

# Normative PROMIS Scores in Healthy Collegiate Athletes

## Establishing a Target for Return to Function in the Young Adult Athlete

Arya Minaie,\* MD, David L. Bernholt,<sup>†</sup> MD, Andrew M. Block,<sup>‡</sup> BS, Ronak M. Patel,<sup>‡</sup> MD, Rick W. Wright,<sup>§</sup> MD, Matthew J. Matava,<sup>‡</sup> MD, and Jeffrey J. Nepple,<sup>‡||</sup> MD

*Investigation performed at Washington University Orthopaedics, St Louis, Missouri, USA*

**Background:** The Patient-Reported Outcomes Measurement Information System (PROMIS) computer-adaptive testing (CAT) has been shown to be a valid and reliable means of assessing patient-reported outcomes. However, normal scores and distributions for a subset of a healthy young athletic population have not been established.

**Purpose:** To establish normative PROMIS scores for the domains of Physical Function (PF-CAT), Mobility (M-CAT), Upper Extremity Function (UE-CAT), and Pain Interference (PI-CAT) and determine the frequency of floor and ceiling effects in a population of healthy collegiate athletes.

**Study Design:** Cross-sectional study; Level of evidence, 3.

**Methods:** Healthy collegiate athletes (18-23 years of age) were prospectively enrolled to complete the 4 PROMIS CAT domains. Additionally, the athletes provided information regarding their age, sex, and sport(s). Mean scores ( $\pm$ SD) and identification of ceiling or floor effects were calculated. Ceiling and floor effects were considered significant if  $>15\%$  of the participants obtained the highest or lowest possible score on a domain.

**Results:** A total of 194 healthy athletes (mean age, 19.1 years) were included in the study: 118 (60.8%) men and 76 (39.2%) women. Mean scores were  $62.9 \pm 6.7$  for PF-CAT,  $58.2 \pm 4.1$  for M-CAT,  $57.4 \pm 5.8$  for UE-CAT, and  $43.2 \pm 6.2$  for PI-CAT. Distributions of scores for M-CAT and UE-CAT indicated strong ceiling effects by 77.3% and 66.0% of the participants, respectively. In healthy athletes, the PF-CAT differed most from the expected population-based mean score (50), with the mean being  $>1$  SD above (62.9), without a ceiling effect observed. There were no significant sex- or age-based differences on any of the PROMIS domain scores.

**Conclusion:** Healthy collegiate athletes scored nearly 1 SD from population-based means for all of the domains tested. M-CAT and UE-CAT demonstrated ceiling effects in more than two-thirds of healthy athletes, which may limit their utility in this population. The PF-CAT did not demonstrate floor or ceiling effects and demonstrated differences in a young adult athletic population from the population mean. The mean PF-CAT score of 62.9 can represent a target for return of function in injured athletes.

**Keywords:** Patient-Reported Outcomes Measurement Information System; PROMIS; upper extremity; physical function; pain interference

With a growing focus in sports medicine on improving patients' quality of life, the use and importance of patient-reported outcome (PRO) instruments have grown. PROs have become a valuable adjunct to functional testing to measure the success and efficacy of orthopaedic procedures.<sup>13,21</sup> PROs provide invaluable information that allow clinicians to objectively monitor athlete rehabilitation and return to preinjury status. However, PRO integration into clinical practice has proved to be challenging, as many of

these instruments result in significant patient burden. Furthermore, nonelectronic administration of lengthy surveys can result in significant burden to the office staff, thereby limiting the routine integration thereof into clinical practice.

An ideal PRO would capture features important to patients and be easy for the patient to complete, simple to interpret, reproducible, highly accurate, and specific to a measured domain.<sup>5,19</sup> To combat patient burden and standardize patient scores, in 2004, the US National Institutes of Health developed the Patient-Reported Outcomes Measurement Information System (PROMIS) with multiple health domains to provide both accurate and reliable PROs.

