patients reporting no pain peaks during the first month. Since pain peaks showed to influence the recall bias, their lower number in terms of function could be the reason for the difference observed between function and pain evaluation in this series.

The subjective evaluation of the pain experience may be influenced by several factors, both disease- and patient-related.[31] In the present study, the presence of pain peaks and a higher number of pain peaks during the previous month significantly influenced the recall bias causing a greater overestimation of the suffered symptoms. This finding confirms also in the field of knee OA the thesis of the literature on psychological sciences that retrospective evaluations of chronic negative symptoms are often dominated by the discomfort perceived at the worst moments of the experience.[24] However, a clear correlation between the level of pain in the last days and the magnitude of the recall bias could not be documented, as well as for the evolution of pain during the previous month. Similarly, the length of symptoms and the mean reported pain also showed no correlation with the recall bias. Regarding patients-related characteristics none of the evaluated ones showed a statistically significant influence on the recall bias: age, sex, BMI, and K-L grade were all unrelated to the magnitude of the documented bias in this series. These factors should be further investigated in larger series. Moreover, other potential influencing factors, such as physician mood and the ambulatory environment are difficult to be quantified but may play an important role as well.[31]

The lack of an analysis on the role of potentially important influencing factors such as the physician's and patient's mood and the ambulatory environment is a limitation of the present study and should be investigated in future trials.[32,33] Another limitation, as previously specified, may be that the 0 to 10 NRS, despite being simple and well tolerated by the patient for a daily assessment, may have a limited capability to quantify a complex variable such as knee function. This may be the reason for the absence of recall bias for the function assessment in the present study, and future studies should clarify the impact of the recall bias in the evaluation of knee function with more complex scales. Furthermore, the experimental contest may have influenced the capability to detect the recall bias: the fact that patients were informed that they will be asked to evaluate their symptoms at the follow-ups and that the evaluation took place after a month in which they performed a daily evaluation of their symptoms may have improved the self-consciousness regarding their symptoms. Another possible limitation is that no information on the physical activity performed during EMA collection or on the concomitant use of painkillers or physical therapy was obtained. The primary aim of this article was to document the presence of recall bias in knee OA independently by the activity performed during collection. Nonetheless, new studies should focus on these aspects to better characterize pain trajectories, also considering activity level and the use of painkillers and physical therapy as possible influencing factors. Finally, while properly powered for the primary outcome, this study was intended as a proof-of-concept study and only 50 patients were included in the trial, possibly hindering the present study in reaching statistical significance in some of the sub-analyses.

Despite these limitations, this pilot study was able to demonstrate important findings. Recall bias influences the retrospective self-assessment of pain at the follow-up visits and the presence of pain peaks during the last month, a common event in patients with OA, increases the magnitude of the recall bias. Thus, EMA performed with a mobile App should be considered as a useful tool to limit the influence of recall bias and improve the interpretation of the patient data in the clinical and research setting of knee OA.

Acknowledgements and Funding

The authors acknowledge Elettra Pignotti for her help with the statistical analysis. This research (and APC) were partially funded by the Italian Ministry of Health — 5 × 1000 Anno 2018, Redditi 2017 "Valutazione dell'esperienza sintomatologica nel paziente con dolore cronico da patologia ortopedica mediante Ecological Momentary Assessment".

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval

The study was approved by the Ethics Committee of the Ospedale Regionale di Lugano, Lugano, Switzerland (Prot. BASEC 2019-00891) and the Ethics Committee of the IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy (Prot. ORL-ORT-010).

ORCID iDs

Davide Previtali https://orcid.org/0000-0002-0284-4368
Angelo Boffa https://orcid.org/0000-0002-1523-6900

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, *et al*. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the global burden of disease study 2010. Lancet. 2012;380(9859):2163-96. doi:10.1016/S0140-6736(12)61729-2.
2. Dillon CF, Rasch EK, Gu Q, Hirsch R. Prevalence of knee osteoarthritis in the United States: arthritis data from the third national health and nutrition examination survey 1991-94. J Rheumatol. 2006;33(11):2271-9.
3. Guillemin F, Rat AC, Mazieres B, Pouchot J, Fautrel B, Euller-Ziegler L, *et al*. Prevalence of symptomatic hip and knee osteoarthritis: a two-phase population-based survey. Osteoarthritis Cartilage. 2011;19(11):1314-22. doi:10.1016/j.joca.2011.08.004.
4. Collins JE, Katz JN, Dervan EE, Losina E. Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2014;22(5):622-30. doi:10.1016/j.joca.2014.03.009.
5. Previtali D, Andriolo L, Di Laura Frattura G, Boffa A, Candrian C, Zaffagnini S, *et al*. Pain trajectories in knee osteoarthritis-a systematic review and best evidence synthesis

