# Preoperative cognitive–behavioral therapy for reducing pain catastrophizing and improving pain outcomes after total knee replacement: a randomized clinical trial

Asokumar Buvanendran,[1] Amanda C Sremac,[1] Patricia A Merriman,[1] Craig J Della Valle,[2] John W Burns,[3] Robert J McCarthy [1]

[1]Anesthesiology, Rush Medical College of Rush University, Chicago, Illinois, USA
[2]Orthopedic Surgery, Rush Medical College of Rush University, Chicago, Illinois, USA
[3]Behavioral Sciences, Rush Medical College of Rush University, Chicago, Illinois, USA

**Correspondence to**
Dr Robert J McCarthy, Anesthesiology, Rush Medical College of Rush University, Chicago, Illinois, USA; Robert_J_McCarthy@rush.edu

Received 3 November 2020
Revised 28 December 2020
Accepted 30 December 2020
Published Online First 15 January 2021

## ABSTRACT

**Introduction** Cognitive–behavioral therapy (CBT) can reduce preoperative pain catastrophizing and may improve postsurgical pain outcomes. We hypothesized that CBT would reduce pain catastrophizing more than no-CBT controls and result in improved pain outcomes.

**Methods** The study was a randomized controlled trial of patients undergoing elective total knee arthroplasty between January 2013 and March 2020. In phase 1, the change in pain catastrophizing scores (PCS) among 4-week or 8-week telehealth, 4-week in person and no-CBT sessions was compared in 80 patients with a PCS >16. In phase 2, the proportion of subjects that achieved a 3-month decrease in Western Ontario and McMaster Universities Osteoarthritis (WOMAC) pain subscale >4 following 4-week telehealth CBT with no-CBT controls were compared in 80 subjects.

**Results** In phase 1, 4-week telehealth CBT had the highest completion rate 17/20 (85%), demonstrated an adjusted median reduction in PCS of −9 (95% CI −1 to −14, p<0.01) compared with no-CBT and was non-inferior to 8-week telehealth CBT at a margin of 2 (p=0.02). In phase 2, 29 of 35 (83%) in the 4-week telehealth CBT and 26 of 33 (79%) subjects in the no-CBT demonstrated a decrease in the WOMAC pain subscale >4 at 3 months, difference 4% (95% CI −18% to 26%, p=0.48), despite a median decrease in the PCS for the 4-week CBT and no-CBT group of −6 (−10 to −2, p=0.02).

**Conclusions** Our findings demonstrate that CBT interventions delivered prior to surgery in person or via telehealth can reduced PCS scores; however, this reduction did not lead to improved 3-month pain outcomes.

**Trial registration number** ClinicalTrials.gov (NCT 01772329, registration date 21 January 2013).



**To cite:** Buvanendran A, Sremac AC, Merriman PA, *et al*. *Reg Anesth Pain Med* 2021;**46**:313–321.

## INTRODUCTION

Total knee arthroplasty (TKA) surgeries were performed on more than 700 000 patients in the USA in 2012 and expected to increase by 143% to 565% by 2050.[1] Although most patients experience pain relief within 6–12 weeks after TKA, approximately 8%–43% of patients report persistent pain lasting longer than 3 months, despite clinical and radiological indicators of surgical success.[2–4]

Acute postsurgical pain (72 hours postoperatively) has been shown to be an independent predictor of chronic pain following TKA.[5] In addition, the presence and the magnitude of psychological factors such as pain catastrophizing have been identified as independent predictors for the development of acute and chronic pain following TKA.[6 7] Pain catastrophizing is a multidimensional construct that combines assessment of rumination, magnification and helplessness. Rumination is the tendency for anxious preoccupation with pain and the inability to inhibit pain-related thoughts, magnification assesses the tendency to amplify the significance of pain with respect to global health and helplessness measures the individuals perceived inability to control the pain experience.[8]

Cognitive–behavioral therapy (CBT) may be a potential method of reducing pain catastrophizing and may be an effective intervention for reducing postsurgical pain.[9–11] CBT is an established modality for treating chronic pain; however, no randomized studies have evaluated the effectiveness of different regimes of CBT prior to surgery on the measurement of pain catastrophizing, nor the effect of reducing catastrophizing preoperatively on clinical pain outcomes following TKA.[11] Our first aim was to determine the relative effectiveness of CBT delivered via 4-week or 8-week telehealth, or four in person CBT sessions with no-CBT controls in patients exhibiting high catastrophizing preoperatively. Our second aim was to compare the reduction in pain catastrophizing as well as postoperative pain outcomes in a group of patients receiving 4-week telehealth CBT compared with no-CBT controls. We hypothesized that CBT would reduce pain catastrophizing more than no-CBT controls and would result in improved pain outcomes following TKA.

## METHODS

This manuscript was prepared following Consolidated Standards of Reporting Trials guidelines. The study was a randomized clinical trial of adults that underwent TKA surgery between 31 January 2013 and 30 March 2020. The study was divided in two phases. In phase 1, subjects were randomized into four groups, three CBT groups and a no-CBT group. The objective of phase 1 was to determine the intervention that produced the largest reduction in pain catastrophizing compared with the no-CBT

intervention. Ideally, this group would be superior or non-inferior to the other CBT groups and have a high rate (>80%) of subject completion. In phase 2, subjects were randomized into a 4-week telehealth CBT or no-CBT groups and the magnitude of catastrophizing reduction and pain and functional outcomes at 3 months following TKA were assessed.

Eligible subjects were scheduled to have a tricompartmental primary TKR; 18–85 years of age; with the joint to be replaced represented the primary source of pain. Excluded were patients that were currently receiving antidepressants or CBT; chronic opioid use >4 weeks of ≥10 mg/day of morphine equivalents; a history of opioid abuse; current enrollment in another study; or a planned elective joint replacement procedure during the study period; or an American Society of Anesthesiology physical status classification >3.

Patients were screened for eligibility by examining their medical record. Included in this record was an intake assessment of catastrophizing using the Pain Catastrophizing Scale (PCS) score. The PCS is a 13-item assessment tool asks respondents to rate, using a 5-point Likert scale ranging from 0 (not at all) to 4 (all the time), the degree to which they have certain thoughts and feelings when experiencing pain. Higher scores indicate greater use of catastrophic thinking. The PCS has strong internal consistency ($\alpha$=0.93),[12] concurrent and discriminant validity, and high test–retest reliability (r=0.78).[13] Research suggests that the PCS is responsive to treatment in patients with chronic pain.[14] Eligible subjects were contacted only if the PCS score was >16.

## Phase 1

Subjects were contacted 10–12 weeks prior to surgery. After obtaining informed consent, subjects were again asked to complete the PCS.[8 15] Subject that scored >16 on the PCS score were eligible for inclusion. The PCS score cut-off value was selected because it represented the upper 33rd percentile of subjects screened using the PCS preoperatively in patients undergoing TKA at the investigator's institution.[5] Randomization to study groups was made by study personnel not involved in study assessment. Group assignment was placed in opaque envelops that were opened by research personnel when the patient was deemed eligible. Patients were randomized into four CBT treatment groups: (1) eight weekly CBT sessions with sessions 1 and 8 made in person and sessions 2 through 7 made via telehealth, (2) four weekly CBT sessions; 1 and 4 made in person and 2 and 3 made via telehealth, (3) 4 weekly CBT sessions made in person, (4) or no-CBT sessions. CBT sessions were scheduled so that the last CBT session was completed approximately 1 week prior to surgery.