The Orthopaedic Journal of Sports Medicine, 9(8), 23259671211017162

DOI: 10.1177/23259671211017162

© The Author(s) 2021

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

Item-response theory was used for the development of a computer-adaptive testing (CAT) PROMIS platform. PROMIS can adjust each question based on previous responses, thereby minimizing survey length and patient burden.<sup>3,4,13,17,18</sup> The PROMIS CAT system continues to ask questions until the confidence interval for the actual PROMIS value is narrowed to a conventional default of SE < 3.0 (which correlates with a reliability of 0.91) or a maximum of 12 questions is reached.<sup>15</sup> Each PROMIS CAT takes only 1 to 2 minutes to complete.<sup>2,18</sup> PROMIS-standardized scoring allows quick and simple interpretation by clinicians with all PROMIS scores centered on a population-normalized mean of 50 points and SD of 10 points.<sup>11</sup> The traditional variation-based estimates to calculate minimally clinically important difference (MCID) (ie,  $\frac{1}{2}$  SD method<sup>16,20</sup>) would suggest an MCID of 5 in the general population. However, the MCID of PROMIS may depend on the population in question.<sup>8</sup> In addition to the benefits of computer administration and length of survey, PROMIS domains have been validated and shown to exhibit fewer ceiling effects than many other PROs in healthy and impaired populations.<sup>13-15</sup>

PROMIS has increasingly been applied in sports medicine, including populations with anterior cruciate ligament tears,<sup>19</sup> foot and ankle trauma,<sup>10</sup> and rotator cuff injuries.<sup>13</sup> However, the performance of PROMIS in a young adult athletic population may be different from the general population, but these differences are less apparent in the setting of injury or pain. The characteristics and performance of PROMIS in a young athletic population have not been well-studied.<sup>9</sup> For example, the PROMIS physical function domain may have a higher normal (mean >50) score in a young adult athletic population than in the overall population (mean PROMIS score of 50). An understanding of the normal PROMIS scores in this population is important to evaluate the results of treatment of sports-related injuries and conditions, to define return to normal function, and to establish cohort-specific norms for research purposes. Additionally, ceiling or floor effects in a healthy athletic population may limit the discriminative ability of PROMIS to accurately follow these patients over the course of their rehabilitation<sup>6</sup> and would be more apparent in an asymptomatic population.

The purpose of this study was 3-fold: (1) to define normative PROMIS scores for the domains of Physical Function (PF-CAT), Mobility (M-CAT), Upper Extremity Function (UE-CAT), and Pain Interference (PI-CAT) of young healthy collegiate athletes; (2) to determine the frequency

of floor and ceiling effects in this population; and (3) to determine if normal PROMIS scores in this population differ by age or sex.

## METHODS

A cross-sectional cohort study of PROMIS scores for the PF-CAT, M-CAT, UE-CAT, and PI-CAT was performed in healthy, collegiate-level athletes. Before prospective data collection, institutional review board approval was granted. The study was performed at a single National Collegiate Athletic Association (NCAA) Division III university, and study participants were collegiate athletes enrolled during the university's annual preseason physical examination appointments, which are required for each athlete. Participants were included in the study if they provided consent, were not injured at time of the examination (defined as disability preventing practice or game participation), completed all 4 PROMIS domains, were between the ages of 18 and 23 years, and competed in at least 1 collegiate sport. Adult PROMIS domains are validated in a population 18 years of age or older. No compensation was given for survey participation.

A total of 262 athletes between the ages of 18 and 23 years were invited to complete the survey, with 213 (81.3%) agreeing to participate and completing the full survey. Nineteen (7.3%) of the 262 athletes were excluded because of current injury. In total, 194 athletes completed all PROMIS domains. The age breakdown for participants 18, 19, 20, and 21 years and over was 88 (45.4%), 47 (24.2%), 26 (13.4%), and 33 (17%), respectively.

The study survey was administered electronically on tablets or smart phones via a web-based research application (Research Electronic Data Capture; Vanderbilt University). Those who consented to be a part of the study were asked characteristic information including age, sex, NCAA primary sport, and secondary sport(s), followed by PROMIS CAT domains including PF-CAT (V2.0), M-CAT (V2.0), UE-CAT (V2.0), and PI-CAT (V1.1). For each domain, the number of questions was recorded to analyze patient burden, in addition to the time to complete (seconds). Outliers who took longer than 5 minutes (300 seconds) were excluded from time calculation because of high likelihood of interruption in the survey contributing to the length of time to complete. A total of 25 (3.2%) individual times were removed from analysis, as the individuals took longer than 5 minutes to complete the survey. The

<sup>||</sup>Address correspondence to Jeffrey J. Nepple, MD, Washington University School of Medicine, Campus Box 8233, 660 Euclid Avenue, St. Louis, MO 63110, USA (email: nepplej@wustl.edu).

