Clin Orthop Relat Res (2022) 480:298-309
DOI 10.1097/CORR.0000000000002052

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

**Papers from the International Consortium for Mental and Social Health in Musculoskeletal Care
Guest Editors: David Ring MD, PhD and Ana-Maria Vranceanu PhD**

Are There Distinct Statistical Groupings of Mental Health Factors and Pathophysiology Severity Among People with Hip and Knee Osteoarthritis Presenting for Specialty Care?

Tom J. Crijns MD[1], Niels Brinkman BS[1], Sina Ramtin MD[1], David Ring MD, PhD[1], Job Doornberg MD, PhD[1], Paul Jutte MD, PhD[1], Karl Koenig MD, MS[1]

Received: 28 May 2021 / Accepted: 26 October 2021 / Published online: 24 November 2021

Abstract

Background There is mounting evidence that objective measures of pathophysiology do not correlate well with symptom intensity. A growing line of inquiry identifies statistical combinations (so-called "phenotypes") of various levels of distress and unhelpful thoughts that are associated with distinct levels of symptom intensity and magnitude of incapability. As a next step, it would be helpful to understand how distress and unhelpful thoughts interact with objective measures of pathologic conditions such as the radiologic severity of osteoarthritis. The ability to identify phenotypes of these factors that are associated with distinct levels of illness could contribute to improved personalized musculoskeletal care in a comprehensive, patient-centered model.

Questions/purposes (1) When measures of mental health are paired with radiologic osteoarthritis severity, are there distinct phenotypes among adult patients with hip and knee osteoarthritis? (2) Is there a difference in the degree of capability and pain self-efficacy among the identified mental health and radiologic phenotypes? (3) When capability (Patient-reported Outcomes Measurement Information System Physical Function [PROMIS PF]) is paired with radiographic osteoarthritis severity, are there distinct phenotypes among patients with hip and knee osteoarthritis? (4) Is there a difference in mental health among patients with the identified capability and radiologic phenotypes?

Methods We performed a secondary analysis of data from a study of 119 patients who presented for musculoskeletal specialty care for hip or knee osteoarthritis. Sixty-seven percent (80 of 119) of patients were women, with a mean age of 62 ± 10 years. Seventy-six percent (91 of 119) of patients had knee osteoarthritis, and 59% (70 of 119) had an advanced radiographic grade of osteoarthritis (Kellgren-Lawrence grade 3 or higher). This dataset is well-suited for our current experiment because the initial study had broad enrollment criteria, making these data applicable to a diverse population and because patients had sufficient variability in radiographic severity of osteoarthritis. All new and returning patients were screened for eligibility. We do not record the percentage of eligible patients who do not participate in cross-sectional surveys, but the rate is typically high (more than 80%). One hundred forty-eight eligible patients started the questionnaires, and 20% (29 of 148) of patients did not complete at least 60% of the questionnaires and were excluded, leaving 119 patients

One of the authors (DR) certifies receipt of personal payments or benefits, during the study period, in an amount of USD 10,000 to 100,000 from Skeletal Dynamics, and in an amount of less than USD 10,000 from Wright Medical.
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.
Ethical approval for this study was obtained from The University of Texas at Austin Institutional Review Board under protocol number 2019-03-0077.

[1]Department of Surgery and Perioperative Care, Dell Medical School at the University of Texas at Austin, Austin, TX, USA

T. J. Crijns ✉, Department of Surgery and Perioperative Care, Dell Medical School, The University of Texas at Austin, Health Discovery Building, 1701 Trinity Street, Austin, TX 78712, USA, Email: tom.j.crijns@gmail.com

Wolters Kluwer

available for analysis. We measured psychologic distress (Patient Health Questionnaire-2 [PHQ-2] and Generalized Anxiety Disorder-2 questionnaire [GAD-2]), unhelpful thoughts about pain (Pain Catastrophizing Scale-4 [PCS-4]), self-efficacy when in pain (Pain Self-Efficacy Questionnaire-2), and capability (PROMIS PF). One of two arthroplasty fellowship–trained surgeons assigned the Kellgren-Lawrence grade of osteoarthritis based on radiographs in the original study. We used a cluster analysis to generate two sets of phenotypes: (1) measures of mental health (PHQ-2, GAD-2, PCS-4) paired with the Kellgren-Lawrence grade and (2) capability (PROMIS PF) paired with the Kellgren-Lawrence grade. We used one-way ANOVA and Kruskal-Wallis H tests to assess differences in capability and self-efficacy and mental health, respectively.

Results When pairing measures of psychologic distress (PHQ-2 and GAD-2) and unhelpful thoughts (catastrophic thinking) with the grade of radiographic osteoarthritis, six distinct phenotypes arose. These groups differed in terms of capability and pain self-efficacy (for example, mild pathology/low distress versus average pathology/high distress [PROMIS PF, mean ± standard deviation]: 43 ± 6.3 versus 33 ± 4.8; $p = 0.003$). When pairing the degree of capability (PROMIS PF) with the Kellgren-Lawrence grade, four distinct phenotypes arose. Patients in three of these did not differ in terms of disease severity but had notable variation in the degree of limitations. Patients with these radiologic and capability phenotypes differed in terms of distress and unhelpful thoughts (for example, moderate pathology/low capability versus mild pathology/high capability [PHQ-2, median and interquartile range]: 3 [1 to 5] versus 0 [0 to 0]; $p < 0.001$).

Conclusion Statistical groupings ("phenotypes") that include both measures of pathology and mental health are associated with differences in symptom intensity and magnitude of incapability and have the potential to help musculoskeletal specialists discern mental and social health priorities. Future investigations may test whether illness phenotype-specific comprehensive biopsychosocial treatment strategies are more effective than treatment of pathology alone.

Level of Evidence Level III, prognostic study.