Patient characteristics, age, sex, race and ethnicity, as well as anxiety and depression and pain and functional status were collected. Pain at rest and with movement of the operated knee were assessed using the Numeric Rating Scale for pain (NRS). The NRS is rated on a 0–10 (11 point) scale, with 0 representing no pain and 10 representing the worse pain imageable. No other anchor point was provided by the assessor. Pain and functions status were assessed using the Western Ontario and McMaster Universities Osteoarthritis (WOMAC 3.1) index. The WOMAC evaluates 3 dimensions: pain (5 items), stiffness (2 items) and physical functioning (17 items). The items were scored on a 11-point NRS scale (0=no interference to 10=extreme interference) with higher scores representing greater interference.[16] The WOMAC pain subscale score contains questions regarding both pain at rest and with movement.

Depression and anxiety were assessed using Becks Depression Inventory (BDI-II) and the State-Trait Anxiety Inventory (STAI). The BDI-II is a 21-item (4-point Likert scale, 0=symptoms absent to 3=severe symptoms) measure of depressive symptoms. A total score range of 0–13 indicates minimal depression, 14–19 indicates mild depression, 20–28 indicates moderate depression and ≥29 indicates severe depression. Patients with a BDI-II ≥29 were excluded from the study.[17] The STAI is a 40-item questionnaire assessing the respondents state (STAI-S) and trait (STAI-T) for anxiety. The STAI-S score reflects a transient elevation in anxiety or anxiety about an event. The STAI-T score assesses the respondent's dominant response pattern with respect to anxious thoughts and feelings. Both the STAI-S and the STAI-T tools use a 4-point Likert scale, for S-anxiety the scale is 1=not at all to 4=very much so, for T-anxiety the scale is 1=almost never to 4=almost always. Range of scores for each subtest is 20–80, the higher score indicating greater anxiety. A cut point of 39–40 has been suggested to detect clinically significant symptoms for the S-Anxiety scale.[18]

The CBT sessions included: a presession retention check, a review of the previous session, an assessment of assigned lessons, a review of the session treatment objectives, a session worksheet, assignment of lessons work on prior to the next session and a postsession retention evaluation. Assignments included instructions to think, do, and write thoughts and reactions to the assigned exercises. The CBT sessions were developed based on a program of cognitive therapy based on program guide developed by Beverly Thorn (see online supplementary appendix).[19] Following the last CBT treatment within 1 week of the scheduled surgery, subjects were re-evaluated using NRS pain assessment at rest and with movement, the PCS, BDI-II, STAI and WOMAC.

## Phase 2

Screening and selection of subjects in phase 2 was the same as in phase 1; however, subjects in phase 2 were randomly assigned into 1 of 2 study groups. The first group received the 4-week telehealth CBT and the second group did not receive CBT. Patient characteristics, age, sex, race and ethnicity, as well as anxiety and depression and pain and functional status were assessed prior to the first CBT session. Pain at rest and with movement of the operated knee over the last 24 hours were assessed using the NRS. Pain and functions status were assessed using the WOMAC 3.1 index. In addition, the Optum SF health survey (SF-36v2) was used to assess the patient's general health status. The SF-36 measures eight domains of health status and principal component analysis has identified two distinct constructs measured by the SF-36: a physical health dimension, and a mental health dimension. Lower scores on the SF-36 (0–100) represent greater interference. Because the physical disability in orthopedic patients can artificially increase the mental component summary, a method specific to orthopedic surgical patients was used to calculate summary scores to mitigate this effect.[20]

Depression and anxiety were assessed using the Patient Health Questionnaire depression scale (PHQ-9) and the Generalized Anxiety Disorder scale (GAD-7). The PHQ-9 is established as a valid diagnostic and severity measure for depressive disorders. Questions are scored on a 0–4 Likert scale. A total score of 0–4 represents no significant depressive symptoms; 5–9, mild symptoms; 10–14, moderate symptoms; 15–19, moderately severe symptoms; and 20–27, severe symptoms.[21] The GAD-7 is a self-report scale for screening, diagnosis and severity assessment of anxiety disorders.[22] Questions are scored on a 0–3 Likert scale. The total score ranges from 0 to 21 and are scored: 0–4 minimal

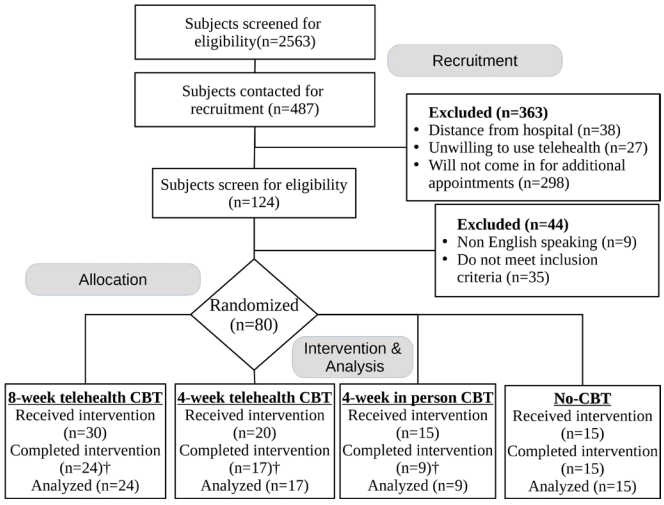


**Figure 1** Consolidated Standards of Reporting Trials flow diagram for subjects enrolled in phase 1. CBT, cognitive–behavioral therapy. †Six subjects withdrew consent prior to completion of the 8-week telehealth session, three subjects withdrew consent prior to completion of the 4-week telehealth session, and six subjects withdrew consent prior to completion of the 4-week in person CBT sessions.

anxiety; 5–9 mild anxiety; 10–14 moderate anxiety; and 15–21 severe anxiety. Pain coping strategies were assessed using the coping strategies questionnaire catastrophizing subscale (CSQ-CAT). The CSQ-CAT is a 7-item assessment using a 0–6 Likert scale. The CSQ-CAT has been demonstrated to reflect the intercorrelation of negative mood and catastrophizing.[23]

Content presented in the CBT sessions were structured as in phase 1. PCS scores were obtained following the final CBT session within 1 week of surgery. Patients received standard anesthesia and operative management. Postoperative analgesia consisted of patient controlled epidural anesthesia or an adductor canal block. Patients were transitioned to oral multimodal analgesia as soon as they were able to take liquids.[5] Outcomes extracted from medical records included: pain assessments, postoperative analgesics, surgical duration, the length of post anesthesia care unit (PACU) stay and length of hospitalization. Pain assessments were made by nursing personnel using the Defense and Veterans Pain Rating Scale (2.0) every 15 min in the post anesthesia recovery room (PACU) and every 4 hour thereafter. Pain burden was calculated as the area under the pain score by time curve using trapezoidal integration and the average pain was calculated by dividing the pain burden by the length of hospitalization.[24 25] Length of stay was the time from admission to actual discharge. Opioid analgesics administered intraoperatively, in the PACU and through hospital discharge were converted to oral milligram morphine equivalents using the conversion tool available from the American Pain Society.[26]

Three-weeks following surgery subjects in the CBT group received a single CBT session to reinforce the lessons taught in the preoperative sessions. At 3 months postoperatively, pain, psychological, physical, functional and overall health assessment were repeated.