<sup>\*</sup>Department of Orthopaedic Surgery, Miller School of Medicine, Miami, Florida, USA.

<sup>†</sup>Campbell Clinic Orthopaedics, University of Tennessee Health Science Center, Memphis, Tennessee, USA.

<sup>‡</sup>Department of Orthopaedic Surgery, Washington University in St Louis, St Louis, Missouri, USA.

<sup>§</sup>Department of Orthopaedic Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, USA.

Final revision submitted January 6, 2021; accepted January 25, 2021.

One or more of the authors has declared the following potential conflict of interest or source of funding: D.B. has received educational support from Smith & Nephew and Arthrex. R.W.W. is a paid consultant for Responsive Arthroscopy. M.J.M. has received consulting fees from Arthrex and Heron Therapeutics and educational support from Elite Orthopedics. J.J.N. has received consulting fees from Smith & Nephew, Ceterix, and Responsive Arthroscopy and research funding from Smith & Nephew, Stryker, and Zimmer Biomet. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from Washington University in St Louis, Missouri (ID No. 201708021).

percentage of participants completing the maximum number of questions (12) was also recorded. This would indicate the percentage of participants for which the PROMIS domain could not accurately reach a mean error  $<3.0$ . In this instance, the survey was stopped since the upper limit of patient burden was reached.

Ceiling and floor effects were assessed by the percentage of participants with the highest or lowest possible score on a domain. A significant ceiling or floor effect was present when greater than 15% of the participants met this criterion.<sup>8,13,19</sup> The maximum and minimum for each PROMIS domain was determined by completing each test and answering with all maximal positive or negative responses. The minimum/maximum values for PROMIS PF-CAT, M-CAT, UE-CAT, and PI-CAT were 14.7/75.6, 18.2/60.2, 14.7/61, and 38.7/83.8, respectively.

### Statistical Analysis

Descriptive statistical analysis was performed to determine the cohort-specific mean and median as well as standard deviation of each PROMIS domain. Statistical testing was performed utilizing SPSS (Version 22.0; IBM). Normality was assessed by visual inspection of histograms, Q-Q plots, and the Shapiro-Wilk test. The distribution of all PROMIS domains in this population was found to be nonnormal. The nonparametric 1-sample Wilcoxon signed-rank test was used to determine if the PROMIS values differed within each domain from the general population-based mean of 50. The nonparametric 2-independent samples Mann-Whitney *U* test was utilized to assess the role of age (grouped as  $<18$  vs  $>19$  years) and sex. Level of significance was established at a 2-sided alpha level of  $P < .05$ .

### RESULTS

A total of 194 completed surveys were included in the study, consisting of 118 (60.8%) men and 76 (39.2%) women. The mean age of the participants was  $19.1 \pm 1.2$  years. Primary sports included cross-country (23%;  $n = 44$ ), soccer (21%;  $n = 41$ ), football (20%;  $n = 39$ ), track and field (12%;  $n = 24$ ), swimming and diving (8%;  $n = 16$ ), baseball/softball (6%;  $n = 12$ ), volleyball (5%;  $n = 10$ ), basketball (3%;  $n = 5$ ), and cheerleading (2%;  $n = 3$ ). Additionally, 68 (35%) participants reported playing a secondary sport, the most common of which being track and field (72%;  $n = 49$ ).

The mean  $\pm$  SD (median) PROMIS scores were as follows: PF-CAT:  $62.9 \pm 6.7$  (62.4), M-CAT:  $58.2 \pm 4.1$  (60.2), UE-CAT:  $57.4 \pm 5.8$  (61.0), and PI-CAT:  $43.2 \pm 6.2$  (38.7). The higher the score of each individual PROMIS section correlates with a higher degree of function in each category except PI-CAT, in which a lower score indicates less pain. The median for each domain was significantly different from the population mean of 50 points ( $P < .001$  for all). The PF-CAT domain differed the most with a mean (median) score of 62.9 (62.4), which was more than 1 SD above the population-based mean of 50. Most healthy athletes had a score  $>50$  for each PROMIS domain (PF-CAT, 97.4%; M-CAT, 94.3%; UE-CAT, 88.1%; and PI-CAT,

71.1%). Figure 1 shows the distribution of each PROMIS domain in this population compared with the established population-based normal curve.