Introduction

There is mounting evidence that objective measures of pathophysiology, such as the grade of osteoarthritis, do not correlate well with symptom intensity [6, 7, 24, 29]. This phenomenon is noted in the wide variability in patient-reported outcome measures for a given disease [13, 26]. Mental health opportunities (such as unhelpful thoughts, like "hurt always indicates harm," and unhelpful feelings, like worry and despair, which are sometimes grouped as distress) account for a substantial proportion of the observed variation in patient-reported outcome measure scores [12-14, 46, 49, 53]. Distress can reinforce unhelpful thoughts about pain and influence recovery [10, 51]. There is increasing interest in identifying statistical subgroups of patients ("phenotypes") based on measures of mental health [24, 42, 54] to help personalize care strategies. In this setting, the word "phenotype" is used metaphorically and refers to the underlying set of characteristics in these statistical groupings of patients; there is no underlying "genotype." Lentz et al. [35] explored this relatively novel concept using a latent class analysis to identify combinations of distress and unhelpful thoughts that form separate illness categories. It may be perceived as a limitation of latent class analysis that it requires dichotomizing measures of mental health, potentially risking the reinforcement of false dichotomies and the perpetuation of the social stigma associated with decreased mental health [45]. On the other hand, a k-means clustering algorithm can identify potentially useful groupings of aspects of mental health such as unhelpful thoughts and symptoms of despair or worry based on continuous scores. For example, one recent study by Miner et al. [41] identified such groupings and confirmed that they are associated with variation in magnitude of incapability.

The next step is to determine whether it is possible to derive phenotypes that combine mental health elements and a measure of pathophysiology, and whether those groupings are associated with variations in symptom intensity and magnitude of incapability. If so, phenotyping has the potential to identify the most effective care strategies. Specifically, groupings based on radiographic severity of hip and knee osteoarthritis in combination with measures of unhelpful thoughts and unhelpful feelings regarding the arthritis might indicate the degree to which strategies to reorient unhelpful thoughts and alleviate symptoms of distress are health priorities. In addition, given that orthopaedic surgeons and patients may expect the level of capability to correspond with the radiographic severity of osteoarthritis, cluster analysis could also be helpful to estimate the likelihood of a mental health opportunity when these two variables diverge.

We therefore asked: (1) When measures of mental health are paired with radiologic osteoarthritis severity, are there distinct phenotypes among adult patients with hip and knee osteoarthritis? (2) Is there a difference in the degree of capability and pain self-efficacy among the identified mental health and radiologic phenotypes? (3) When capability (Patient-reported Outcomes Measurement Information System Physical Function [PROMIS PF]) is paired with radiographic osteoarthritis severity, are there distinct phenotypes among patients with hip and knee osteoarthritis? (4) Is there a difference in mental health among patients with the identified capability and radiologic phenotypes?

Patients and Methods

Study Overview

More broadly, mental health phenotyping is a statistical method used to classify patients into groups with comparable characteristics (such as levels of symptoms of depression and anxiety and levels of unhelpful thoughts) (Fig. 1). This concept can be used in a clinical setting by identifying which statistical grouping best represents a person's current illness, and the software can inform the clinician the degree to which mental health opportunities are important health opportunities. Given that phenotyping has the potential to give patients and clinicians real-time information about the illness at the point of care, it could assist in setting treatment priorities regarding relative attention to the physical compared to the psychological and social aspects of the illness.

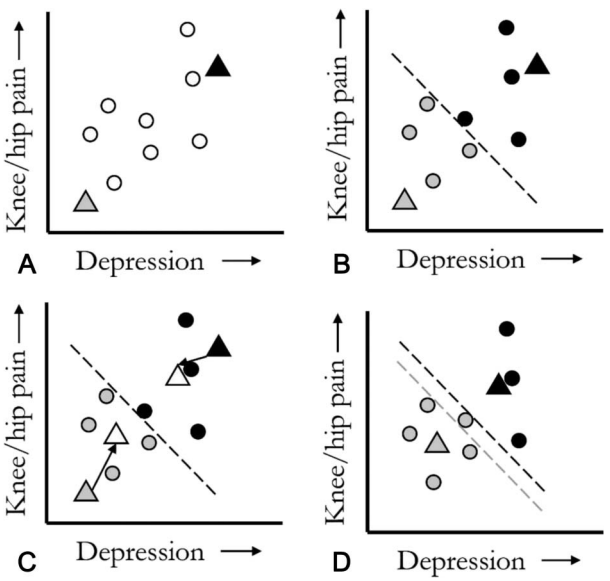

Fig. 1 A-D An example of a model with two clusters. (**A**) Step 1: The black and gray triangles are "centroids" (that is, the middle of the cluster), which will be relocated to the cluster mean. To start, they are placed in a random location. (**B**) Step 2: The field is divided in two (dashed line) based on the location of the centroids. The black and gray shades represent our two clusters. Each datapoint is assigned to the closest centroid. (**C**) Step 3: The centroids are moved from their (random) starting location to the mean value of all datapoints in their respective cluster (gray or black). (**D**) Step 4: Because the centroids moved, the dashed line moves from its old location (light gray) to its new location (black). Note that redividing the field causes one of the black datapoints to be reassigned to the gray cluster because it is now closer to the gray centroid. This process is repeated until the centroids stop moving.

The primary goal of this study was to identify statistical groupings ("phenotypes") that combine elements of both mental health and pathophysiology (radiographic osteoarthritis). These phenotypes can be considered clinically important if the degree of capability (PROMIS PF) and pain self-efficacy (accommodation of potentially painful pathology [nociception]) differ among phenotypes. In addition, we set out to test whether pairing the degree of radiographic osteoarthritis with the level of capability would generate patient profiles with measurable differences in mental health. This is important because such phenotypes could help practicing orthopaedic surgeons assess the likelihood of mental health opportunities when comparing the level of incapability and the degree of radiographic osteoarthritis to increase the chance that such opportunities are addressed.

Study Design and Setting

We performed a secondary analysis of data collected from adult English-speaking patients between 40 and 89 years old who presented with hip or knee osteoarthritis to an urban orthopaedic clinic [16]. This dataset is well-suited for our current experiment because the initial study had broad enrollment criteria, making these data applicable to a diverse population of patients with osteoarthritis. In addition, patients had sufficient variability in radiographic severity of osteoarthritis to warrant entering this as a vector in the k-means clustering algorithm. Our cohort included both patients with hip and knee osteoarthritis. Although research suggests these patient populations may have different levels of satisfaction after total joint arthroplasty [1, 5, 21, 47], there is ample evidence that distress and coping strategies are important factors influencing the magnitude of capability for both of these conditions [9, 16, 31, 35].

