


## Original Article

# Validation of the Calgary Postoperative Pain after Spine Surgery Score for Poor Postoperative Pain Control after Spine Surgery

Michael M.H. Yang<sup>1,2,3</sup> , Jay Riva-Cambrin<sup>1,3</sup>, Jonathan Cunningham<sup>1</sup> and Steven Casha<sup>1,4</sup>

<sup>1</sup>Department of Clinical Neurosciences, Section of Neurosurgery, University of Calgary, Calgary, AB, Canada, <sup>2</sup>O'Brien Institute for Public Health, University of Calgary, Calgary, AB, Canada, <sup>3</sup>Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada and <sup>4</sup>Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada

**ABSTRACT: Objective:** The Calgary Postoperative Pain after Spine Surgery (CAPPS) score was developed to identify patients at risk of experiencing poorly controlled pain after spine surgery. The goal of this study was to independently validate the CAPPS score on a prospectively collected patient sample. **Methods:** Poor postoperative pain control was defined as a mean numeric rating scale (NRS) for pain >4 at rest in the first 24 hours after surgery. Baseline characteristics in this study (validation cohort) were compared to those of the development cohort used to create the CAPPS score. Predictive performance of the CAPPS score was assessed by the area under the curve (AUC) and percentage misclassification for discrimination. A graphical comparison between predicted probability vs. observed incidence of poorly controlled pain was performed for calibration. **Results:** Fifty-two percent of 201 patients experienced poorly controlled pain. The validation cohort exhibited lower depression scores and a higher proportion using daily opioid medications compared to the development cohort. The AUC was 0.74 [95% CI = 0.68–0.81] in the validation cohort compared to 0.73 [95% CI = 0.69–0.76] in the development cohort for the eight-tier CAPPS score. When stratified between the low- vs. extreme-risk and low- vs. high-risk groups, the percentage misclassification was 21.2% and 30.7% in the validation cohort, compared to 29.9% and 38.0% in the development cohort, respectively. The predicted probability closely mirrored the observed incidence of poor pain control across all scores. **Conclusions:** The CAPPS score, based on seven easily obtained and reliable prognostic variables, was validated using a prospectively collected, independent sample of patients.

**RÉSUMÉ : Validation du score CAPPS relatif au soulagement insuffisant de la douleur après une opération du rachis. Objectif :** Le score Calgary Postoperative Pain after Spine Surgery (CAPPS) a été conçu pour permettre le repérage des patients susceptibles d'éprouver un soulagement insuffisant de la douleur après une opération du rachis. L'étude avait pour but de valider le score CAPPS dans un échantillon distinct de patients, formé de façon prospective. **Méthode :** Un soulagement insuffisant de la douleur postopératoire a été défini comme une valeur moyenne > 4 au repos, au cours des 24 premières heures suivant l'opération, sur une échelle d'évaluation numérique de la douleur. Il y a eu comparaison des caractéristiques de base de la cohorte de validation, formée pour l'étude ici présentée, avec celles de la cohorte ayant servi à l'élaboration du score CAPPS, dite de référence. La performance prévisionnelle du score a été évaluée en fonction de la surface sous la courbe (SSC) et du pourcentage de classement erroné aux fins de différenciation. Par ailleurs, une comparaison graphique entre les risques prédits de soulagement insuffisant de la douleur et l'incidence observée a servi à l'étalonnage du score. **Résultats :** Des 201 patients opérés, 52 % ont éprouvé un soulagement insuffisant de la douleur. Dans la cohorte de validation, on a observé des scores plus faibles de dépression et une plus forte proportion de patients faisant un usage quotidien d'opioïdes que dans la cohorte de référence. La SSC était de 0,74 [IC à 95 % = 0,68–0,81] dans la cohorte de validation contre 0,73 [IC à 95 % = 0,69–0,76] dans la cohorte de référence au score CAPPS à huit niveaux. Une fois répartis les groupes à faible risque contre à risque extrême et à faible risque contre à risque élevé, le pourcentage de classement erroné s'élevait à 21,2 % et à 30,7 % dans la cohorte de validation, comparativement à 29,9 % et à 38,0 % dans la cohorte de référence, respectivement. Les risques prédits concordaient étroitement avec l'incidence observée du soulagement insuffisant de la douleur, et ce, à tous les niveaux du score. **Conclusion :** Le score CAPPS, qui repose sur sept variables pronostiques dignes de confiance, faciles à obtenir, a été validé d'après les résultats obtenus dans un échantillon distinct de patients, constitué de façon prospective.