On average, the UE-CAT domain required the largest number of questions to determine the participants' score, followed by the M-CAT, PI-CAT, and PF-CAT domains (Table 1). Likewise, this same trend was seen when examining the percentage of participants answering the maximum number of questions on each domain (12). Time to complete each domain (Table 1) was a mean of between 62.5 and 88.6 seconds.

There were no significant differences according to participant age regarding any of the PROMIS domain scores ( $P > .05$  for all). Similarly, no significant differences ( $P > .05$  for all) were present in PROMIS domain scores when stratified by sex (Table 2).

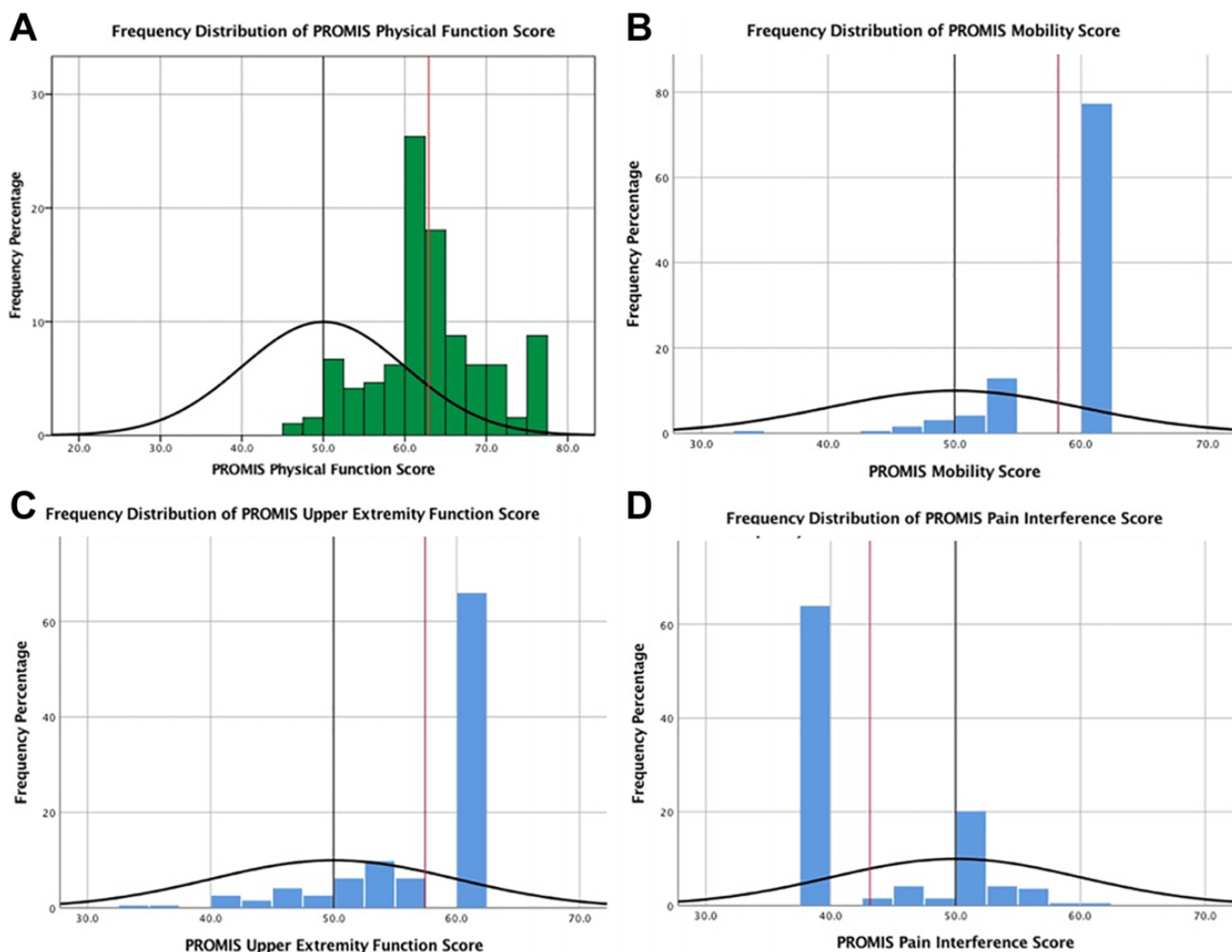
### Ceiling and Floor Effects

M-CAT and UE-CAT domain scores had strong ceiling effects, with maximal scores in 77.3% and 66% of the athletes, respectively (Table 3). Similarly, 63.9% of the athletes exhibited a floor effect for PI-CAT. However, PF-CAT scores were not found to be affected by a ceiling effect, as only 8.8% of the participants exhibited the highest or lowest scores in this domain. In terms of ceiling effect by sex (male/female), neither M-CAT (73%/84%) nor UE-CAT (69%/62%) showed any differences by sex, and again, a floor effect was found in both sexes in the PI-CAT (69%/55%) domain (all  $P > .05$ ). The PF-CAT did not display a ceiling effect in either sex.

### DISCUSSION

PROMIS domains have created a simple, time-efficient way to track meaningful PROs in orthopaedic practices and are increasingly utilized in sports medicine. In the current study, we demonstrated that the PROMIS PF-CAT, M-CAT, UE-CAT, and PI-CAT are significantly different in young, healthy collegiate athletes, compared with established adult population norms. The magnitude of differences for all 4 domains exceeds the classically established MCID,<sup>16,20</sup> indicating that the magnitude of difference is likely clinically relevant. In a healthy, young adult athletic population, the PF-CAT domain appears to have an acceptably low rate of ceiling effects, while PROMIS domains including M-CAT and UE-CAT appear to have unacceptably high rates of ceiling effects that limit their utility in this population.