**Table 1** Clinical characteristics, psychological, pain and functional assessment of subjects prior to cognitive–behavioral therapy intervention

| | Intervention | | | No intervention | |
|---|---|---|---|---|---|
| Subjects randomized to group (n) | 8-week telehealth 30 | 4-week telehealth 20 | 4-week in-person 15 | 15 | Effect size (CI)* |
| Median age (IQR) in years | 58 (54 to 66) | 58 (52 to 62) | 62 (60 to 69) | 65 (50 to 74) | 0.06 (–0.01 to 0.23) |
| Female sex n (%) | 19 (65) | 16 (80) | 4 (27) | 7 (47) | 0.38 (0.10 to 0.57) |
| Race n (%) | | | | | |
| Caucasian | 24 (80) | 11 (55) | 6 (40) | 7 (47) | |
| African American | 5 (17) | 9 (45) | 8 (53) | 8 (53) | 0.29 (0.00 to 0.40) |
| Asian | – | – | 1 (7) | – | |
| Missing | 1 (3) | – | – | – | |
| Median BDI (IQR) on a 0–63 scale | 13 (8 to 19) | 14 (10 to 19) | 9 (4 to 18) | 13 (6 to 19) | –0.01 (–0.03 to 0.16) |
| Missing (n) | 2 | 2 | 1 | 1 | |
| Median STAI-S (IQR) on a 0–80 scale | 35 (26 to 51) | 44 (33 to 54) | 38 (31 to 47) | 44 (33 to 49) | –0.01 (–0.04 to 0.14) |
| Missing (n) | 2 | 2 | 1 | 1 | |
| Median STAI-T(IQR) on a 0–80 scale | 38 (26 to 48) | 36 (32 to 49) | 36 (29 to 41) | 42 (25 to 50) | –0.03 (–0.04 to 0.1) |
| Missing (n) | 2 | 2 | 1 | 1 | |
| Median PCS score (IQR) on a 0–52 scale | 32 (22 to 37) | 28 (21 to 37) | 25 (18 to 36) | 24 (21 to 34) | –0.01 (–0.03 to 0.14) |
| Missing (n) | 1 | 1 | 0 | 0 | |
| Median NRS knee pain (IQR) on a 0–10 scale | | | | | |
| At rest | 3 (1 to 5) | 5 (3 to 7) | 5 (4 to 7) | 6 (2 to 8) | 0.04 (–0.02 to 0.23) |
| With movement | 7 (6 to 8) | 8 (6 to 8) | 7 (6 to 9) | 8 (6 to 9) | –0.02 (–0.04 to 0.13) |
| Missing (n) | 2 | 3 | 1 | 1 | |
| Median WOMAC (IRQ) score | | | | | |
| Pain on a 0–50 scale | 27 (23 to 38) | 31 (23 to 38) | 37 (28 to 41) | 34 (24 to 42) | 0.01 (–0.03 to 0.19) |
| Stiffness on a 0–20 scale | 14 (9 to 16) | 14 (12 to 19) | 13 (10 to 17) | 16 (12 to 19) | –0.01 (–0.03 to 0.14) |
| Physical function on a 0–170 scale | 91 (74 to 110) | 88 (49 to 136) | 96 (66 to 117) | 85 (63 to 130) | –0.03 (–0.04 to 0.11) |
| Total score on a 0–240 scale | 133 (109 to 158) | 141 (90 to 187) | 148 (106 to 164) | 117 (113 to 189) | –0.03 (–0.04 to 0.10) |
| Missing (n) | 2 | 3 | 1 | 1 | |

Data presented as median (first to third quartile) or n (% of column).
*Effect size for continuous data based on Kruskal-Wallis H statistic and given as eta-squared [H] with 95% CIs. Effect size for contingency table analysis given as Cramer's V with 95% CIs.
BDI, Beck's Depression Index; STAI-S, State-Trait Anxiety Inventory State; STAI-T, State-Trait Anxiety Index Trait; PCS, Pain Catastrophizing Scale; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

**Table 2** Pain, functional and psychological assessment of subjects following cognitive–behavioral therapy intervention

| | Intervention | | | | |
|---|---|---|---|---|---|
| | 8-week telemedicine | 4-week telemedicine | 4-week in-person | No intervention | |
| Subjects randomized to group (n) | 30 | 20 | 15 | 15 | P value |
| Median BDI (IQR) on a 0–63 scale | 11 (6 to 16) | 10 (8 to 19) | 8 (5 to 15) | 11 (5 to 27) | 0.5 |
| Post–preintervention | −5 (−8 to 2) | −1 (−3 to 1) | 0 (−3 to 4) | −1 (−4 to 4) | 0.33 |
| Missing | 6 | 3 | 6 | 2 | |
| Median STAI-S(IQR) on a 0–80 scale | 40 (27 to 50) | 34 (29 to 41) | 37 (26 to 41) | 45 (27 to 50) | 0.68 |
| Median (IQR) post–preintervention | −1 (−7 to 7) | −4 (−8 to 4) | 1 (−6 to 3) | 0 (−5 to 2) | 0.87 |
| | 6 | 3 | 6 | 2 | |
| Median STAI-T(IQR) on a 0–80 scale | 38 (25 to 45) | 37 (31 to 43) | 37 (31 to 40) | 42 (24 to 52) | 0.88 |
| Median (IQR) post–preintervention | −3 (−10 to 2) | 1 (−1 to 2) | 1 (−4 to 8) | −1 (−5 to 4) | 0.27 |
| Missing | 6 | 3 | 6 | 2 | |
| Median PCS score (IQR) on a 0–52 scale | 19 (12 to 27) | 19 (9 to 25) | 22 (11 to 28) | 24 (16 to 31) | 0.4 |
| Unadjusted median (IQR) post–preintervention | −11 (−18 to −8) | −10 (−18 to −3) | −7 (−17 to −2) | −3 (−5 to 0) | 0.02 |
| Adjusted median (IQR) post–preintervention* | −12 (−19 to −9) | −16 (−20 to −9) | −12 (−15 to −8) | −7 (−12 to 0) | 0.02 |
| Missing | 6 | 3 | 6 | 2 | |
| NRS pain of the knee on a 0–10 scale | | | | | |
| Median NRS (IQR) at rest | 4 (2 to 6) | 4 (2 to 6) | 4 (3 to 6) | 5 (4 to 6) | 0.65 |
| Median (IQR) post–preintervention | 0 (−1 to 1) | −1 (−3 to 1) | −1 (−2 to 1) | 1 (−2 to 1) | 0.64 |
| Median NRS (IQR) with movement | 7 (5 to 8) | 7 (5 to 10) | 6 (5 to 7) | 8 (3 to 10) | 0.47 |
| Median (IQR) post–preintervention | 0 (−1 to 1) | 0 (−2 to 20 | −2 (−4 to 0) | 0 (−2 to 2) | 0.1 |
| Missing | 6 | 3 | 6 | 2 | |
| Median WOMAC (IQR) scores | | | | | |
| Pain subscale on a 0–50 scale | 26 (18 to 38) | 26 (13 to 42) | 28 (10 to 43) | 26 (12 to 36) | 0.29 |
| Median (IQR) post–preintervention | −2 (−14 to 5) | 0 (−4 to 5) | −2 (−14 to 1) | −1 (−5 to 3) | 0.7 |
| Stiffness subscale on a 0–20 scale | 13 (7 to 16) | 12 (10 to 17) | 14 (5 to 19) | 14 (5 to 18) | 0.52 |
| Median (IQR) post–preintervention | −2 (−4 to 1) | −2 (−3 to 2) | 0 (−4 to 1) | 0 (−3 to 1) | 0.98 |
| Physical function on a 0–170 scale | 88 (58 to 118) | 74 (33 to 145) | 78 (33 to 145) | 87 (33 to 106) | 0.97 |
| Median (IQR) post–preintervention | −2 (−23 to 10) | −1 (−12 to 11) | −12 (−20 to 2) | 6 (−36 to 14) | 0.7 |
| Total score on a 0–240 scale | 129 (91 to 169) | 108 (69 to 170) | 110 (59 to 202) | 130 (47 to 152) | 0.9 |
| Median (IQR) post–preintervention | −9 (−30 to 16) | −6 (−27 to 54) | −14 (−42 to 1) | 1 (−39 to 14) | 0.49 |
| Missing | 6 | 3 | 6 | 2 | |

