

# Surgeon identification of pain catastrophizing versus the Pain Catastrophizing Scale in orthopedic patients after routine surgical consultation

Marlis T. Sabo, MD, MSc  
Mili Roy, MSc

This study was presented at the Canadian Orthopaedic Association annual meeting, June 20–23, 2018, Victoria, BC, and the American Association of Orthopaedic Surgeons annual meeting, Mar. 12–16, 2019, Las Vegas, Nev.

Accepted Oct. 31, 2018; Early-released June 1, 2019; subject to revision

## Correspondence to:

M. Sabo  
South Health Campus  
Cumming School of Medicine  
University of Calgary  
4448 Front St SE  
Calgary AB T3M 1M4  
marlis.sabo@gmail.com

DOI: 10.1503/cjs.009918

**Background:** A high level of pain catastrophizing has negative influences on outcomes in many surgical disciplines. Our purpose was to determine whether surgeons are able to accurately identify high catastrophizing in orthopedic patients after routine clinical consultation.

**Methods:** In this prospective study, English-literate patients aged 18 years or more were assessed by 1 of 11 orthopedic surgeons. Patients completed the Pain Catastrophizing Scale (PCS), and the surgeon rated each patient as having a high or low level of catastrophizing after the clinical encounter. We calculated accuracy and agreement of surgeon assessment with the PCS at a cut-off score of 30 (score  $\geq$  30 = high level of catastrophizing) and used multivariate testing to determine whether patient age or sex, surgeon experience or subscores of the PCS (rumination, magnification and helplessness) influenced surgeon accuracy.

**Results:** Among 203 patients (109 women and 94 men), the mean PCS score was 18.4 (standard deviation 12.9), with no sex difference and no significant correlation to patient age. Of the 40 patients who scored 30 or more on the PCS, 22 (55%) were not identified as having high levels of catastrophizing by their surgeon. Accuracy was 0.72, and agreement was 0.2. Female patients were more likely than male patients to be identified as high catastrophizing regardless of PCS score (odds ratio 2.0, 95% confidence interval 1.04–4.0).

**Conclusion:** Surgeons were not able to accurately identify patients with high levels of pain catastrophizing during routine initial consultation. In considering which patients may most benefit from interventions to improve coping and reduce catastrophizing, explicitly measuring pain catastrophizing will be required.

**Contexte :** Un niveau élevé de dramatisation face à la douleur a une influence négative sur les résultats dans plusieurs disciplines chirurgicales. Notre objectif était de déterminer si les chirurgiens orthopédistes sont capables d'identifier avec précision un niveau élevé de dramatisation de la douleur chez les patients après une consultation clinique de routine.

**Méthodes :** Au cours de cette étude prospective, des patients capables de communiquer en anglais âgés de 18 ans ou plus ont été évalués par un de 11 chirurgiens orthopédistes. Les patients ont répondu au questionnaire PCS (Pain Catastrophizing Scale), une échelle de mesure de la dramatisation face à la douleur, et après la consultation clinique, le chirurgien assignait à chaque patient un niveau faible ou élevé de dramatisation. Nous avons calculé la justesse et la concordance de l'évaluation du chirurgien avec le score PCS, avec un score seuil de 30 (score  $\geq$  30 = niveau élevé de dramatisation) et utilisé un test multivarié pour déterminer si l'âge ou le sexe, l'expérience du chirurgien ou certaines sous-échelles de la PCS (rumination, amplification et sentiment d'impuissance) influençaient sur le jugement du chirurgien.

**Résultats :** Sur 203 patients (109 femmes et 94 hommes), le score PCS moyen a été de 18,4 (écart-type 12,9), sans différence liée au sexe ni corrélation significative avec l'âge des patients. Sur les 40 patients ayant obtenu un score de 30 ou plus à l'échelle PCS, 22 (55 %) n'ont pas été identifiés comme présentant des niveaux élevés de dramatisation par les chirurgiens. La précision a été de 0,72, et la concordance de 0,2. Les patientes étaient plus susceptibles que les patients d'être

identifiées comme présentant un niveau élevé de dramatisation, indépendamment du score à l'échelle PCS (rapport des cotes 2,0, intervalle de confiance à 95 % 1,04–4,0).