The lack of floor and ceiling effects is a significant proposed advantage of PROMIS compared with some legacy measures.<sup>7,8,19</sup> PROMIS domains have demonstrated small floor and ceiling effects in the general population. However, a highly functional healthy athletic population may experience more floor and ceiling effects than the general population, thus limiting the tool's applicability for certain patient cohorts. Thus, it should be noted that the prevalent definition of ceiling and floor effects used in the literature has limitations. The definition used most frequently is a



**Figure 1.** Distribution of PROMIS domains in healthy athletes (bars) compared with established distributions in the general adult population (normal curve) for (A) physical function, (B) mobility, (C) upper extremity function, and (D) pain interference. The population mean (50) is shown by the black vertical line, and the mean for healthy athletes is shown as a red vertical line for each domain. PROMIS, Patient-Reported Outcomes Measurements Information System.

**TABLE 1**  
Patient Burden by PROMIS Domain<sup>a</sup>

PROMIS Domain	Total Questions in REDCap CAT Bank <sup>b</sup>	Number of Questions Answered, Mean ± SD (range)	Participants Answering Maximum Number of Questions (12), n (%)	Time to Complete, s, Mean ± SD
PF-CAT	165	8.8 ± 2.9 (4-12)	65 (33.5)	68.2 ± 34.8
M-CAT	15	11.3 ± 1.9 (4-12)	169 (87.1)	88.6 ± 48.7
UE-CAT	46	11.5 ± 1.6 (4-12)	174 (89.7)	62.5 ± 50.1
PI-CAT	40	9.3 ± 3.7 (4-12)	126 (65)	74.4 ± 49.2

<sup>a</sup>CAT, computer-adaptive testing; M, mobility; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcomes Measurements Information System; UE, upper extremity function.

<sup>b</sup>Total questions in the bank indicates the total number of possible questions while, with CAT, each patient answered a subset of questions.

binary answer of presence versus absence based on >15% of the individuals reaching the maximum (ceiling) or minimum (floor) score on an instrument.<sup>7</sup> This definition,

however, does not allow for quantification of the degree of a ceiling or floor effect on a specific cohort. As shown in this study, PF-CAT demonstrated an acceptable ceiling effect