All patients with hip and knee osteoarthritis were invited to participate in this study by a research assistant who was not directly involved in their care and who obtained verbal consent before data collection. The osteoarthritis diagnosis was made by a nonspecialist prompting a referral to specialty care, then confirmed by the orthopaedic surgeon. Participants were asked to complete measures of mental health, coping in response to nociception, capability, self-efficacy, and basic demographics. We excluded patients with a fracture or ligament injury and those who had cognitive deficiencies or language barriers that would preclude completion of questionnaires. Completion of the questionnaires implied consent. Participants were invited to participate before or after their visit with the surgeon, and all surveys were administered in a private examination room on a tablet device through the Heath Insurance and Portability Accounting Act–compliant Research Electronic Data Capture internet-based application.

Participants

All new and returning patients were screened for eligibility. We do not record the percentage of eligible patients that does not participate in cross-sectional surveys, but the rate is typically high (more than 80%). One hundred forty-eight eligible patients started the questionnaires, and 20% (29 of 148) of patients did not complete at least 60% of the questionnaires and were thus excluded. These incomplete questionnaires were most likely caused by logistical lapses in which the patient had already left the examination room before questionnaire completion. There were 119 patients with complete questionnaires available for analysis. Sixty-seven percent (80 of 119), of participants were women, with a mean age of 62 ± 10 years (Table 1). A notable percentage of patients received high school education or less (38% [45 of 119]), and 39% (47 of 119) of patients had an annual household income of USD 30,000 or less. Seventy-six percent (91 of 119) of patients had knee osteoarthritis, and 59% (70 of 119) had an advanced radiographic grade of osteoarthritis (Kellgren-Lawrence grade 3 or higher).

Measurements

We measured the degree of capability using the PROMIS PF computerized adaptive test (PROMIS PF CAT) [8, 15, 50]. The PROMIS questionnaires are normed to the general population in the United States; a score of 50 represents the average, with each 10 points above or below 50 representing one SD. Higher scores indicate greater capability (fewer activity limitations).

We collected two measures of psychologic distress: the two-item Patient Health Questionnaire (PHQ-2), which measures symptoms of depression on a continuous scale (range 0 to 6) [33, 37], and the two-item version of the Generalized Anxiety Disorder (GAD-2) questionnaire to measure symptoms of anxiety (range 0 to 6) [36]. In addition, all patients completed the four-item Pain Catastrophizing Scale (PCS-4) to measure unhelpful thoughts (worst case thinking) regarding symptoms (range 0 to 16) [23, 43]. For each of these measures, higher scores indicate greater symptom intensity. We do not report minimum clinically important differences for these scores because these are highly specific to the pathology and timing of their measurement [44]. Patients completed questionnaires before their visit with the specialist and were not aware of their Kellgren-Lawrence grade.

Two of the treating surgeons, both fellowship trained in total joint replacement, rated the severity of osteoarthritis according to the Kellgren-Lawrence grade [30] after the patient visit based on radiographs. The Kellgren-Lawrence grading system has moderate interobserver agreement and has a weak-to-moderate correlation with intraoperatively confirmed cartilage loss [28, 55]. Doubtful joint space narrowing with possible osteophyte lipping is rated as Grade 1; definite osteophytes and possible joint space narrowing is marked as Grade 2; multiple osteophytes, definite joint space narrowing, and possible bony deformity is rated as Grade 3; and Grade 4 is assigned to radiographs demonstrating large osteophytes, marked joint space narrowing, severe subchondral sclerosis, and bony contour deformity [30].

Table 1. Patient demographics (n = 119 patients)

Variable	Value
Age in years	62 ± 10
Gender	
Women	67 (80)
Men	33 (39)
Education	
High school or less	38 (45)
2-year college	16 (19)
4-year college	26 (31)
Postgraduate degree	20 (24)
Work status	
Employed	36 (43)
Retired	37 (44)
Other	27 (32)
Annual household income	
< USD 30,000	39 (47)
USD 30,000 to USD 99,999	36 (43)
> USD 100,000	24 (29)
Insurance	
Private	38 (45)
Public or no insurance	62 (74)
GAD-2	2 (0-3)
PHQ-2	1 (0-3)
PCS-4	4 (2-7)
PSEQ-2	6 (4-10)
PROMIS Physical Function (t-score)	38 ± 6.9
Joint	
Knee	76 (91)
Hip	24 (28)
Kellgren-Lawrence grade	
1	21 (25)
2	20 (24)
3	25 (30)
4	34 (40)

Continuous variables are presented as mean ± SD or median (interquartile range); discrete variables are presented as % (n); GAD-2 = Generalized Anxiety Disorder, two-item; PHQ-2 = Patient Health Questionnaire, two-item; PCS-4 = Pain Catastrophizing Scale, four-item; PSEQ-2= Pain Self-efficacy Questionnaire, two-item.

Ethical Approval

Ethical approval for this study was obtained from The University of Texas at Austin Institutional Review Board under protocol number 2019-03-0077.

Statistical Analysis

Descriptive statistics were performed for all participants. Continuous variables are reported as the mean ± SD or median (interquartile range), depending on their distribution. We used a cluster analysis (Stata 13.0, StataCorp) to identify two sets of phenotypes: (1) measures of mental health (PHQ-2, GAD-2, and PCS-4) paired with the Kellgren-Lawrence grade and (2) the degree of capability (PROMIS PF CAT) paired with the Kellgren-Lawrence grade. Cluster analysis is a method for identifying statistical groupings in continuous data (see Fig. 1) and can parse patients into subgroups at the point of care. For the first set of phenotypes (mental health and radiologic osteoarthritis), two measures of psychologic distress (PHQ-2 and GAD-2), one measure of coping in response to nociception (PCS-4), and one measure of osteoarthritis (Kellgren-Lawrence grade) were entered into the k-means clustering algorithm. Scree plots were generated for the mental health and radiologic osteoarthritis phenotypes, and we allowed up to 20 unique clusters. We used the elbow method described by Yuan and Yang [56] to determine the optimal number of clusters, minimizing prediction error and the number of unique clusters. Several sets of clusters were generated around the optimum point (a combination of low prediction error and a low number of phenotypes), and the authors collectively decided on the set of phenotypes that seemed most consistent with daily practice and most useful for guiding health strategies. This process was repeated for the second set of phenotypes, using the PROMIS PF CAT and Kellgren-Lawrence grades to generate independent clusters. The principles used in this study to select phenotypes include generating a number of phenotypes with sufficient variation (that is, having at least one group that is one SD above or below the population average) in measures of mental health and selecting a number of phenotypes at which (most) groups are of sufficient size. We used one-way ANOVA to test for differences in capability (PROMIS PF CAT) between the identified phenotypes; we used the Kruskal-Wallis H test to test for differences in symptoms of depression (Pain Self-Efficacy Questionnaire-2 [PHQ-2]), anxiety (GAD-2), pain self-efficacy (PSEQ-2), and pain catastrophizing (PCS-4). We performed post hoc pairwise comparisons using the Tukey range test and the Dunn test for parametric and non-parametric variables, respectively. Statistical significance was set at $p = 0.05$.