**Conclusion :** Les chirurgiens ont été incapables d'identifier avec précision les patients présentant des niveaux élevés de dramatisation face à la douleur durant la consultation initiale de routine. Pour savoir quels patients sont plus susceptibles de bénéficier d'interventions visant à améliorer leurs capacités d'adaptation et réduire la dramatisation, il faudra des mesures précises du niveau de dramatisation face à la douleur.

**M**ounting evidence suggests that factors other than objective disease severity or technical success of a surgical intervention can influence surgical outcomes. Among these factors is a catastrophizing coping style. High levels of pain catastrophizing leading to an exaggerated negative mental set when pain is experienced or expected is independently associated with higher probability of chronic postoperative pain, lower patient satisfaction following surgery and poorer patient-reported surgical outcomes in various patient populations.<sup>1–7</sup> Components of catastrophizing are rumination (continued focus on pain), magnification (reported severity of pain is higher than expected) and helplessness (perceived inability to positively affect one's situation). Catastrophizing is regarded as a maladaptive coping strategy.<sup>6</sup> Although surgeons may intuitively identify some patients as likely to struggle with their postoperative course owing to factors such as catastrophizing, this factor is not routinely quantified in orthopedic patients.

There are several instruments validated to assess catastrophizing, with the Pain Catastrophizing Scale (PCS) having been validated for multiple musculoskeletal conditions.<sup>6</sup> The PCS measures total catastrophizing and also the components of rumination, magnification and helplessness. Given that researchers have identified pervasive and substantial effects of a catastrophizing coping style on treatment outcome,<sup>5</sup> it is important to determine whether surgeons can identify at-risk patients during their routine consultation practice or whether a catastrophizing coping style needs to be formally quantified during clinical assessment. The purpose of this study was to determine whether surgeons can accurately identify patients as having low or high levels of pain catastrophizing after a routine initial consultation. We hypothesized that surgeons have greater accuracy at the extremes of the measurement scale but have poor diagnostic discrimination for most patients with respect to levels of pain catastrophizing.

## METHODS

This prospective study was conducted at a single academic hospital site from February to March 2017. All English-literate patients over 18 years of age presenting for an initial consultation with a participating surgeon were eligible. A consecutive sample of at least 200 patients was sought. We selected this number owing to the absence of relevant pilot data and to ensure adequate sample size for the multivariate

analysis. Potentially eligible patients were prescreened by M.T.S. to ensure that their surgeon had never met them. Identified patients received a letter of information after registration in the clinic. Implied consent was given if the patient continued to complete the PCS. After completing the PCS, the patient placed the form in a closed box.

Following the consultation, the assessing surgeon (who was blinded to the patient's PCS answers) recorded the patient's sex and age and indicated whether he or she felt that the patient had high or low levels of pain catastrophizing. The surgeon form was also placed in a closed box. Forms were collected and linked by deidentified identification numbers.

Eleven fellowship-trained orthopedic surgeon volunteers (8 men and 3 women) from the site practice group in subspecialty practices including hip and knee reconstruction, adult trauma, shoulder reconstruction, foot and ankle reconstruction, and hand surgery participated. The surgeons had 8–41 years' experience (mean 18.4 [standard deviation (SD) 11.6] yr, median 12 yr). An educational rounds was held before the start of the study, and each surgeon viewed the same educational review module about catastrophizing coping styles and study protocol before the start of recruitment. The rounds highlighted different coping styles and the impact of high levels of pain catastrophizing on surgical outcome, and the review module provided basic information about the definition of catastrophizing and description of the components of rumination, magnification and helplessness, as well as review of the potential clinical consequences of a catastrophizing coping style on surgical outcomes. This also provided a review of key material that had been presented in an education rounds 6 months earlier. Learning from the module was not formally assessed. Clinic staff completed an educational module about correct implementation of the study to minimize problems related to blinding of the surgeons to patient responses.