TABLE 2  
PROMIS Domain Scores by Age and Sex<sup>a</sup>

PROMIS Domain	Age, y				Sex	
	18	19	20	≥21 <sup>b</sup>	Male	Female
PF-CAT	62.9 ± 5.8 (62.1)	62.2 ± 7.4 (62.1)	62.9 ± 7.7 (63.9)	63.9 ± 8.0 (63.2)	63.3 ± 6.8 (62.9)	62.3 ± 6.5 (62.1)
M-CAT	58.5 ± 3.5 (60.2)	57.6 ± 5.4 (60.2)	58.3 ± 3.8 (60.2)	58.3 ± 4.0 (60.2)	57.9 ± 4.4 (60.2)	58.7 ± 3.6 (60.2)
UE-CAT	56.9 ± 5.7 (61.0)	57.9 ± 6.3 (61.0)	57.5 ± 6.4 (61.0)	58.2 ± 5.0 (61.0)	57.6 ± 6.0 (61.0)	57.1 ± 5.6 (61.0)
PI-CAT	42.6 ± 5.8 (38.7)	43.3 ± 6.8 (38.7)	42.9 ± 5.6 (38.7)	44.5 ± 7.1 (38.7)	42.5 ± 6.0 (38.7)	44.2 ± 6.5 (38.7)

<sup>a</sup>Data are reported as mean ± SD (median). CAT, computer-adaptive testing; M, mobility; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcomes Measurements Information System; UE, upper extremity function.

<sup>b</sup>Participants 21 years and older were grouped together for analysis owing to fewer participants in this age group.

TABLE 3  
Ceiling and Floor Effects by PROMIS Domain<sup>a</sup>

PROMIS Domain	Score			Ceiling or Floor Effect <sup>b</sup>
	Mean ± SD (Median)	Mean ± SE	Range	
PF-CAT	62.9 ± 6.7 (62.4)	3.1 ± 0.6	14.7-75.6	8.8%
M-CAT	58.2 ± 4.1 (60.2)	5.2 ± 1.3	18.2-60.2	77.3% <sup>c</sup>
UE-CAT	57.4 ± 5.8 (61.0)	5.4 ± 1.3	14.7-61.0	66% <sup>c</sup>
PI-CAT	43.2 ± 6.2 (38.7)	4.4 ± 1.6	38.7-83.8	63.9% <sup>c</sup>

<sup>a</sup>CAT, computer-adaptive testing; M, mobility; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcomes Measurements Information System; UE, upper extremity function.

<sup>b</sup>Floor effect was only observed for the PI-CAT.

<sup>c</sup>Significant ceiling/floor observed if >15% of the participants reached maximum or minimum achievable value, respectively.

rate of 8.8%. However, a significant ceiling effect for UE-CAT (66.0%) and M-CAT (77.3%), as well as floor effects for PI-CAT (63.9%), was found to be present in a healthy population. The non-normal distribution and ceiling/floor effects of these domains remain important to recognize.

Interestingly, PF-CAT was the only PROMIS domain that lacked both ceiling and floor effects in our study. This is consistent with other studies<sup>1,2,11,12,18,19</sup> that have shown similar findings for PF-CAT in other populations. Indeed, despite being the most common domain assessed, not a single study has found PROMIS PF-CAT to have a substantial ceiling (0%-9%) or floor (0%-3.9%) in any population studied (8.8% in the current study).<sup>7</sup> In a similar population of collegiate athletes, Madsen et al<sup>12</sup> reported similar findings: PF-CAT scores (61.7 ± 6.0 vs current study 62.9 ± 6.7) and a similar lack of ceiling/floor effect. It is worth noting that PROMIS PF-CAT's questions appear to be highly relevant to athletic activities, unlike domains of UE-CAT or M-CAT, which demonstrated unacceptably high ceiling effects. PF-CAT attempts to qualify an athlete's ability to run both long and short distances, jump rope, show endurance for swimming or dance, and so forth, rather than activity from a single sport. Conversely, UE-CAT and M-CAT ask questions that are less athletically demanding (eg, "Does

your health now limit you in climbing one flight of stairs?" "Are you able to stand for one hour?" "Are you able to run a short distance, such as to catch a bus?").

A recent systematic review reported PROMIS UE-CAT to have variable ceiling effects (0%-28.2%) depending on the version of UE-CAT used. Previous studies<sup>1,2,14</sup> noted that the UE-CAT domain had a significant ceiling effect even in a nonhealthy population. Previously, Patterson et al<sup>14</sup> found that on the UE-CAT (Version 1.0), the maximum achievable score from their cohort of 164 patients with shoulder injuries was 56, in line with previous studies<sup>16,21</sup> using a similar CAT. With the ceiling being <1 SD above the reference population, this domain appeared to lack generalizability in the normal population. More recently, a newer version (2.0) attempted to correct the ceiling effect; however, our results using Version 2.0 showed that 66% of the healthy participants still achieved the ceiling score (mean, 61 points). Our findings of the significant ceiling effect in the UE-CAT domain (2.0) are likely indicative of a lack of specificity for our younger, healthier athletes and limit its utility in this population.