Results

Phenotypes Pairing Mental Health with Radiographic Osteoarthritis

When pairing measures of psychologic distress (PHQ-2 and GAD-2) and unhelpful thoughts (catastrophic thinking) with the grade of radiographic osteoarthritis, six distinct statistical groupings arose (Table 2).

Table 2. Characteristics of mental health and radiologic phenotypes

Phenotype	Number	Radiographic OA Kellgren-Lawrence grade	Symptoms of distress		Unhealthy thoughts about pain[c]
			Depression[a]	Anxiety[b]	
Advanced disease and low distress	31	High (+0.42 SD)	Low (-0.32 SD)	Low (-0.35 SD)	Low (-0.33 SD)
Advanced disease and high distress	18	High (+0.20 SD)	High (+1.4 SD)	High (+1.7 SD)	High (+1.2 SD)
Average disease and high distress	8	Moderate (+0.078 SD)	High (+1.6 SD)	Moderate (+0.21 SD)	High (+0.91 SD)
Average disease and unhealthy thoughts	23	Moderate (-0.092 SD)	Low (-0.29 SD)	Low (-0.34 SD)	High (+0.92 SD)
Average disease and notable distress	10	Low (-0.18 SD)	Moderate (+0.21 SD)	High (+1.2 SD)	Low (-0.44 SD)
Mild disease and low distress	29	Low (-0.41 SD)	Low (-0.81 SD)	Low (-0.87 SD)	Low (-1.2 SD)

The value between parentheses indicates the number of SDs patients in each cluster differ from the mean value of the entire patient population.
[a]Patient Health Questionnaire, two-item.
[b]Generalized Anxiety Disorder, two-item.
[c]Pain Catastrophizing Scale, four-item.

Capability and Pain Self-efficacy Differences Among Mental Health/Radiographic Osteoarthritis Phenotypes

There were differences in the degree of capability (for example, mild pathology/low distress and unhelpful thoughts versus average pathology/high distress unhelpful thoughts [PROMIS PF]: 43 ± 6.3 versus 33 ± 4.8; p = 0.003) and pain self-efficacy (for example, mild pathology/low distress unhelpful thoughts versus average pathology/high distress unhelpful thoughts [PSEQ-2]: 9 [IQR 7 to 12] versus 3.5 [IQR 1.5 to 5]; p < 0.001) for phenotypes that paired measures of psychologic distress, unhelpful thoughts, and the grade of radiographic osteoarthritis (Table 3). The magnitude of capability (PROMIS PF CAT) and pain self-efficacy (PSEQ-2) were lower for phenotypes with greater distress and unhelpful thoughts.

Phenotypes Pairing Capability (PROMIS PF CAT) With Radiographic Osteoarthritis

When pairing the degree of capability (PROMIS PF) with the Kellgren-Lawrence grade, four unique phenotypes arose (Table 4). There was little variation in radiographic osteoarthritis severity among three of the four clusters (a quarter of an SD an less), but there were differences in the degree of capability (for example, moderate pathology/low capability versus mild pathology/high capability [PROMIS PF]: 29 ± 3.1 versus 53 ± 3.3; p < 0.001).

Mental Health Differences Among Radiographic OA/Capability Phenotypes

Symptoms of depression and anxiety and catastrophic thinking were higher and pain self-efficacy was lower in phenotypes that had greater divergence between pathologic severity and the level of capability (Table 5). For example, the PHQ-2 had a median value of 3 (IQR 1 to 5) for the phenotype "moderate pathology/low capability" compared with 0 (IQR 0 to 0) for "mild pathology/high capability" (p < 0.001).

Discussion

Statistical grouping of unhelpful thoughts and symptoms of distress (worry or despair) has shown promise for potential tailoring of musculoskeletal care to an individual's health priorities [24, 35, 42, 54]. As a next step, we can further our understanding of how symptoms of distress and unhelpful thoughts interact with objective measures of disease, such as the radiographic grade of osteoarthritis. We performed a cluster analysis of patients with hip and knee osteoarthritis and identified six distinct phenotypes of mental health and radiographic severity and found differences in the degree of capability and self-efficacy between these groups. An alternative cluster analysis of the magnitude of capability and radiographic severity identified four distinct groups of levels of unhelpful thoughts and distress about symptoms. These findings indicate that identification of phenotypes based on measures of mental health and pathology severity is a feasible method for identifying health priorities for people seeking musculoskeletal specialty care.

Limitations

This study has several limitations. First, our analysis only included patients who felt that their hip or knee symptoms were sufficiently bothersome to seek specialty care and had access to care. It is possible that specific combinations of pathology severity and mental health factors did not exist in this cohort, and population-based cohort studies that include people not presenting for care might yield different results [29]. Second, some patients do not answer questionnaires that quantify symptoms of depression and anxiety forthrightly, contributing to notable floor effects that are associated with less time to completion [3, 4, 19]. Third, 20% (29 of 148) of patients who started the surveys did not complete at least 60% of the questionnaire. We know there were logistical lapses related to an inexperienced enroller, and so we believe that these introduced data loss at random, but it is also possible that patients were less likely to complete the questionnaires if they had greater symptoms of depression or anxiety. These two factors might contribute to an underestimation of the impact of symptoms of depression and anxiety on the degree of capability and symptom intensity.