## Statistical analysis

We performed descriptive statistics regarding demographic characteristics and distribution of catastrophizing scores in the patient population. We assessed the influence of age and sex on levels of catastrophizing using Pearson product correlation (age) and Wilcoxon signed rank (sex) analyses, as the PCS scores did not follow a normal distribution. We analyzed categorical data using  $\chi^2$  tests. Alpha was set at 0.05 for all comparison testing. We calculated

sensitivity, specificity, positive predictive ability and negative predictive ability of surgeons to correctly identify levels of catastrophizing for a cut-off point of 30, with a score of 30 or more indicating a high level of catastrophizing (as per the PCS handbook<sup>6</sup>). We assessed agreement between the surgeon's assessment and the PCS score by calculating a  $\kappa$  statistic.<sup>8</sup> We used multivariate logistic regression analysis to assess associations between patient age and sex and accuracy of the surgeon's assessment, the effect of surgeon experience on accuracy and potential influences of PCS subscores (rumination, magnification, helplessness) on accuracy and true-positive rates. Analyses were performed with Microsoft Excel for Mac (version 14.7.3) and GraphPad for descriptive statistics, and SAS 9.4 (SAS Institute) for the multivariate analyses. The study methodology was reviewed and approved by the local research ethics board, and patient consent was implied by completion of the PCS.

## RESULTS

A total of 203 eligible patients (109 women and 94 men with a mean age of 52.9 [SD 14.9] yr [range 19–80 yr]) presenting over a 5-week period agreed to participate. The number of patients rated by each surgeon ranged from 5 to 43 (mean 18.3 [SD 11.2], median 17), with the 3 most senior surgeons rating 17 of the 203 patients. Patient characteristics by surgeon are shown in Table 1.

The mean PCS score was 18.4 (SD 12.9) (median 16, range 0–52). There was no difference in mean score between men (17.6 [SD 12.6]) and women (19.1 [SD 13.1]) ( $p = 0.4$ ) and no clinically significant correlation between patient age and PCS score ( $r = -0.072$ ,  $p = 0.3$ ). Forty patients (19.7%) (24 women and 16 men) scored 30 or more on the PCS. There was no sex difference in the number who scored 30 or more ( $p = 0.4$ ).

The specificity of surgeon accuracy in identifying high- and low-catastrophizing patients was 0.43, sensitivity was 0.79, positive predictive value was 0.33, negative predictive value was 0.85, and accuracy was 0.72 (Table 1). The false-negative rate was 12% for the overall cohort of patients. Agreement between surgeons' assessment and the PCS score was slight ( $\kappa = 0.2$ ). Of the 40 patients with a PCS score of 30 or more, 22 (55%) were not identified as high catastrophizing by their surgeons. There was no effect of patient sex on accuracy ( $p = 0.7$ ). Female patients were more likely than male patients to be assessed as high catastrophizing (34 [31.2%] v. 18 [19.1%]) (odds ratio 2.0, 95% CI 1.04–4.0).

Multivariate analysis did not show a significant combined effect of patient age and sex on accuracy, false-positive rates or false-negative rates ( $p > 0.05$ ). Surgeon experience was also not associated with improved accuracy ( $p > 0.05$ ). Surgeon sex was not a significant influence on

**Table 1. Patient demographic characteristics, Pain Catastrophizing Scale scores and surgeon accuracy in identifying high- and low-catastrophizing patients, by surgeon**