Data presented as median (first to third quartile).
*Adjusted differences determined using a generalized estimating equation model with group, sex and race as fixed effects and age, STAI-T and baseline PCS as random factors.
BDI, Beck's Depression Index; STAI-S, State-Trait Anxiety Inventory State; STAI-T, State Trait Anxiety Index Trait; PCS, Pain Catastrophizing Scale; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

## Statistical analysis

The primary outcome for phase 1 was the difference in differences in the PCS score defined as the post intervention score minus the preintervention score. Difference in differences were examined for deviation from normality by examining q–q plots and the Shapiro-Wilks test. Imbalance in group allocations were assessed using Cramer's V (>0.1) for dichotomous and ordinal data and eta-squared [H] (>0.01) for continuous data. Eta-squared [H] was calculated from the Kruskal-Wallis H statistic using the formulae eta-squared [H] = (H − k+1)/(n − k); where H is the value obtained in the Kruskal-Wallis test; k is the number of groups; n is the total number of observations. Unadjusted difference in differences in PCS scores were compared among groups using the Kruskal-Wallis test with post-hoc comparisons made using Dunn's test with multiple comparisons controlled for using the Holm-Šidák method. A generalized estimating equation fitted by maximum likelihood was used to adjusted differences in differences in PCS scores for imbalances in group allocation in patient demographics, psychological assessments of anxiety and depression as well as baseline PCS scores. Sex, race and group were evaluated as fixed factors and the STAI-T and baseline PCS scores as random factors. Multiple comparisons in the primary analysis were controlled for using the Holm-Šidák method. Non-inferiority io the reduction in the PCS score between the 8-week and 4-week telehealth groups was evaluated using the Wilcoxon rank-sum location difference test for Non-Inferiority assuming a non-inferiority margin of 2.

Secondary outcomes include the difference in differences in NRS pain, BDI-II, STAI-S, STAI-T and WOMAC scores. Postintervention to preintervention measures of NRS pain, BDI-II, STAI-S, STAI-T and WOMAC scores were compared within groups using the Wilcoxon signed-rank test. Post to preintervention differences were compared among groups using the Kruskal-Wallis test. A $p<0.05$ was required to reject the null hypothesis.

The sample size for phase 1 was estimated to be 20 per group and a total sample of 80. This design would achieve any pair power of 0.804 for comparison of any group versus the no-CBT group assuming a decrease in the PCS score by −10 to −6 and −6 in the three CBT groups compared with 0 in the no-CBT group, with a SD across groups of 10. The family-wise error rate was set at 0.05. The changes in PCS scores following CBT were selected within the range 0 to −19 consistent with of interventions for changing PCS scores in surgical patients.[9]

The primary outcome for phase 2 was the difference in the proportion of subjects that achieved a 3-month decrease in WOMAC pain subscale >4. The primary outcome was compared between groups using the proportions test and the relative risk of an increase in the proportion of patients with

a WOMAC pain subscale >4 compared with no intervention was determined using binary logistic regression. Imbalances in baseline characteristics were assessed by examining the mean standardized difference and 95% CI of the standardized difference. Standardized differences in baseline characteristics were determined using Hedges' g for continuous variables and Cliff's delta for ordinal or dichotomous data.

Secondary outcomes included the difference in differences in the PCS scores. Differences in PCS scores were compared between the groups using the Mann-Whitney U test. An exploratory analysis was performed to examine difference in differences of the three PCS subscales, rumination, magnification and helplessness. Additional exploratory outcomes include the difference in NRS pain, PHQ-9, GAD-7, WOMAC scores, SF-36 PCS and MCS score, PCS scores and CSQ-CAT scores at 3 months postoperatively.

Primary and secondary outcomes that were interval data were examined for deviation from normality by examining q–q plots and the Shapiro-Wilks test. Three months postoperative to preintervention differences in NRS pain, functional and psychological tests were compared using the Wilcoxon signed-rank test. Sex and race were compared using a $\chi^2$ statistic. CIs for differences in proportions were calculated using the Pearson-Klopper method. Differences in medians and 95% CI interval of the median difference were calculated using a 10 000-sample bootstrap. A $p<0.05$ was required to reject the null hypothesis. Statistical analysis was performed using RStudio V.1.3.1093 (Integrated Development for R. RStudio, Boston, Massachusetts; URL: http://www.rstudio.com/) and R V.4.0.3, release date 10 October 2020 (The R Foundation for Statistical Computing, Vienna, Austria).

The sample size for phase 2 was estimated based on the study of Riddle *et al*[27] that demonstrated an OR of 6.0 for a dichotomized PCS score at 16 for predicting an increase in the proportion of patients that did not demonstrated a decrease in the WOMAC pain subscale >4 at 6 months following a TKA. The OR of 6.04 represent a common language effect size of 0.65.[28] Based on this common effect size group samples of 38 achieves 81% power to reject the null hypothesis of a zero-effect size and a significance level of 0.05 using a two-sided z-test. Sample size calculations were made using PASS 2008, release date 27 January 2011 (Power Analysis and Sample Size Software (2008). NCSS, LLC. Kaysville, Utah, USA, ncss.com/software/pass).