In addition, most patients presenting to our urban clinic had a low annual income and had public or no insurance, perhaps limiting the generalizability of these findings. On the other hand, the experimental technique depends more on variation within the sample than on representativeness of the sample, so most populations with sufficient diversity in mental health and pathology will have identifiable phenotypes. The associations, and the ability of cluster analysis to generate mental health phenotypes that have distinct levels of illness, are likely more reproducible than the percentage of patients in each category [42]. Another limitation pertains to how to frame these models; as with other statistical techniques such as propensity scoring and multivariable models, the authors' judgment helps determine the construction and interpretation of the statistical models. Statisticians indicate that this aspect of analysis is unavoidable, and they emphasize the importance of preanalysis expert input to inform statistical models. In our early use of cluster analysis to explore the potential utility of statistically derived phenotypes, we have experienced the choices and interpretations that go into the model and suggest that consensus principles for making these determinations would be helpful. For instance, there is some

Table 3. Side-by-side comparison of patient-reported outcome measures by mental health and radiologic phenotypes

Capability		Advanced disease and low distress	Advanced disease and high distress	Average disease and high distress	Average disease and unhealthy thoughts	Average disease and notable distress	Mild disease and low distress
	PROMIS PF CAT	39 ± 6.7	34 ± 4.9	33 ± 4.8	35 ± 5.1	38 ± 8.6	43 ± 6.3
Advanced disease and low distress	39 ± 6.7	-					
Advanced disease and high distress	34 ± 4.9	p = 0.13	-				
Average disease and high distress	33 ± 4.8	p = 0.29	p > 0.99	-			
Average disease and unhealthy thoughts	35 ± 5.1	p = 0.48	p = 0.97	p = 0.96	-		
Average disease and notable distress	38 ± 8.6	p > 0.99	p = 0.55	p = 0.61	p = 0.89	-	
Mild disease and low distress	43 ± 6.3	p = 0.09	p < 0.001	p = 0.003	p = 0.001	p = 0.29	-

Pain self-efficacy		Advanced disease and low distress	Advanced disease and high distress	Average disease and high distress	Average disease and unhealthy thoughts	Average disease and notable distress	Mild disease and low distress
	PSEQ-2	7 (5-10)	4.5 (3-8)	3.5 (1.5-5)	6 (4-8)	7.5 (5-10)	9 (7-12)
Advanced disease and low distress	7 (5-10)	-					
Advanced disease and high distress	4.5 (3-8)	p = 0.042	-				
Average disease and high distress	3.5 (1.5-5)	p = 0.009	p = 0.17	-			
Average disease and unhealthy thoughts	6 (4-8)	p = 0.15	p = 0.22	p = 0.057	-		
Average disease and notable distress	7.5 (5-10)	p = 0.47	p = 0.11	p = 0.028	p = 0.25	-	
Mild disease and low distress	9 (7-12)	p = 0.042	p < 0.001	p < 0.001	p = 0.005	p = 0.098	-

PROMIS PF CAT presented as the mean ± SD, and the PSEQ-2 presented as the median (IQR). One-way ANOVA/Tukey range tests were performed to compare PROMIS PF CAT; Kruskal-Wallis H tests/Dunn tests were performed for PSEQ-2; PROMIS PF CAT = Patient-reported Outcomes Measurement Information System Physical Function computerized adaptive test; PSEQ-2 = Pain Self-efficacy Questionnaire, two-item.

Table 4. Characteristics of capability and radiologic phenotypes

Phenotype	Number	Radiographic OA Kellgren-Lawrence grade	Capability PROMIS PF CAT
Moderate disease and low capability	27	Moderate (+0.088 SD)	Low (-1.3 SD)
Moderate disease and moderate capability	56	Moderate (-0.047 SD)	Moderate (-0.13 SD)
Moderate disease and high capability	29	Moderate (+0.069 SD)	High (+0.93 SD)
Mild disease and high capability	7	Low (-0.25 SD)	High (+2.2 SD)

The value between parentheses indicates the number of SDs patients in each cluster differ from the mean value of the entire patient population; PROMIS PF CAT = Patient-reported Outcomes Measurement Information System Physical Function computerized adaptive test; OA = osteoarthritis.

degree of subjectivity in selecting the number and size of statistical groupings given that cluster analysis attempts to balance limiting prediction error with having a practical number of distinct phenotypes. The number of clusters was chosen based on the sense that it matches what is encountered in patient care, is subject to author bias, and merits testing in other settings for reliability. The specific set of phenotypes selected may be relatively unimportant as long as each of the several possible cluster selections map to associations with patient-reported outcome measures or mental health measures that prove useful for patient care and research. This study used two different phenotyping methods and both mapped to useful associations.

We also recognize that assigning Kellgren-Lawrence grades has limited interobserver reliability. Since there is not a better measure of pathology available—and it is difficult to conceive of one—there may be fixed limits to the ability to incorporate pathology severity in statistical categorization. Given that patients and surgeons use radiographic severity of arthritis to make treatment decisions, our study using Kellgren-Lawrence grades assigned by fellowship-trained arthroplasty surgeons seems relevant to current practice. Finally, the cohort of patients with hip osteoarthritis was relatively small, which precluded a separate subanalysis. This was an exploratory analysis on a suitable data set and can lead to more specific analyses separating patients with hip and knee arthritis.