Similarly, we demonstrated unacceptable ceiling effects for the M-CAT domain in this population (77%). The extent of ceiling/floors for M-CAT in the literature has not been established. Our findings of the significant ceiling effect in the M-CAT domain are likely indicative of a lack of specificity for our younger, healthier athletes and limit its utility in this population.

The PI-CAT has been shown to have significant floor effects (ranging from 0% to 19.0% in the literature) and no ceiling effects (0-4.7%) in symptomatic populations. In the current study, we demonstrated a floor effect of 64% for the PI-CAT domain. However, we feel this floor effect is acceptable compared with domains assessing function. The floor effect found in PI-CAT is likely due to many athletes truly having minimal to no pain despite participation in high-level athletics. In the current study, we excluded injured athletes. Indeed, other orthopaedic pain models, such as the visual analog scale pain scale, have similar floor effects (0/10 pain) but are still commonplace in practice.

Furthermore, decreased patient burden is another advantage of PROMIS CAT compared with established legacy measures, in that PROMIS CAT requires only approximately 4 to 12 questions to complete.<sup>11,19</sup> In the current study, we demonstrate mean times to completion of 1-1.5

minutes per domain. Specifically, the PF-CAT has exhibited a low patient burden,<sup>1,11,19</sup> requiring a mean of only 4.06 questions in a cohort of hip preservation patients compared with the international Hip Outcome Tool (12 questions), modified Harris Hip Score (8 questions), and Hip Outcome Score (26 questions). Likewise, similar studies<sup>1,4,11,18,19</sup> evaluating shoulder and knee conditions showed a lack of floor and ceiling effects for PF-CAT with only approximately 4 questions being asked. UE-CAT and PI-CAT have also been shown to have a relatively low patient burden in the orthopaedic population,<sup>2</sup> with a mean time to completion/number of questions required of 69 seconds/4.8 and 43 seconds/4.3, respectively. However, the current study demonstrated that a significantly greater question burden is present with the use of PROMIS in a healthy athletic population that is likely similar to that of a fully recovered athlete after an injury. The current study population answered a mean of 10.2 questions per domain (Table 1), with the majority of participants completing the maximum number of questions possible (12). Therefore, approximately 2 to 3 times as many questions are needed to complete the CAT in a high-functioning athlete. The time to complete data seems to suggest that while more questions were answered by the young athletic population, the time burden is still relatively low compared with the general population estimate of 1 to 2 minutes.<sup>2,18</sup> Furthermore, even with the increased number of questions, a previous study<sup>19</sup> of patients that have since approached return to their baseline have shown that PROMIS is still less burdensome than established PROs.

This study is not without limitations. Our sampling strategy, which involved the convenience of available athletes, led to an unequal distribution of age and sex among our participants as well as a skewed prevalence of sports participation. Yet, the current study demonstrates no apparent differences by age, sex, or sport. Additionally, this study sample of nearly 200 participants came from a single Midwest Division III university and may not be representative of other collegiate athletes in different geographic regions or level of competition, although any differences are likely small. However, Division III schools outnumber their Division I and II counterparts in the NCAA. Therefore, their student-athletes would potentially be the most representative of a healthy young adult athletic population. It is possible that Division I and II athletes could demonstrate PROMIS functional scores even farther from the population mean; however, the literature has not shown this to be the case based on comparison with Madsen et al.<sup>12</sup> Finally, the fact that only 81% of those participants invited actually completed the survey may have introduced selection bias into our results. However, this is to some degree unavoidable, and the relatively high rate of participation minimizes this potential bias.

## CONCLUSION

PROMIS has been shown to be valid and reliable in quantitatively evaluating multiple health domains of the general adult population. The current study shows that

healthy collegiate athletes scored nearly 1 SD from population-based means for all of the PROMIS domains tested. M-CAT and UE-CAT demonstrated significant ceiling and floor effects in more than two-thirds of healthy athletes that limited their utility in this population. The PF-CAT domain demonstrated an acceptable rate of ceiling effects while also having the largest difference between the healthy athletic population and the general population mean. The PROMIS PI-CAT likely has utility in this population, despite high floor rates, due to minimal pain in many healthy athletes. Therefore, these findings must be strongly considered when using PROMIS to evaluate functional outcomes in a healthy active cohort.