**Keywords:** Postoperative pain; Spine surgery; Clinical prediction score; Pain measurement

(Received 14 January 2022; final revisions submitted 6 September 2022; date of acceptance 10 October 2022; First Published online 24 October 2022)

**Corresponding author:** Dr. Steven Casha, Department of Clinical Neurosciences, University of Calgary, 12th Floor, 29 Street NW, Calgary, AB, T2N 2T9, Canada. Email: [steve.casha@albertahealthservices.ca](mailto:steve.casha@albertahealthservices.ca)

**Cite this article:** Yang MMH, Riva-Cambrin J, Cunningham J, and Casha S. (2023) Validation of the Calgary Postoperative Pain after Spine Surgery Score for Poor Postoperative Pain Control after Spine Surgery. *The Canadian Journal of Neurological Sciences* 50: 687–693, <https://doi.org/10.1017/cjn.2022.305>

© The Author(s), 2022. Published by Cambridge University Press on behalf of Canadian Neurological Sciences Federation. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

## Introduction

In a study of 179 surgical procedures, spinal fusion was identified as the second most painful procedure (following open calcaneus reduction) with a median numeric rating scale (NRS) for pain of 7/10 on the first day after surgery.<sup>1</sup> Further, it has been estimated that between 30 and 64% of patients undergoing a spinal operation experience poor pain control in the acute postoperative period.<sup>2</sup> In a study by Grob et al.,<sup>3</sup> 26% of patients identified poorly controlled postoperative pain as a complication at their 12-month follow-up after spine surgery. Poorly controlled postoperative pain has been associated with impaired mobility, prolonged length of hospital stay, thromboembolic diseases, respiratory complications, and development of chronic pain.<sup>4-6</sup> The high incidence of poorly controlled pain could be explained by the lack of personalized treatment strategies, limitation in understanding of patient and surgical risk factors, and limited preoperative education about expected pain after surgery.<sup>6</sup>

Prediction models for poor postoperative pain control have been developed for other surgical disciplines;<sup>7-9</sup> however, none are specific for spine surgery.<sup>6</sup> Further, the literature lacks risk prediction tools that facilitate risk stratification for poor postoperative pain control. A recent scoping review found no studies that calculated a personalized risk for poor pain control after surgery to inform patients, healthcare professionals, or both.<sup>10</sup> The ability to risk-stratify patients according to their risk profile allows for personalized healthcare interventions to improve outcomes.

We previously developed and internally validated the Calgary Postoperative Pain after Spine Surgery (CAPPS) score in an effort to stratify patients by their risk of experiencing poorly controlled pain after elective spine surgery.<sup>11</sup> The CAPPS score was based on seven prognostic factors: age <70 years, female sex, preoperative daily use of opioid medication, preoperative neck or low back pain >7 measured by the NRS, Patient Health Questionnaire-9 (PHQ-9) depression score  $\geq 10$ ,  $\geq 3$  motion segment surgery, and fusion surgery. A score between 0 and 13 categorized patient within an eight-tier or a more simplified three-tier scale with known predicted probability of poorly controlled postoperative pain.<sup>11</sup> The goal of this study was to independently validate the CAPPS score using a new, temporally independent cohort undergoing elective spine surgery at the same center.

## Methods

Ethics approval was provided by the University of Calgary's Conjoint Health Research Ethics Board, and all subjects provided informed consent. This study was reported according to the Transparent Reporting of a multivariable model for Individual Prognosis Or Diagnosis (TRIPOD) statement.<sup>12</sup>

### Study Population

Consecutive patients were enrolled prospectively at the Foothills Medical Center in Calgary, Alberta, between June 2019 and October 2020. That tertiary care center performs the vast majority of spine surgeries in southern Alberta with a catchment area of two million people. Data collection was suspended between March 2020 and August 2020 due to COVID-19 pandemic-related constraints on surgical and research activity.

### Eligibility Criteria

We used the same study eligibility criteria as those used to develop the CAPPS score.<sup>11</sup> Adult patients ( $\geq 18$  years) who underwent

elective spine surgery and required hospital admission for at least 24 hours were included. Patients were excluded if they received intraoperative intrathecal or postoperative epidural analgesia, required admission to the intensive care unit, had spine surgery for acute trauma, or underwent more than one operation on the same day. Only patients with complete data were included in the final analysis.