Mental Health and Radiologic Phenotypes in Patients with Osteoarthritis

The observation that six statistical groupings of mental health factors and radiographic severity among patients with hip and knee osteoarthritis corresponded with differences in the magnitude of capability and accommodation of pain highlights the potential use of these groupings to help develop health strategies with individual patients. Our results agree with a latent class analysis that pooled patients into the classes of high distress, low distress, low self-efficacy and acceptance, and negative pain coping [35]. In contrast to the study of Miner et al. [41], their analysis dichotomized each trait in the Optimal Screening for Prediction of Referral and Outcome Yellow Flag tool [34] and evaluated the prevalence of each of those characteristics in these latent classes. Our current study adds to this line of evidence by demonstrating that surgeon-rated radiographic severity of osteoarthritis interacts with mental health factors on their continuum to form subgroups associated with levels of incapability. The incorporation of radiographic severity in the groupings allows patients and surgeons to notice and discuss when incapability corresponds more or less with mental health opportunities such as unhelpful thoughts about symptoms of feelings of worry or despair about symptoms than with pathology severity. Both patients and surgeons might be more interested in prioritizing mental health after receiving this information. Our results are consistent with a population-based study by Nigoro et al. [42], which identified six distinct clusters in patients with OA based on their BMI, radiologic grade of osteoarthritis, bone mineral density, high-sensitivity C-reactive protein level, and depression state. In contrast to our study, only one of the clusters identified by Nigoro et al. [42] had a mental health index that the authors labeled as "having depressive symptoms," whereas our current analysis generated multiple phenotypes of patients who had noteworthy symptoms of depression or anxiety, interacting with levels of unhelpful thoughts. Our research group strives to avoid categorizing mental health because categories risk reinforcing false dichotomies and the social stigma associated with mental health [45]. The multiple statistical groupings identified can help clinicians and patients personalize health strategies while limiting stigma. Next steps might include repeating this type of experiment in other settings and separating hip and knee osteoarthritis. If reproducible groupings are identified, we can test the use of these groupings as points of discussion with an individual patient and whether this process helps musculoskeletal specialists and patients develop more effective health strategies.

The finding of four identifiable capability or radiologic phenotypes, with differences in measures of distress

Table 5. Side-by-side comparison of measures of mental health by capability and radiologic phenotypes

Depression		Moderate disease and low capability	Moderate disease and moderate capability	Moderate disease and high capability	Mild disease and high capability
	PHQ-2	3 (1-5)	1 (0-3)	0 (0-1)	0 (0-0)
Moderate disease and low capability	3 (1-5)	-			
Moderate disease and moderate capability	1 (0-3)	p < 0.001	-		
Moderate disease and high capability	0 (0-1)	p < 0.001	p = 0.006	-	
Mild disease and high capability	0 (0-0)	p < 0.001	p = 0.008	p = 0.17	-

Anxiety		Moderate disease and low capability	Moderate disease and moderate capability	Moderate disease and high capability	Mild disease and high capability
	GAD-2	3 (1-6)	1 (0-3)	0 (0-1)	0 (0-0)
Moderate disease and low capability	3 (1-6)	-			
Moderate disease and moderate capability	1 (0-3)	p = 0.001	-		
Moderate disease and high capability	0 (0-1)	p = 0.005	p = 0.46	-	
Mild disease and high capability	0 (0-0)	p = 0.036	p = 0.44	p = 0.42	-

Self-efficacy		Moderate disease and low capability	Moderate disease and moderate capability	Moderate disease and high capability	Mild disease and high capability
	PSEQ-2	3 (1-5)	6 (5-9)	10 (8-12)	12 (8-12)
Moderate disease and low capability	3 (1-5)	-			
Moderate disease and moderate capability	6 (5-9)	p = 0.001	-		
Moderate disease and high capability	10 (8-12)	p < 0.001	p < 0.001	-	
Mild disease and high capability	12 (8-12)	p < 0.001	p = 0.003	p = 0.33	-

Unhealthy thoughts about pain		Moderate disease and low capability	Moderate disease and moderate capability	Moderate disease and high capability	Mild disease and high capability
	PCS-4	7 (4-8)	4 (3-6)	3 (1-4)	1 (0-3)
Moderate disease and low capability	7 (4-8)	-			
Moderate disease and moderate capability	4 (3-6)	p = 0.007	-		
Moderate disease and high capability	3 (1-4)	p < 0.001	p = 0.004	-	
Mild disease and high capability	1 (0-3)	p < 0.001	p = 0.003	p = 0.12	-

Data presented as median (IQR). Dunn tests were performed for all variables; PHQ-2 = Patient Health Questionnaire, two-item; GAD-2 = Generalized Anxiety Disorder, two-item; PSEQ-2 = Pain Self-efficacy Questionnaire, two-item; PCS-4= Pain Catastrophizing Scale, four-item.

(symptoms of depression and anxiety) and unhelpful thoughts (like catastrophic thinking), suggests that musculoskeletal specialty care may be improved by personalizing treatment based on the alignment between symptom intensity and objective evidence of osteoarthritis. Indeed, a musculoskeletal specialist may be the first clinician to discern mental health opportunities because the discordance between symptom intensity and disease may not be discerned by nonspecialists. This represents an important role for musculoskeletal specialists in the mental health of individuals and populations, and our current study suggests that this discernment could be facilitated by using mental health phenotyping. Relative incongruency between the radiographic severity of osteoarthritis and degree of capability is associated with notable psychologic distress and misconceptions about pain, which is consistent with evidence that psychologic factors account for a substantial proportion of variation in symptom intensity and degree of incapability [12, 25, 26, 32, 40]. Given the correlation between mental health and symptom intensity [52], one might assume that biomedical treatment that addresses osteoarthritis will decrease symptoms and improve symptoms of psychologic distress [17, 39], but there is evidence that operative treatment is not associated with more relief of depression symptoms than nonoperative treatment [11], and that preoperative depression symptoms are associated with less favorable recovery trajectories [20, 27]. There is ample evidence to suggest that performing elective procedures on patients with untreated or poorly controlled symptoms of distress (symptoms of anxiety and depression in particular) is associated with persistent pain and dissatisfaction [2, 12, 25, 38, 48]. On the other hand, there is insufficient evidence to suggest that biomedical treatment alone reduces emotional distress. Therefore, we speculate that among the subset of patients with discordance between pain, activity limitations, and objective evidence of disease, the mental health aspects of the illness should be addressed in parallel and perhaps prioritized before discretionary surgery [18], which might decrease the proportion of surgical interventions that result in less desirable patient-reported outcomes; future investigations might test this idea.

Conclusion

We performed a secondary analysis of people with hip and knee osteoarthritis presenting for musculoskeletal care and found that there is identifiable clustering of mental health factors and radiologic severity that correspond with magnitude of capability. We also found groupings of the magnitude of capability and radiologic severity that correspond with mental health opportunities such as unhelpful thoughts and feelings of worry or despair. Our results demonstrate that combined mental health and pathophysiology groupings might be a viable tool for patients and specialists to simplify the complexity of such measurements and discussions, thereby helping to personalize treatment strategies in the biopsychosocial model [18]. As the ability to identify meaningful groupings of thoughts, feelings, and pathology is established by this line of research, future efforts can develop and test treatment strategies for using a health phenotype tool to set treatment priorities (that is, the balance and timing of interventions that address mental health relative to interventions to address pathophysiology).