Surgeon no.	Patient age, yr, mean $\pm$ SD (range)	Patient sex, male:female	Pain Catastrophizing Scale score, mean $\pm$ SD (range)	No. with Pain Catastrophizing Scale score $\geq$ 30/ $<$ 30*	PPV	NPV	Accuracy
1	61.7 $\pm$ 13.3 (27–80)	23:20	21.9 $\pm$ 14.8 (0–52)	12/31	0.27	0.71	0.56
2	54 $\pm$ 19 (29–80)	1:4	12.6 $\pm$ 10.3 (4–30)	1/4	0.50	1.00	0.80
3	53.9 $\pm$ 11.5 (28–80)	11:10	16.0 $\pm$ 12.5 (1–49)	3/18	0.00	0.85	0.81
4	42.7 $\pm$ 14.8 (20–61)	9:8	15.5 $\pm$ 13.0 (1–46)	2/15	0.50	1.00	0.88
5	46.2 $\pm$ 13.9 (21–65)	5:22	20.9 $\pm$ 12.7 (0–49)	8/19	0.60	0.77	0.74
6	49.7 $\pm$ 13 (30–67)	5:8	17.4 $\pm$ 14.9 (0–43)	3/10	1.00	0.83	0.85
7	48.6 $\pm$ 11.7 (27–67)	13:9	15.4 $\pm$ 12.5 (1–47)	3/19	0.13	0.86	0.59
8	50.4 $\pm$ 16.1 (19–69)	10:7	19.2 $\pm$ 12.3 (0–45)	3/14	0.40	0.92	0.76
9	66.7 $\pm$ 8.2 (53–80)	4:3	19 $\pm$ 10.5 (9–35)	2/5	NA	0.71	0.71
10	61.1 $\pm$ 13.6 (33–73)	3:4	13 $\pm$ 7.2 (7–28)	0/7	NA	1.00	1.00
11	51.8 $\pm$ 16.2 (21–80)	10:14	18.7 $\pm$ 11.4 (0–42)	3/21	0.33	1.00	0.75
Overall	52.9 $\pm$ 14.9 (19–80)	94:109	18.4 $\pm$ 12.9 (0–52)	40/163	0.33	0.85	0.72

NA = not applicable; NPV = negative predictive value; PPV = positive predictive value; SD = standard deviation.  
\*A score  $\geq$  30 denotes a high level of catastrophizing.

accuracy ( $p = 0.2$ ). Surgeon accuracy and patient characteristics varied across surgeon practices.

We examined the potential association between the PCS subdomains (rumination, magnification and helplessness) and surgeon accuracy to determine whether surgeons were better at detecting a specific element of catastrophizing. The rumination score was associated with a lower degree of accuracy than the overall PCS score (odds ratio 0.87, 95% confidence interval 0.75–0.10). The other subscores had no influence on accuracy ( $p > 0.05$ ).

## DISCUSSION

Surgeons were not good at recognizing patients with high levels of pain catastrophizing during routine clinical assessment. Accuracy was moderate, but agreement between surgeons' assessments and the PCS score was poor. More than half (56%) of the patients with a PCS score of 30 or more were misidentified by their surgeon as having low levels of catastrophizing. This represents an undesirable situation: the patients with the highest self-reported levels of pain catastrophizing are being missed and are the most at risk for the adverse effects of catastrophizing on treatment outcome. A better approach is needed.

Explaining the difficulties that surgeons have in identifying which patients have a catastrophizing coping style is challenging. Patient affect and presentation in clinic is often dependent on many factors consciously or unconsciously registered by the assessing surgeon. This study was designed to look specifically and only at catastrophizing. However, given the poor accuracy and agreement of the surgeons' assessments compared to the PCS score, there are clearly other factors that influence the overall clinical impression. For example, in this study, although patient sex did not influence accuracy, female patients were more likely than male patients to be rated as having high levels of catastrophizing. This may be the result of differing verbal and nonverbal communications about pain between the sexes.<sup>9</sup> There may also be factors such as personality (introversion or extroversion) or anxiety levels that contribute to the assessment along with an impression of catastrophizing. There is the potential for surgeon bias based on how well patient-reported symptom severity and observed structural disease align.<sup>10</sup> Being incorrectly identified as having high levels of catastrophizing could adversely affect a patient's care if the patient is perceived to benefit less from a given surgical treatment, another undesirable situation. It is important to remember that the presence of higher levels of catastrophizing should not stigmatize a patient. There are interventions that have shown the ability to reduce pain catastrophizing in multiple surgical populations,<sup>11</sup> and making the tools and resources available to such patients offers a potential road to success after surgery.