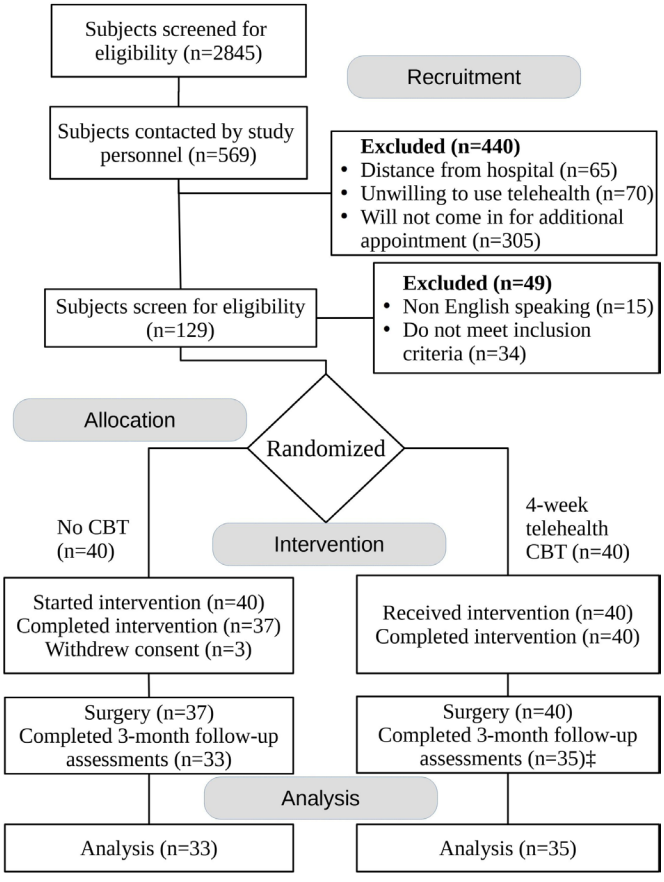


**Figure 2** Consolidated Standards of Reporting Trials flow diagram for subjects enrolled in phase 2. CBT, cognitive–behavioral therapy. †Three subjects withdrew consent prior to presurgery assessments. ‡Four subjects lost to follow-up at 3 months in the no CBT group and five subjects lost to follow-up at 3 months in the 4-week CBT group.

## RESULTS

### Phase 1

The flow of subjects in the study is shown in figure 1. Due to an error in group assignment, 30 subjects were randomized to the 8-week telehealth CBT, 20 to the 4-week telehealth CBT and 15 to the 4 weeks in person and no-CBT groups. The completion rates for the 3 CBT intervention groups were 24 of 30 (80%), 17 of 20 (85%), and 9 of 15 (60%) for the 8 weeks, 4 weeks telehealth and the 4 weeks in person CBT sessions, respectively.

Preintervention clinical characteristics and psychological assessment are shown in table 1. A greater percentage of women were randomized into 4-week telehealth and the fraction of Caucasian participants was greater in the 8-week telehealth compared with the no intervention group. Post intervention, there were no differences in postintervention to preintervention measures of the BDI, STAI-S, STAI-T, NRS pain or WOMAC scores (table 2). PCS scores were decreased from preintervention values in the 8-week and 4-week telehealth as well as in the 4 weeks in person groups, but not in the no-CBT group. Unadjusted median difference in differences in the PCS scores for the 8-week and 4-week telehealth and 4-week in-person compared with no-CBT group were −8 (95% CI −5 to −15, $p<0.01$), –7 (95% CI −1 to −15, $p=0.02$) and −4 (95% CI −19 to 2, $p=0.54$), respectively. Adjusted PCS scores for the 4-week in-person, the 8-week and 4-week telehealth groups compared with no-CBT were −5 (95% CI −12 to −0, $p=0.03$), –9 (95% CI −15 to −1, $p<0.01$) and −5 (95% CI −19 to 1, $p=0.17$), respectively. The adjusted difference between the 8-week and 4-week telehealth groups was −4 (95% CI −7 to 3, $p=0.69$). The 4-week telehealth group was non-inferior to the 8-week telehealth group at a non-inferiority margin of 2 ($p=0.02$). Based on the high completion rate and the observed non-inferiority, the 4-week telehealth intervention was selected for phase 2.

### Phase 2

The flow of subjects is shown in figure 2. Forty subjects were randomized to each group with all CBT participants receiving the intervention. Two subjects in the no-CBT group withdrew consent at the time of preoperative assessment and one subject did not undergo a TKA. Preintervention standardized differences between were negligible (effect size <0.1) for sex, body mass index, race, ethnicity, PCS, preoperative opioid use and CSQ-CAT scores (table 3). Preintervention difference in psychological, functional or pain assessments were small (effect size ≤0.5).

Differences in differences in total PCS scores and in the rumination, magnification and helplessness subscales following

**Table 3** Clinical, psychological, pain and physical function of study groups prior to surgery

| | Cognitive–behavioral therapy | No intervention | |
|---|---|---|---|
| **Number of subjects** | **40** | **37** | **Standardized difference (95% CI)*** |
| Median age (IQR) in years | 66 (58 to 70) | 62 (55 to 73) | −0.14 (−0.58 to 0.31) |
| Sex n (%) | | | |
| Male | 14 (35) | 12 (32) | 0.03 (−0.19 to 0.24) |
| Female | 26 (65) | 25 (68) | |
| Median BMI (IQR) in kg/m$^2$ | 30 (26 to 37) | 31 (28 to 36) | 0.07 (−0.37 to 0.52) |
| Race n (%) | | | |
| White | 28 (70) | 24 (65) | |
| African American | 11 (28) | 12 (32) | 0.05 (−0.16 to 0.26) |
| Asian | 1 (2) | 1 (3) | |
| Ethnicity n(%) | | | |
| Hispanic | 1 (2) | 2 (5) | 0.03 (−0.06 to 0.12) |
| Median GAD-7 (IQR) (0–21) | 4 (2 to 9) | 3 (1 to 6) | −0.24 (−0.69 to 0.21) |
| Median PHQ-9 (IQR) on a 0–27 scale | 6 (3 to 11) | 5 (3 to 8) | −0.17 (−0.63 to -0.27) |
| Median PCS (IQR) on a 0–52 scale | | | |
| Preintervention | 27 (20 to 36) | 28 (21 to 35) | −0.01 (−0.23 to 0.44) |
| Presurgery | 18 (11 to 25) | 21 (16 to 31) | 0.48 (0.02 to 0.93) |
| Median WOMAC (IRQ) score | | | |
| Pain on a 0–50 scale | 23 (16 to 34) | 27 (16 to 38) | 0.21 (−0.25 to 0.67) |
| Stiffness on a 0–20 scale | 12 (7 to 15) | 13 (6 to 16) | −0.01 (−0.46 to 0.46) |
| Physical function on a 0–170 scale | 86 (60 to 108) | 107 (55 to 118) | 0.15 (−0.31 to 0.62) |
| Total score on a 0–240 scale | 125 (84 to 164) | 141 (73 to 164) | 0.16 (−0.30 to 0.62) |
| Median SF-36 domains (IQR) | | | |
| Physical health on a 0–100 scale | 35 (30 to 43) | 34 (26 to 39) | −0.37 (−0.84 to 0.08) |
| Mental health on a 0–100 scale | 47 (40 to 52) | 47 (40 to 51) | −0.11 (−0.57 to 0.35) |
| Median NRS pain (IQR) in the knee on a 0–10 scale | | | |
| Median NRS (IQR) at rest | 2 (1 to 5) | 3 (1 to 6) | 0.24 (−0.22 to 0.71) |
| Median NRS (IQR) with movement | 7 (4 to 9) | 6 (4 to 8) | 0.14 (−0.32 to 0.61) |
| Median preoperative opioid use (IQR) in MME | 15 (0 to 15) | 15 (0 to 15) | −0.02 (−0.47 to 0.43) |
| Median CSQ-CAT (IQR) on a 0–36 scale | 8 (3 to 15) | 7 (3 to 15) | 0.01 (−0.45 to 0.47) |