## REFERENCES

1. Anthony CA, Glass NA, Hancock K, Bollier M, Wolf BR, Hettrich CM. Performance of PROMIS instruments in patients with shoulder instability. *Am J Sports Med.* 2017;45(2):449-453.
2. Bernstein DN, Houck JR, Hammert WC. A comparison of PROMIS UE versus PF: correlation to PROMIS PI and depression, ceiling and floor effects, and time to completion. *J Hand Surg Am.* 2019;44(10):901.
3. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res.* 2007;16(suppl 1):133-141.
4. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care.* 2007;45(5)(suppl 1):S3-S11.
5. Deshpande PR, Rajan S, Sudeepthi BL, Abdul Nazir CP. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res.* 2011;2(4):137-144.
6. Fries J, Rose M, Krishnan E. The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. *J Rheumatol.* 2011;38(8):1759-1764.
7. Gausden EB, Levack A, Nwachukwu BU, Sin D, Wellman DS, Lorich DG. Computerized adaptive testing for patient reported outcomes in ankle fracture surgery. *Foot Ankle Int.* 2018;39(10):1192-1198.
8. Gullledge CM, Lizzio VA, Smith DG, Guo E, Makhni EC. What are the floor and ceiling effects of Patient-Reported Outcomes Measurement Information System computer adaptive test domains in orthopaedic patients? A systematic review. *Arthroscopy.* 2020;36(3):901-912. e907.
9. Hancock KJ, Glass N, Anthony CA, et al. Performance of PROMIS for healthy patients undergoing meniscal surgery. *J Bone Joint Surg Am.* 2017;99(11):954-958.
10. Hung M, Voss MW, Bounsanga J, Crum AB, Tyser AR. Examination of the PROMIS upper extremity item bank. *J Hand Ther.* 2017;30(4):485-490.
11. Kollmorgen RC, Hutyra CA, Green C, Lewis B, Olson SA, Mather RC III. Relationship between PROMIS computer adaptive tests and legacy hip measures among patients presenting to a tertiary care hip preservation center. *Am J Sports Med.* 2019;47(4):876-884.
12. Madsen LP, Evans TA, Snyder KR, Docherty CL. Patient-Reported Outcomes Measurement Information System physical function item bank, version 1.0: physical function assessment for athletic patient populations. *J Athl Train.* 2016;51(9):727-732.
13. Nicholson AD, Kassam HF, Pan SD, Berman JE, Blaine TA, Kovacevic D. Performance of PROMIS Global-10 compared with legacy instruments for rotator cuff disease. *Am J Sports Med.* 2019;47(1):181-188.
14. Patterson BM, Orvets ND, Aleem AW, et al. Correlation of Patient-Reported Outcomes Measurement Information System (PROMIS) scores with legacy patient-reported outcome scores in patients undergoing rotator cuff repair. *J Shoulder Elbow Surg.* 2018; 27(6)(suppl):S17-S23.

15. PROMIS Instrument Development and Validation Scientific Standards Version 20 (2013, May) Retrieved March 16, 2019. [http://www.healthmeasures.net/images/PROMIS/PROMISStandards\\_Vers20\\_Final.pdf](http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers20_Final.pdf)
16. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102-109.
17. Revicki DA, Cella DF. Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. *Qual Life Res*. 1997;6(6):595-600.
18. Robins RJ, Anderson MB, Zhang Y, Presson AP, Burks RT, Greis PE. Convergent validity of the Patient-Reported Outcomes Measurement Information System's physical function computerized adaptive test for the knee and shoulder injury sports medicine patient population. *Arthroscopy*. 2017;33(3):608-616.
19. Scott EJ, Westermann R, Glass NA, Hettrich C, Wolf BR, Bollier MJ. Performance of the PROMIS in patients after anterior cruciate ligament reconstruction. *Orthop J Sports Med*. 2018;6(5):2325967118774509.
20. Sloan JA. Assessing the minimally clinically significant difference: scientific considerations, challenges and solutions. *COPD*. 2005;2(1):57-62.
21. Weldring T, Smith SM. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). *Health Serv Insights*. 2013;6:61-68.