There are 2 potential approaches to improving the ability to correctly identify patients with high levels of catastroph-

izing. The first would be to explicitly measure the levels with an instrument. This has the advantages of removing the subjectivity of the surgeon's assessment, accurate longitudinal assessment during and after treatment, and the ability to compare effects between studies. The second would be to define what catastrophizing looks like in patients and educate surgeons to better identify the features clinically. To an extent, the ability to code pain behaviours associated with high levels of catastrophizing can be made reliable in a research setting with trained personnel.<sup>12</sup> In principle, this would be superior to asking the surgeon to identify catastrophizing by patient affect and presentation. Ensuring that clinicians are sufficiently educated about physical manifestations of catastrophizing and pain behaviours to match the accuracy of a research environment would be challenging.

The education provided to the surgeons in this study was not in any way intended to create "expert" catastrophizing detectors but, rather, to mimic the "real-world" scenario of a clinician who is aware that catastrophizing is a potential negative prognostic factor and has a basic understanding of what catastrophizing is, similar to what one might take away from a rounds presentation (the most likely scenario for surgeons). Surgeons with only a passing familiarity with the concepts of catastrophizing represent the current standard of care, whereas use of a validated tool represents the gold standard. The fact that the surgeons in our study were aware that their abilities were being assessed may have led to the Hawthorne effect in some cases. If our findings reflect the situation that the surgeons were focused on assessing for catastrophizing (rumination, magnification and helplessness) more than usual, our mediocre results represent the best-case scenario.

Difficulty identifying features of patient distress in the clinical setting has also been shown by Kwon and colleagues,<sup>13</sup> who found that surgeon ratings had poor concordance with patient-reported scores on the Distress Risk Assessment Method but that surgeon impressions of distress influenced the likelihood of the patient's being offered surgical intervention. In that study, surgeons classified the patients into 3 categories: normal, at risk and distressed. The  $\kappa$  coefficient was 0.2, and the authors noted that surgeons tended to underestimate patient distress, as surgeons underestimate catastrophizing in the current study. Further investigation may shed light on how surgeons perceive their patients and how this may influence decision-making in unexpected ways.

Other important findings included no difference in catastrophizing scores between male and female patients, and no significant relation to patient age. Some studies have suggested that female patients have higher levels of catastrophizing than male patients,<sup>2</sup> but others have not corroborated this.<sup>14,15</sup> This may be a function of variation in sample composition. Jacobsen and colleagues<sup>16</sup> proposed that younger patients manifest higher levels of pain catastrophizing than older patients and suggested that this may be due

to less experience coping with painful events. We did not find a significant age relation in our cohort, nor did patient age influence the accuracy of surgeons' assessments. Further work may clarify whether a relation actually exists.

### Strengths and limitations

This study has several strengths. First, it took a broad cross-section of surgeons and patients and measured the ability to recognize patients with high levels of catastrophizing in a real-world situation. Second, it did not differentiate between patients selected or opting for surgical management from those who were not offered surgical treatment or who opted against it, which resulted in a wider range of patient than those usually included in surgical studies of catastrophizing (usually only patients undergoing surgery). This provides a new insight into patients seeking orthopedic care as opposed to orthopedic surgery.

Limitations include the potential for accidental loss of blinding, exclusion of patients who were not literate in English and inability to capture information concerning patients who chose not to participate in the study. The second and third limitations may have introduced bias in terms of patient characteristics but should have had minimal effect on the accuracy of surgeon identification of catastrophizing traits. The number of surgeon raters was relatively small but included the majority of surgeons working in the study environment. Ensuring high levels of protocol adherence and blinding and minimizing logistical issues presented substantial feasibility issues for a multisite trial. Furthermore, the consistency of assessment between surgeons was not examined in this work. To achieve this, we would have had to remove surgeon-patient encounters from the usual clinical environment, which would have affected the applicability of the data. It is also expected that interrater agreement would be poor. This could be examined in further work. A deliberate limitation of this study is that patient diagnosis, disease severity and treatment choice were not recorded for the participating patients. Measuring disease-specific scores for all patients was well beyond the scope of this investigation. Future work will look at the relation between disease severity, treatment choices and catastrophizing levels. Finally, the potential influence of surgeon experience on the ability to correctly identify patients with high levels of catastrophizing must be interpreted with caution. This study revealed a difference in practice pattern between the senior and junior surgeons in that the senior surgeons had a much smaller proportion of new patients versus reassessment consultations. Given that the most senior surgeons rated a small portion of the sample, the strength of inference that may be drawn regarding experience and accuracy is limited. Future work with more evenly matched sample sizes may more conclusively speak to the effect of experience on accuracy.