*Standardized difference reported as Hedges G for continuous variables and Cliff's delta for ordinal and dichotomous variables.
.BMI, body mass index; CSQ-CAT, coping strategies questionnaire-catastrophizing subscale; GAD-7, General Anxiety Disorder 7 question survey; IRQ, first to third quartile; MCS, mental component summary of the SF-36; MME, oral morphine equivalents; PCS, physical component summary of the SF-36; PHQ-9, Patient Health Questionnaire 9 question survey; SF-36, Optum SF Short Form Health Survey; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

**Table 4** Difference in differences in pain catastrophizing scale scores for cognitive therapy compared with no intervention

| | Presurgery minus preintervention difference* | P value | Unadjusted differences between groups (95% CI)† | P value |
|---|---|---|---|---|
| **Pain Catastrophizing Scale total score (0–52 scale)** | | | | |
| 4-week telehealth CBT | −9 (−16 to −3) | <0.01 | −6 (−10 to −2) | 0.02 |
| No intervention | −3 (−8 to 0) | <0.01 | | |
| **Pain Catastrophizing Scale rumination subscore (0–16 scale)** | | | | |
| 4-week telehealth CBT | −3 (−6 to −1) | <0.01 | −3 (−4 to −1) | <0.01 |
| No intervention | 0 (-3 to 1) | 0.02 | | |
| **Pain Catastrophizing Scale magnification sub-score (0–12 scale)** | | | | |
| 4-week telehealth CBT | −1 (−3 to 0) | <0.01 | 0 (−2 to 1) | 0.37 |
| No intervention | −1 (−2 to 0) | <0.01 | | |
| **Pain Catastrophizing Scale helplessness sub-score (0–24 scale)** | | | | |
| 4-week telehealth CBT | −4 (−8 to 0) | <0.01 | −1 (−6 to 0) | 0.05 |
| No intervention | −1 (−6 to 0) | <0.01 | | |

Data for presurgery minus preintervention pain catastrophizing scores reported as median (first to third quartile).
Differences between groups reported as median difference and 95% CIs.
*Comparison made using the Wilcoxon signed-rank test.
†Comparisons made using the Mann-Whitney U-test.
CBT, cognitive–behavioral therapy.

**Table 5** Postsurgical outcomes

| | Cognitive–behavioral therapy | No intervention | Unadjusted differences between groups | |
|---|---|---|---|---|
| **Number of subjects** | **40** | **37** | **(95% CI)*** | **P value** |
| **Hospital course** | | | | |
| Median duration of surgery (IQR) in minutes | 79 (71 to 94) | 80 (70 to 93) | −1 (−12 to 6) | 0.96 |
| Median PACU time (IQR) in minutes | 128 (94 to 171) | 110 (82 to 174) | 18 (−26 to 50) | 0.29 |
| Median pain burden (IQR) in score*hours | | | | |
| 0–12 hours | 10.9 (1.9 to 23.0) | 18.8 (7.3 to 31.7) 39.9 (19.8 to 62.2) | −7.9 (−19.2 to 3.5) | 0.10 |
| 12–24 hours | 25.4 (15.8 to 48.3) | 37.9 (22.1 to 90.0) | −14.5 (−31.7 to 3.2) | 0.22 |
| 24–48 hours | 58.4 (18.8 to 134.3) | 40.0 (27.1 to 129.5) | 20.4 (−14.6 to 59.9) | 0.38 |
| 48–72 hours | 34.9 (17.9 to 84.1) | | −5.1 (−89.2 to 30.6) | 0.31 |
| Median average pain score (IQR) on a 0–10 scale | 2.9 (1.4 to 3.9) | 3.2 (1.7 to 4.3) | −0.3 (−1.2 to 0.9) | 0.67 |
| Acute postoperative pain>4 n (%) | 9 (22) | 10 (27) | −5 (−27 to 17) | 0.79 |
| Median opioid analgesia (IQR) in MME | 300 (214 to 401) | 368 (210 to 453) | −68 (−124 to 43) | 0.42 |
| Median length of stay (IQR) in hours | 36 (28 to 55) | 29 (26 to 53) | 7 (−3 to 23) | 0.44 |
| **Three-month follow-up** | | | | |
| Number of subjects | 35 | 33 | | |
| Median average pain score (IQR) on a 0–10 scale | | | | |
| Pain at rest | 1 (0 to 3) | 1 (0 to 2) | 0 (−1 to 2) | 0.43 |
| Pain with movement | 2 (1 to 4) | 2 (1 to 4) | 0 (−1 to 1) | 0.91 |
| Pain at movement>4 on a 0–10 scale | 9 (26) | 8 (24) | 2 (−22 to 26) | 0.79 |
| Median GAD-7 (IQR) (0–21) | 2 (0 to 6) | 2 (0 to 4) | 0 (−2 to 3) | 0.43 |
| Median PHQ-9 (IQR) on a 0–27 scale | 4 (2 to 6) | 3 (1 to 6) | 1 (−2 to 3) | 0.24 |
| Median WOMAC (IQR) | | | | |
| Pain on a 0–50 scale | 6 (3 to 13) | 8 (5 to 14) | −2 (−5 to 2) | 0.43 |
| Stiffness on a 0–20 scale | 5 (2 to 9) | 5 (3 to 10) | 0 (−3 to 2) | 0.57 |
| Physical function on a 0–170 scale | 29 (17 to 50) | 30 (17 to 47) | −1 (−16 to 11) | 0.94 |
| Total score on a 0–240 scale | 40 (23 to 73) | 46 (25 to 69) | −6 (−23 to 15) | 0.65 |
| Decrease in WOMAC pain subscale>4, n (%) | 29 (83) | 26 (79) | 4 (−18 to 26) | 0.76 |
| Median SF-36 (IQR) | | | | |
| Physical health on a 0–100 scale | 46 (41 to 51) | 43 (38 to 51) | 3 (−5 to 7) | 0.48 |
| Mental health on a 0–100 scale | 52 (46 to 58) | 51 (47 to 58) | 1 (−6 to 5) | 0.71 |
| Median PCS (IQR) on a 0–52 scale | 3 (0 to 8) | 2 (0 to 7) | 1 (−2 to 5) | 0.47 |
| Median CSQ-CAT (IQR) on a 0–36 scale | 2 (0 to 5) | 2 (0 to 3) | 0 (−1 to 2) | 0.55 |

*Difference reported as median difference and 95% CI determined by bootstrapping or difference in percentages with 95% CIs determined using the Clopper-Pearson method.
CSQ-CAT, coping strategies questionnaire-catastrophizing subscale; GAD-7, General Anxiety Disorder 7 question survey; IRQ, first to third quartile; MME, oral morphine equivalents; PACU, postanesthesia recovery unit; PCS, pain catastrophizing scale; PHQ-9, Patient Health Questionnaire 9 question survey; SF-36, Optum SF Short Form Health Survey; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

4-week telehealth and no-CBT are shown in table 4. Differences in the overall PCS score as well as the PCS rumination subscale were greater in the 4-week telehealth group. The PCS magnification and helplessness subscales were not different.

There were no differences in surgical duration or the hospital course (table 5). Five patients were lost to follow-up at 3 months in the CBT and four in the no-CBT group. In the 4-week telehealth CBT, 29 of 35 (83%) subjects compared with 26 of 33 (79%) subjects in the no-CBT demonstrated a decrease in the WOMAC pain subscale >4 at 3 months, difference 4% (95% CI −18% to 26%, p=0.48; table 5). The relative risk of an increase in the WOMAC pain subscale at 3 months in subjects that received 4-week CBT compared with the no-CBT group was 1.30 (95% CI 0.38 to 4.37, p=0.67). There was significant improvement in pain and functional outcomes compared with the preoperative values, but no differences between groups. Psychological assessments were also improved at 3 months, but were not different between groups.