Acknowledgments We thank Kenneth Furlough MD, Harisson Miner MD, Devin Garza, Kara Titus, Devin Patel, Niki Mehra, Ray Kitziger, Dhairye Dave, Joshua Mathew, Krishna Anand, Sarah Grant, and Monica Trevino for collecting data. We thank Karl Koenig MD, MS, and Kevin Bozic MD, MBA, for assigning the Kellgren-Lawrence grade of osteoarthritis.

References

1. Ali A, Sundberg M, Robertsson O, et al. Dissatisfied patients after total knee arthroplasty: a registry study involving 114 patients with 8-13 years of followup. *Acta Orthop*. 2014;85:229-233.
2. Bae J-Y, Kim JK, Yoon JO, Kim JH, Ho BC. Preoperative predictors of patient satisfaction after carpal tunnel release. *Orthop Traumatol Surg Res*. 2018;104:907-909.
3. Beleckas CM, Prather H, Guattery J, Wright M, Kelly M, Calfee RP. Anxiety in the orthopedic patient: using PROMIS to assess mental health. *Qual Life Res*. 2018;27:2275-2282.
4. Bernstein DN, Atkinson J, Fear K, et al. Determining the generalizability of the PROMIS depression domain's floor effect and completion time in patients undergoing orthopaedic surgery. *Clin Orthop Relat Res*. 2019; 477:2215-2225.
5. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clin Orthop Relat Res*. 2010;468:57-63.
6. Briet JP, Houwert RM, Hageman MGJS, Hietbrink F, Ring DC, Verleisdonk EJJM. Factors associated with pain intensity and physical limitations after lateral ankle sprains. *Injury*. 2016;47:2565-2569.
7. Cho HJ, Chang CB, Yoo JH, Kim SJ, Kim TK. Gender differences in the correlation between symptom and radiographic severity in patients with knee osteoarthritis. *Clin Orthop Relat Res*. 2010;468:1749-1758.
8. Cook KF, Jensen SE, Schalet BD, et al. PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. *J Clin Epidemiol*. 2016;73:89-102.
9. Costello CA, Liu M, Furey A, Rahman P, Randell EW, Zhai G. Association between epidemiological factors and nonresponders to total joint replacement surgery in primary osteoarthritis patients. *J Arthroplasty*. 2021;36:1502-1510.e5.
10. Cremers T, Zoulfi Khatiri M, van Maren K, Ring D, Teunis T, Fatehi A. Moderators and mediators of activity intolerance related to pain. *J Bone Joint Surg Am*. 2021;103:205-212.
11. Crijns TJ, Bernstein DN, Gonzalez R, Wilbur D, Ring D, Hammert WC. Operative treatment is not associated with more relief of depression symptoms than nonoperative treatment in patients with common hand illness. *Clin Orthop Relat Res*. 2020;478:1319-1329.