### CONCLUSION

Although catastrophizing is an important patient-specific factor affecting outcomes of management of orthopedic disease, surgeons had difficulty correctly identifying patients with high levels of catastrophizing. Catastrophizing may require deliberate measurement to avoid missing patients who could benefit from individualized care to optimize their treatment outcomes.

**Affiliations:** From the Cumming School of Medicine, University of Calgary, Calgary, Alta. (Sabo, Roy).

**Competing interests:** None declared.

**Contributors:** M. Sabo designed the study and acquired and analyzed the data, which M. Roy also analyzed. Both authors wrote and reviewed the article and approved the final version for publication.

### References

1. Das De S, Vranceanu AM, Ring DC. Contribution of kinesiophobia and catastrophic thinking to upper-extremity-specific disability. *J Bone Joint Surg Am* 2013;95:76-81.
2. Pavlin DJ, Sullivan MJ, Freund PR, et al. Catastrophizing: a risk factor for postsurgical pain. *Clin J Pain* 2005;21:83-90.
3. Papaioannou M, Skapinakis P, Damigos D, et al. The role of catastrophizing in the prediction of postoperative pain. *Pain Med* 2009; 10:1452-9.
4. Forsythe ME, Dunbar MJ, Hennigar AW, et al. Prospective relation between catastrophizing and residual pain following knee arthroplasty: two-year follow-up. *Pain Res Manag* 2008;13:335-41.
5. Khan RS, Ahmed K, Blakeway E, et al. Catastrophizing: a predictive factor for postoperative pain. *Am J Surg* 2011;201:122-31.
6. Sullivan MJL. *The Pain Catastrophizing Scale: user manual*. Montréal: McGill University; 2009.
7. Burns LC, Ritvo SE, Ferguson MK, et al. Pain catastrophizing as a risk factor for chronic pain after total knee arthroplasty: a systematic review. *J Pain Res* 2015;8:21-32.
8. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360-3.
9. Keogh E. Gender differences in the nonverbal communication of pain: A new direction for sex, gender, and pain research? *Pain* 2014; 155:1927-31.
10. Tait RC, Chibnall JT. Physician judgments of chronic pain patients. *Soc Sci Med* 1997;45:1199-205.
11. Gibson E, Sabo MT. Can pain catastrophizing be changed in surgical patients? A scoping review. *Can J Surg* 2018;61:311-8.
12. Martel MO, Trost Z, Sullivan MJ. The expression of pain behaviors in high catastrophizers: the influence of automatic and controlled processes. *J Pain* 2012;13:808-15.
13. Kwon B, Li L, Kim D. Influence of a patient's psychological distress on surgical decision making. *Spine* 2011;1:161S-25S.
14. Sullivan M, Tanzer M, Stanish W, et al. Psychological determinants of problematic outcomes following total knee arthroplasty. *Pain* 2009;143:123-9.
15. Høvik LH, Winther SB, Foss OA, et al. Preoperative pain catastrophizing and postoperative pain after total knee arthroplasty: a prospective cohort study with one year follow-up. *BMC Musculoskelet Disord* 2016;17:214.
16. Jacobsen PB, Butler RW. Relation of cognitive coping and catastrophizing to acute pain and analgesic use following breast cancer surgery. *J Behav Med* 1996;19:17-29.