## DISCUSSION

The important findings of this study are that 4-week CBT therapy delivered by telehealth reduced catastrophizing preoperatively; however, the reduction in catastrophizing did not influence pain, physical or psychological outcomes at 3 months following TKA. A second finding of our study was that 4-week and 8-week telehealth sessions appear to confer similar reductions in catastrophizing. In addition, our findings demonstrate that CBT reductions in catastrophizing primarily affect rumination or focused attention of the symptoms of the pain and on its courses and consequence, and not the extent that patients feel helpless or on the extent that they magnify the consequences of the pain. Taken together, our findings do not support the large-scale application of CBT as a sole therapeutic intervention to improve outcomes even in high catastrophizing patients.

Prior studies have suggested an association between increased PCS scores and poor outcomes following TKA. Riddle *et al* found that a PCS score dichotomized at a value of 16 was associated with reduced improvement in WOMAC pain scores and functional outcomes post TKA.[27] In a review of pain catastrophizing on outcomes following TKA, Burns *et al* found that in five of six studies, catastrophizing was associated with at least one pain outcome.[6] In a prior study from our group, we found that increased PCA scores were associated with acute pain (72 hours

postoperatively).[6] These studies concluded that pain catastrophizing is a potentially modifiable response, and to the extent that it has a causal relationship with pain outcomes following TKA, interventions aimed at reducing catastrophizing might influence pain outcomes.[6]

Two prior studies have investigated the use of CBT sessions 2 weeks pre and for up to 6 weeks postoperatively in patients undergoing TKA. Riddle *et al* in a multicentered study of 402 patients randomized subjects with PCS scores>16 to receive 8 CBT sessions, a single arthritis education session or standard of care.[29] No difference in pain or functional outcomes was seen at 2, 6 or 12 months postoperatively. Birch *et al* examined a presurgery and postsurgery CBT intervention in high catastrophizing patients (PCS >22), allocating patients to two groups, one receiving seven CBT sessions, three preoperative and four postoperative, or standard of care.[30] Unlike the current study, neither of these studies delivered the majority of the CBT intervention prior to surgery, nor did they report the change in catastrophizing that occurred from the CBT interventions independent of the surgery. Our findings concur with studies suggesting that CBT interventions alone are not likely sufficient to improve post TKA pain or functional outcomes. In addition, similar to the findings of these studies, we found that catastrophizing scores were substantially reduced postoperatively in association with the decrease in pain and improvement in function outcomes.

In a systematic review of studies examining interventions to reduce catastrophizing, Gibson and Sabo found that CBT interventions reduced PCS scores from 0 to −19.[9] The effects were assessed across multiple study types and was not adjusted for confounding variables. They also reported that the mean clinically important reduction in the PCS score was −9.1. In our study, the unadjusted change in PCS values was on the order of −9 to −11 for 4-week or 8-week telehealth sessions and reduced to −5 after adjusting for baseline PCS, PHQ-8 and GAD-7 scores. Taken together, these data suggest that the effect of CBT sessions on PCS sore reductions is below the threshold needed to demonstrate clinically important differences in outcomes.

The results of our study should only be interpreted in the context of its limitations. We studied only patients undergoing primary TKA that did not expect to undergo an additional orthopedic procedure within the follow-up period to improve the fidelity of our follow-up assessment. We excluded patients that were chronically taking opioids prior to surgery and those with clinically severe depression and anxiety, factors that have been shown to influence catastrophizing as well as pain and functional outcomes following TKA. We only examined 4-week and 8-week CBT interventions and longer interventions or ones that used different behavioral modification methods may have had a greater effect on reducing catastrophizing. Due to an error in randomization, we were unable to assess if in person session would have reduced catastrophizing to a greater extent that those that were administered via telehealth; although other studies that employed CBT used telephone delivered interventions. Finally, although our sample was based on prior studies examining the influence of catastrophizing on outcomes following TKA, given the limited effect of the intervention on PCS scores, it is likely that it was underpowered to detect small differences in outcomes.

## CONCLUSION

Our findings demonstrate that a 4-week CBT intervention prior to surgery can reduced catastrophizing as assessed using the PCS; however, this reduction alone is may be insufficient to reduce pain outcomes. We also showed that increasing the CBT sessions from 4 to 8 does not likely increase the effectiveness of the intervention. Our results suggest that CBT reductions in catastrophizing primarily affect rumination or focused attention of the symptoms of the pain and its courses and consequence, and not on the extent that patients feel helpless or on the extent that they magnify the consequences of the pain. Taken together, our findings do not support the large-scale application of CBT as a sole intervention to improve outcomes even in high catastrophizing patients.

**Contributors** AB was involved in the conception or design of the study, drafting the article, critical revision of the article and final approval of the version to be published. AS was involved in the acquisition of the data, drafting the article, critical revision of the article and final approval of the version to be published. PM was involved in the preforming CBT sessions, acquisition of the data, drafting the article, critical revision of the article and final approval of the version to be published. CJDV was involved in data interpretation, drafting the article, critical revision of the article and final approval of the version to be published. JB was involved in conception or design of the study and data interpretation, drafting the article, critical revision of the article and final approval of the version to be published. RJM was involved in the oversight of the study, acquisition and analysis of the data, drafting the article, critical revision of the article and final approval of the version to be published.

**Funding** This study was supported in part by an investigator-initiated unrestricted grant from Pfizer (WI173357), to Rush University Medical Center with Principal investigator as Asokumar Buvanendran of the Department of Anesthesiology.

**Disclaimer** The funding organization had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This study was approved by the Institutional Review Board of Rush University (12031901-IRB03).

**Data availability statement** Data are available upon reasonable request. Deidentified participant data may be available from RJM (ORCID id:0000-0002-0966-5311), the corresponding author, on request and execution of a data use agreement.

**ORCID iD**

Robert J McCarthy http://orcid.org/0000-0002-0966-5311

## REFERENCES

1 Inacio MCS, Paxton EW, Graves SE, *et al*. Projected increase in total knee arthroplasty in the United States - an alternative projection model. *Osteoarthritis Cartilage* 2017;25:1797–803.
2 Wylde V, Hewlett S, Learmonth ID, *et al*. Persistent pain after joint replacement: prevalence, sensory qualities, and postoperative determinants. *Pain* 2011;152:566–72.
3 Beswick AD, Wylde V, Gooberman-Hill R, *et al*. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. *BMJ Open* 2012;2:e000435.
4 Liu SS, Buvanendran A, Rathmell JP, *et al*. A cross-sectional survey on prevalence and risk factors for persistent postsurgical pain 1 year after total hip and knee replacement. *Reg Anesth Pain Med* 2012;37:415–22.
5 Buvanendran A, Della Valle CJ, Kroin JS, *et al*. Acute postoperative pain is an independent predictor of chronic postsurgical pain following total knee arthroplasty at 6 months: a prospective cohort study. *Reg Anesth Pain Med* 2019;44:e100036–96.
6 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.
7 Vissers MM, Bussmann JB, Verhaar JAN, *et al*. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. *Semin Arthritis Rheum* 2012;41:576–88.
8 Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess* 1995;7:524–32.
9 Gibson E, Sabo MT. Can pain catastrophizing be changed in surgical patients? A scoping review. *Can J Surg* 2018;61:311–8.
10 Keefe FJ, Shelby RA, Somers TJ, *et al*. Effects of coping skills training and sertraline in patients with non-cardiac chest pain: a randomized controlled study. *Pain* 2011;152:730–41.