### Data Collection

Baseline patient characteristics collected included patient age, sex, smoking history, body mass index (BMI), date of surgery, chief complaint (back pain, neck pain, radiculopathy, myelopathy, neurogenic claudication, or other), principal pathology (disc herniation, degenerative disc disease, spinal stenosis, spondylolisthesis, deformity, tumor, or other), length of hospital stay, oral morphine equivalent dose (MED) in the first 24 hours after surgery, preoperative neck or back pain intensity measured by the NRS, frequency of preoperative opioid use, and Oswestry Disability Index (ODI) or the Neck Disability Index (NDI). Surgical characteristics collected included number of motion segment operation (i.e., number of levels), surgical approach (i.e., anterior, posterior, or anterior and posterior), fusion surgery, minimally invasive surgery, and revision surgery. A procedure was considered minimally invasive if soft tissue dilation rather than stripping dissection was performed. All case report forms were compared to patients' hospital records to ensure accuracy.

### Outcome

The primary outcome was poor postoperative pain control defined as mean NRS for pain  $>4$  at rest in the first 24 hours after surgery (postoperative day 1). The rationale for choosing this threshold was previously described.<sup>11</sup> The NRS records pain intensity on a 11-point scale (where 0 indicates no pain and 10 indicates worst pain possible). The postoperative NRS was recorded prospectively by clinical care nurses, without knowledge of the research objectives or patient risk factors, as part of usual patient care after surgery. Patients were asked while recumbent to verbally rate their pain by selecting an integer on a scale of 0–10 as defined above. While the frequency of pain assessments were not standardized, they were performed according to the *American Pain Society* guideline.<sup>13</sup> Pain assessments were performed more frequently when pain was poorly controlled, before the administration of any analgesic medications, and at least every 6 hours in line with institutional policy. Pain assessments were deferred during sleep or when off clinical units (e.g., for imaging studies).

### Prognostic Variables and the CAPPS Score

We previously found seven prognostic factors to be associated with poor pain control after elective spine surgery, and these were transformed to create the 14-point CAPPS score: (1) age <70 years, (2) female sex, (3) preoperative daily use of opioid medication, (4) preoperative axial neck or low back pain  $>7$  on NRS, (5) PHQ-9 depression score  $\geq 10$ , (6)  $\geq 3$  motion segment surgery, and (7) fusion surgery (Table 1).<sup>11</sup> These prognostic variables were prospectively collected by research coordinators after patient admission in the preoperative holding area. These research coordinators were excluded from subsequent outcomes data collection. The CAPPS score consists of an eight-tier and a simplified three-tier risk-based score. The three-tier CAPPS score trichotomizes patients into low-risk (scores 0–4), high-risk (scores 5–8), or

**Table 1:** The Calgary Postoperative Pain after Spine Surgery Score modified from Yang et al. (2020)<sup>11</sup>

Predictor	Score
Age <70 years	2
Female sex	2
Daily preoperative opioid use	3
Preoperative NRS neck/back pain>7	1
Moderate to severe depression (PHQ-9≥10)	1
≥3 motion segment operation	2
Fusion Surgery	2
<b>Total Score</b>	<b>0 to 13</b>

PHQ-9- patient health questionnaire-9.

extreme-risk (scores 9–13) groups. The planned and performed procedure were verified for all patients from the dictated operative records.

### Statistical Analysis

Descriptive statistics were used to summarize patient demographics and study characteristics. Baseline patient characteristics in this study cohort (validation cohort) were compared to those of the score’s model development cohort in our previous study (development cohort).<sup>11</sup> The two-sample student *t*-test and the chi-square test were used to compare continuous and categorical variables, respectively. One-way analysis of variance was used to compare oral MEDs across the three CAPPs groups.