12. Crijns TJ, Bernstein DN, Ring D, Gonzalez RM, Wilbur D, Hammert WC. Depression and pain interference correlate with physical function in patients recovering from hand surgery. *Hand (N Y)*. 2018:14:830-835.
13. Crijns TJ, Bernstein DN, Ring D, Gonzalez R, Wilbur D, Hammert WC. Factors associated with a discretionary upper-extremity surgery. *J Hand Surg Am*. 2019;44:155.e1-155.e7.
14. Das De S, Vranceanu A-M, Ring DC. Contribution of kinesophobia and catastrophic thinking to upper-extremity-specific disability. *J Bone Joint Surg Am*. 2013;95:76-81.
15. Driban JB, Morgan N, Price LL, Cook KF, Wang C. Patient-Reported Outcomes Measurement Information System (PROMIS) instruments among individuals with symptomatic knee osteoarthritis: a cross-sectional study of floor/ceiling effects and construct validity. *BMC Musculoskelet Disord*. 2015;16:253.
16. Furlough K, Miner H, Crijns TJ, Jayakumar P, Ring D, Koenig K. What factors are associated with perceived disease onset in patients with hip and knee osteoarthritis? *J Orthop*. 2021;26:88-93.
17. Goh GS, Zeng GJ, Chen JY, Lo N-N, Yeo S-J, Liow MHL. Ten-year results of unicompartmental knee arthroplasty in patients with psychological distress. *J Arthroplasty*. 2020;35:2830-2836.e1.
18. Gonzalez AI, Kortlever JTP, Brown LE, Ring D, Queralt M. Can crafted communication strategies allow musculoskeletal specialists to address health within the biopsychosocial paradigm? *Clin Orthop Relat Res*. 2021;479:1217-1223.
19. Guattery JM, Dardas AZ, Kelly M, Chamberlain A, McAndrew C, Calfee RP. Floor effect of PROMIS depression CAT associated with hasty completion in orthopaedic surgery patients. *Clin Orthop Relat Res*. 2018;476:696-703.
20. Hafkamp FJ, de Vries J, Gosens T, den Oudsten BL. The relationship between psychological aspects and trajectories of symptoms in total knee arthroplasty and total hip arthroplasty. *J Arthroplasty*. 2021;36:78-87.
21. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. Patient dissatisfaction after primary total joint arthroplasty: the patient perspective. *J Arthroplasty*. 2019;34:1093-1096.
22. Hall M, Doherty S, Courtney P, Latief K, Zhang W, Doherty M. Synovial pathology detected on ultrasound correlates with the severity of radiographic knee osteoarthritis more than with symptoms. *Osteoarthritis Cartilage*. 2014;22:1627-1633.
23. Hampton SN, Nakonezny PA, Richard HM, Wells JE. Pain catastrophizing, anxiety, and depression in hip pathology. *Bone Joint J*. 2019;101-B:800-807.
24. de Hoop AMS, Kloek CJJ, Pisters MF, Veenhof C. Movement behaviour patterns in patients with hip and/or knee osteoarthritis in the physical therapy setting: a cross-sectional study. *BMC Musculoskelet Disord*. 2020;21:651.
25. Jayakumar P, Overbeek CL, Lamb S, et al. What factors are associated with disability after upper extremity injuries? A systematic review. *Clin Orthop Relat Res*. 2018;476:2190-2215.
26. Jayakumar P, Teunis T, Vranceanu A-M, et al. Psychosocial factors affecting variation in patient-reported outcomes after elbow fractures. *J Shoulder Elbow Surg*. 2019;28:1431-1440.
27. Kamalapathy P, Kurker KP, Althoff AD, Browne JA, Werner BC. The impact of mental illness on postoperative adverse outcomes after outpatient joint surgery. *J Arthroplasty*. 2021;36:2734-2741.
28. Keenan OJF, Holland G, Maempel JF, Keating JF, Scott CEH. Correlations between radiological classification systems and confirmed cartilage loss in severe knee osteoarthritis. *Bone Joint J*. 2020;102:301-309.
29. Kim KW, Han JW, Cho HJ, et al. Association between comorbid depression and osteoarthritis symptom severity in patients with knee osteoarthritis. *J Bone Joint Surg Am*. 2011;93:556-563.
30. Kohn MD, Sassoon AA, Fernando ND. Classifications in brief: Kellgren-Lawrence classification of osteoarthritis. *Clin Orthop Relat Res*. 2016;474:1886-1893.
31. Kopp B, Furlough K, Goldberg T, Ring D, Koenig K. Factors associated with pain intensity and magnitude of limitations among people with hip and knee arthritis. *J Orthop*. 2021;25:295-300.
32. Kortlever JTP, Janssen SJ, van Berckel MMG, Ring D, Vranceanu AM. What is the most useful questionnaire for measurement of coping strategies in response to nociception? *Clin Orthop Relat Res*. 2015;473:3511-3518.
33. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;6:606-613.
34. Lentz TA, Beneciuk JM, Bialosky JE, et al. Development of a yellow flag assessment tool for orthopaedic physical therapists: results from the Optimal Screening for Prediction of Referral and Outcome (OSPRO) cohort. *J Orthop Sports Phys Ther*. 2016;46:327-343.
35. Lentz TA, George SZ, Manickas-Hill O, et al. What general and pain-associated psychological distress phenotypes exist among patients with hip and knee osteoarthritis? *Clin Orthop Relat Res*. 2020;478:2768-2783.
36. Löwe B, Decker O, Müller S, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Med Care*. 2008;46:266-274.
37. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res*. 2005;58:163-171.
38. Lozano Calderón SA, Paiva A, Ring D. Patient satisfaction after open carpal tunnel release correlates with depression. *J Hand Surg Am*. 2008;33:303-307.
39. Mahdi A, Hälleberg-Nyman M, Wretenberg P. Reduction in anxiety and depression symptoms one year after knee replacement: a register-based cohort study of 403 patients. *Eur J Orthop Surg Traumatol*. 2021;31:1215-1224.
40. Menendez ME, Ring D. Factors associated with greater pain intensity. *Hand Clin*. 2016;32:27-31.
41. Miner H, Rijk L, Thomas J, Ring D, Reichel LM, Fatehi A. Mental-health phenotypes and patient-reported outcomes in upper-extremity illness. *J Bone Joint Surg Am*. 2021;103:1411-1416.
42. Nigoro K, Ito H, Kawata T, et al. Obesity with radiological changes or depression was associated with worse knee outcome in general population: a cluster analysis in the Nagahama study. *Arthritis Res Ther*. 2020;22:280.
43. Ong WJ, Kwan YH, Lim ZY, et al. Measurement properties of pain catastrophizing scale in patients with knee osteoarthritis. *Clin Rheumatol*. 2020;40:295-301.
44. Ring D. Benchmarking changes in symptom intensity and magnitude of disability. *J Hand Surg Am*. 2013;38:650-651.
45. Ring D. Invited commentary: depression is continuous, not categorical. *J Orthop Trauma*. 2016;30:233-234.
46. Ring D, Kadzielski J, Fabian L, Zurakowski D, Malhotra LR, Jupiter JB. Self-reported upper extremity health status correlates with depression. *J Bone Joint Surg Am*. 2006;88:1983-1988.
47. Scott CEH, Oliver WM, MacDonald D, Wade FA, Moran M, Breusch SJ. Predicting dissatisfaction following total knee arthroplasty in patients under 55 years of age. *Bone Joint J*. 2016;98:1625-1634.
48. Shah RF, Gwilym SE, Lamb S, Williams M, Ring D, Jayakumar P. Factors associated with persistent opioid use after an upper extremity fracture. *Bone Joint Open*. 2021;2:119-124.

49. Shin YH, Yoon JO, Kim YK, Kim JK. Psychological status is associated with symptom severity in patients with carpal tunnel syndrome. *J Hand Surg Am*. 2018;43:484.e1-484.e8.
50. Stone AA, Broderick JE, Junghaenel DU, Schneider S, Schwartz JE. PROMIS fatigue, pain intensity, pain interference, pain behavior, physical function, depression, anxiety, and anger scales demonstrate ecological validity. *J Clin Epidemiol*. 2016;74:194-206.
51. Teunis T, Bot AGJ, Thornton ER, Ring D. Catastrophic thinking is associated with finger stiffness after distal radius fracture surgery. *J Orthop Trauma*. 2015;29:e414-20.
52. Trinh JQ, Carender CN, An Q, Noiseux NO, Otero JE, Brown TS. Resilience and depression influence clinical outcomes following primary total joint arthroplasty. *J Arthroplasty*. 2021;36:1520-1526.
53. Vissers MM, Bussmann JB, Verhaar JAN, Busschbach JJV, Bierma-Zeinstra SMA, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. *Semin Arthritis Rheum*. 2012;41:576-88.
54. Vongsirinavarat M, Nilmart P, Somprasong S, Apinonkul B. Identification of knee osteoarthritis disability phenotypes regarding activity limitation: a cluster analysis. *BMC Musculoskelet Disord*. 2020;21:237.
55. Wright RW. Osteoarthritis classification scales: interobserver reliability and arthroscopic correlation. *J Bone Joint Surg Am*. 2014;96:1145-1151.
56. Yuan C, Yang H. Research on k-value selection method of k-means clustering algorithm. *J*. 2019; 2:226-235.