11 Riddle DL, Keefe FJ, Nay WT, *et al*. Pain coping skills training for patients with elevated pain catastrophizing who are scheduled for knee arthroplasty: a quasi-experimental study. *Arch Phys Med Rehabil* 2011;92:859–65.
12 Osman A, Barrios FX, Kopper BA, *et al*. Factor structure, reliability, and validity of the pain Catastrophizing scale. *J Behav Med* 1997;20:589–605.
13 Van Damme S, Crombez G, Bijttebier P, *et al*. A confirmatory factor analysis of the pain Catastrophizing scale: invariant factor structure across clinical and non-clinical populations. *Pain* 2002;96:319–24.
14 de Jong JR, Vlaeyen JWS, Onghena P, *et al*. Reduction of pain-related fear in complex regional pain syndrome type I: the application of graded exposure in vivo. *Pain* 2005;116:264–75.
15 Edwards RR. Bingham CO 3rd, Bathon J, Haythornthwaite JA. Catastrophizing and pain in arthritis, fibromyalgia, and other rheumatic diseases. *Arthritis Rheum* 2006;55:325–32.
16 Copsey B, Thompson JY, Vadher K, *et al*. Problems persist in reporting of methods and results for the WOMAC measure in hip and knee osteoarthritis trials. *Qual Life Res* 2019;28:335–43.
17 Lincoln N, Moreton B, Turner K, *et al*. The measurement of psychological constructs in people with osteoarthritis of the knee: a psychometric evaluation. *Disabil Rehabil* 2017;39:372–84.
18 Julian LJ. Measures of anxiety. *Arthritis Care Res* 2011;63.
19 Thorn BE. *Cognitive therapy for chronic pain, first edition: a step-by-step guide*. New York, NY: The Guilford Press, 2004.
20 Laucis NC, Hays RD, Bhattacharyya T. Scoring the SF-36 in orthopaedics: a brief guide. *J Bone Joint Surg Am* 2015;97:1628–34.
21 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606–13.
22 Jordan P, Shedden-Mora MC, Löwe B. Psychometric analysis of the generalized anxiety disorder scale (GAD-7) in primary care using modern item response theory. *PLoS One* 2017;12:e0182162.
23 Hirsh AT, George SZ, Riley JL, *et al*. An evaluation of the measurement of pain catastrophizing by the coping strategies questionnaire. *Eur J Pain* 2007;11:75–81.
24 Matthews JN, Altman DG, Campbell MJ, *et al*. Analysis of serial measurements in medical research. *Br Med J* 1990;300:230–5.
25 Bell ML, King MT, Fairclough DL. Bias in area under the curve for longitudinal clinical trials with missing patient reported outcome data: summary measures versus summary statistics. *Sage Open* 2014:1–12.
26 Pain management and dosing guide, 2016. Available: http://americanpainsociety.org/uploads/education/PAMI_Pain_Mangement_and_Dosing_Guide_02282017.pdf [Accessed 15 Jul 2020].
27 Riddle DL, Wade JB, Jiranek WA, *et al*. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. *Clin Orthop Relat Res* 2010;468:798–806.
28 Lenhard W, Lenhard A. Calculation of effect sizes Dettelbach (Germany): Psychometrica, 2016. Available: https://www.psychometrica.de/effect_size.html
29 Riddle DL, Keefe FJ, Ang DC, *et al*. Pain coping skills training for patients who Catastrophize about pain prior to knee arthroplasty. *J Bone Joint Surg Am* 2019;101:218–27.
30 Birch S, Stilling M, Mechlenburg I, *et al*. No effect of cognitive behavioral patient education for patients with pain catastrophizing before total knee arthroplasty: a randomized controlled trial. *Acta Orthop* 2020;91:98–103.

Appendix 1: A description of the weekly activities and topics covered in the cognitive behavioral therapy (CBT) sessions. The program was developed from Beverly Thorn's 10 session program ("Cognitive Therapy for Chronic Pain: A Step-by-Step Guide," New York: Guilford Publications, 2004.) with modification to fit into either a 4 or 8 weekly session format. Workbooks were provided to the participants at the first session. Topics and exercises undertaken at the session are outlined in the table.

Table: Topics covered in the 4-week and 8-week CBT sessions.

<table>
<tr><th colspan="2">8-week CBT sessions</th><th colspan="2">4-week CBT sessions</th></tr>
<tr><th colspan="2">Topics covered</th><th colspan="2">Topics Covered</th></tr>
<tr><td>Week 1</td><td>Gate Control Theory<br>Rationale for Cognitive Therapy<br>The Stress Response</td><td rowspan="2">Week 1</td><td>Gate Control Theory<br>Rationale for Cognitive Therapy<br>The Stress Response<br>The Stress-Appraisal-Coping Model of Pain<br>Working with Stressors</td></tr>
<tr><td>Week 2</td><td>The Stress-Appraisal-Coping Model of Pain<br>Working with Stressors</td></tr>
<tr><td>Week 3</td><td>Working with Negative Automatic Thoughts</td><td rowspan="2">Week 2</td><td>Working with Negative Automatic Thoughts<br>Creating Coping Self Statements</td></tr>
<tr><td>Week 4</td><td>Creating Coping Self Statements</td></tr>
<tr><td>Week 5</td><td>Working with Negative Automatic Thoughts<br>Intermediate Beliefs</td><td rowspan="2">Week 3</td><td>Working with Negative Automatic Thoughts<br>Intermediate Beliefs<br>Core Beliefs</td></tr>
<tr><td>Week 6</td><td>Core Beliefs</td></tr>
</table>

| | | | |
|---|---|---|---|
| | Working with Core Beliefs | | Working with Core Beliefs |
| Week 7 | Review of Gate Control Model<br>Review of Stress Appraisal | Week 4 | Review of Gate Control Model<br>Review of Stress Appraisal |
| Week 8 | Review of Automatic Thoughts<br>Review of Intermediate Beliefs<br>Review of Core Beliefs | | Review of Automatic Thoughts<br>Review of Intermediate Beliefs<br>Review of Core Beliefs |

Subjects were given workbooks at the first session. The workbooks topic included: education about the relationships between pain, stress, and cognitions, as well as the somatic, emotional and cognitive aspects of stress and pain. Subjects were assigned take home worksheets at each session, in which they recorded stressful events and their "automatic thoughts" triggered by the event, along with changes in emotions, somatic reactions, and behaviors that followed the event. At the beginning of each session the therapist provided an overview of the session and this was followed by a review of the previous session and the worksheet completed in the past week, review of the of assigned lessons, and a review of the session treatment objectives. At the end of each session the therapist assigned homework to be completed before the next session for and a post-session retention evaluation.

The eight-session regime allowed for additional education and practice in identifying, evaluating, and altering maladaptive thought patterns. Simple relaxation exercises were also introduced in session 4 and session 7 of the 4- and 8-week sessions, respectively.