on pain predictors. J Clin Med. 2020;9(9):2828. doi:10.3390/jcm9092828.
6. Davis AM, King LK, Stanaitis I, Hawker GA. Fundamentals of osteoarthritis: outcome evaluation with patient-reported measures and functional tests. Osteoarthritis Cartilage. 2022;30:775-85. doi:10.1016/j.joca.2021.07.016.
7. Stone AA, Schwartz JE, Broderick JE, Shiffman SS. Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic. Pers Soc Psychol Bull. 2005;31(10):1340-6. doi:10.1177/0146167205275615.
8. Teirlinck CH, Sonneveld DS, Bierma-Zeinstra SMA, Luijsterburg PAJ. Daily pain measurements and retrospective pain measurements in hip osteoarthritis patients with intermittent pain. Arthritis Care Res (Hoboken). 2019;71(6):768-76. doi:10.1002/acr.23711.
9. Kihlstrom JF, Eich E, Sandbrand D, Tobias BA. Emotion and memory: implications for self-report. Mahwah (NJ): Lawrence Erlbaum; 2000.
10. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol. 2008;4:1-32. doi:10.1146/annurev.clinpsy.3.022806.091415.
11. May M, Junghaenel DU, Ono M, Stone AA, Schneider S. Ecological momentary assessment methodology in chronic pain research: a systematic review. J Pain. 2018;19(7):699-716. doi:10.1016/j.jpain.2018.01.006.
12. Domingo-Salvany A. The science of real-time data capture: self-reports in health research. J Epidemiol Community Health. 2008;62(5):471. doi:10.1136/jech.2007.068551.
13. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain. 2001;91(3):277-85. doi:10.1016/S0304-3959(00)00450-4.
14. Garcia-Palacios A, Herrero R, Belmonte MA, Castilla D, Guixeres J, Molinari G, *et al*. Ecological momentary assessment for chronic pain in fibromyalgia using a smartphone: a randomized crossover study. Eur J Pain. 2014;18(6):862-72. doi:10.1002/j.1532-2149.2013.00425.x.
15. Perrot S, Marty M, Legout V, Moyse D, Henrotin Y, Rozenberg S. Ecological or recalled assessments in chronic musculoskeletal pain? a comparative study of prospective and recalled pain assessments in low back pain and lower limb painful osteoarthritis. Pain Med. 2011;12(3):427-36. doi:10.1111/j.1526-4637.2011.01052.x.
16. Karcioglu O, Topacoglu H, Dikme O. A systematic review of the pain scales in adults: which to use? Am J Emerg Med. 2018;36(4):707-14. doi:10.1016/j.Ajem.2018.01.008.
17. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain. 2004;8(4):283-91. doi:10.1016/j.ejpain.2003.09.004.
18. Trouvin AP, Marty M, Goupille P, Perrot S. Determinants of daily pain trajectories and relationship with pain acceptability in hip and knee osteoarthritis. a national prospective cohort study on 886 patients. Joint Bone Spine. 2019;86(2):245-50. doi:10.1016/j.jbspin.2018.06.009.
19. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New York: Routledge; 1988.
20. van den Brink M, Bandell-Hoekstra EN, Abu-Saad HH. The occurrence of recall bias in pediatric headache: a comparison of questionnaire and diary data. Headache. 2001;41(1):11-20. doi:10.1046/j.1526-4610.2001.111006011.x.
21. Koren G, Maltepe C, Navioz Y, Wolpin J. Recall bias of the symptoms of nausea and vomiting of pregnancy. Am J Obstet Gynecol. 2004;190(2):485-8. doi:10.1016/j.ajog.2003.08.039.
22. Stull DE, Leidy NK, Parasuraman B, Chassany O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions. Curr Med Res Opin. 2009;25(4):929-42. doi:10.1185/03007990902774765.
23. Pellise F, Vidal X, Hernandez A, Cedraschi C, Bago J, Villanueva C. Reliability of retrospective clinical data to evaluate the effectiveness of lumbar fusion in chronic low back pain. Spine (Phila Pa 1976). 2005;30(3):365-8. doi:10.1097/01.brs.0000152096.48237.7c.
24. Kahneman D, Fredrickson BL, Schreiber CA, Redelmeier DA. When more pain is preferred to less: adding a better end. Psychol Sci. 1993;4(6):401-5. doi:10.1111/j.1467-9280.1993.tb00589.x.
25. Gotlin MJ, Kingery MT, Baron SL, McCafferty J, Jazrawi LM, Meislin RJ. Recall bias in retrospective assessment of preoperative patient-reported American Shoulder and Elbow Surgeons scores in arthroscopic rotator cuff repair surgery. Am J Sports Med. 2020;48(6):1471-5. doi:10.1177/0363546520913491.
26. Singh JA. Epidemiology of knee and hip arthroplasty: a systematic review. Open Orthop J. 2011;5:80-5. doi:10.2174/1874325001105010080.
27. Gaertner J, Elsner F, Pollmann-Dahmen K, Radbruch L, Sabatowski R. Electronic pain diary: a randomized crossover study. J Pain Symptom Manage. 2004;28(3):259-67. doi:10.1016/j.jpainsymman.2003.12.017.
28. Ornetti P, Dougados M, Paternotte S, Logeart I, Gossec L. Validation of a numerical rating scale to assess functional impairment in hip and knee osteoarthritis: comparison with the WOMAC function scale. Ann Rheum Dis. 2011;70(5):740-6. doi:10.1136/ard.2010.135483.
29. Roos EM, Lohmander LS. The knee injury and osteoarthritis outcome score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. 2003;1:64. doi:10.1186/1477-7525-1-64.
30. Bellamy N. Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index. Semin Arthritis Rheum. 1989;18(4 Suppl 2):14-17. doi:10.1016/0049-0172(89)90010-3.
31. Turk DC, Okifuji A. Assessment of patients' reporting of pain: an integrated perspective. Lancet. 1999;353(9166):1784-8. doi:10.1016/S0140-6736(99)01309-4.
32. Riddle DL, Kong X, Fitzgerald GK. Psychological health impact on 2-year changes in pain and function in persons with knee pain: data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2011;19(9):1095-101. doi:10.1016/j.joca.2011.06.003.
33. Sharma A, Kudesia P, Shi Q, Gandhi R. Anxiety and depression in patients with osteoarthritis: impact and management challenges. Open Access Rheumatol. 2016;8:103-13. doi:10.2147/OARRR.S93516.