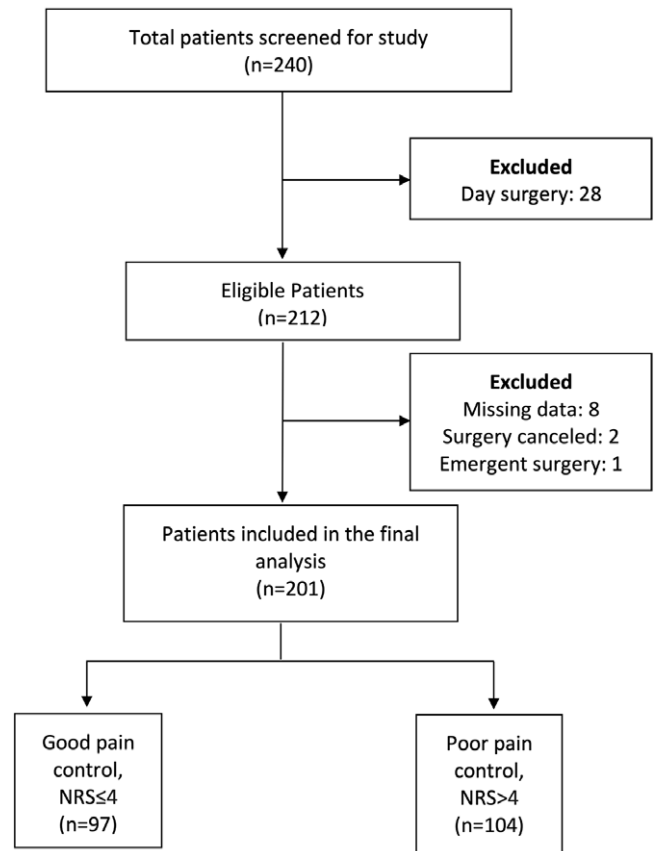
### Sample Size

Simulation studies have shown a minimum of 100 events (e.g., patients with poor pain control) are required to adequately validate prediction scores developed based on logistic regression models.<sup>14</sup> Using an estimated incidence of poor pain control of 57% after spine surgery,<sup>11</sup> a minimum sample size of 176 was required to adequately validate the CAPPs score. We elected to enroll an *a priori* determined conservative sample of 200 patients for this study.

### Validation of the CAPPs Score

All patients in this study were evaluated using the CAPPs score based on seven prognostic factors (Table 1). The sum of the numeric values determined by presence or absence of each dichotomized variable yielded a numeric score between 0 and 13. Patient’s prognostic factors were entered into our web-based CAPPs score calculator (<https://calgaryspine.ca/research/capps>) to determine the numeric sum for each patient and the associated predicted probability of poor pain control following elective spine surgery.<sup>11</sup> The sum value of the CAPPs score was evaluated for normality using the Shapiro–Wilk test.

The predictive performance of the CAPPs score on the validation cohort was quantified using measures of discrimination and calibration.<sup>15</sup> For calibration, the degree of agreement between the predicted and the observed incidence of poor pain control was compared graphically.<sup>16</sup> For discrimination (the ability to differentiate patients who do and do not experience poorly controlled pain), the area under the curve (AUC) of a receiver operating characteristic curve was used for the eight-tier CAPPs score, and the percentage misclassification with the low-risk group as the



**Figure 1:** Study flow diagram.

reference group was used for the three-tier CAPPs score.<sup>17</sup> Percentage misclassification was defined as the sum of false positives (CAPPs high- or extreme-risk group but observed to have good pain control) and false negatives (CAPPs low-risk group but observed to be in poor pain control) divided by the total sample, multiplied by 100%.

The level of significance was alpha = 0.05, and 95% confidence intervals (CIs) were reported. All statistical analyses were performed using STATA version 15.1.

## Results

### Patient Sample

Two-hundred forty patients were screened, and 201 met eligibility criteria (Figure 1). Excluded patients included 28 who underwent out-patient surgery, eight who had missing data, two for whom surgery was canceled, and one who underwent emergent surgery. The mean age was 58.6 years, and 52.7% were female (Table 2). The incidence of poor pain control after surgery was 51.7% [95% CI = 44.8%–58.6%], and the mean NRS for pain in the first 24 hours was 4.1 (standard deviation 2.0). Fifty-two percent underwent a fusion operation, 18.4% had a procedure including ≥3 motion segments (any part of two adjacent articulating vertebrae), 19.9% underwent a minimally invasive procedure, and 6.0% had both anterior and posterior surgery. The mean number of postoperative pain evaluations within the first 24 hours was 8.3 evaluations per patient (standard deviation 3.2). There was no difference in length of hospital stay between patients with mean NRS >4 (poor pain control) and NRS ≤4 (3.1 days vs. 3.0 days, *p* = 0.91). Patients with

**Table 2:** Baseline characteristics of original cohort used to develop the CAPPS score<sup>11</sup> (n=910) and the validation cohort from this study (n=201).

Characteristic	Development Cohort <sup>11</sup>	Validation Cohort	p-value
Age in years (mean ± SD)	59.8 ± 13.8	58.6 ± 12.9	0.26
Female sex (n, %)	442 (48.6)	106 (52.7)	0.29
Principal pathology (n, %)			
Disc herniation	124 (13.6)	25 (12.4)	<b>0.002</b>
Degenerative disc disease	111 (12.2)	28 (13.9)	
Spinal stenosis	360 (39.6)	106 (52.7)	
Spondylolisthesis	219 (24.1)	25 (12.4)	
Deformity	44 (4.8)	8 (4.0)	
Tumor	15 (1.7)	0 (0)	
Others	37 (4.1)	9 (4.5)	
Chief Complaint (n, %)			
Back pain	161 (17.7)	16 (8.0)	<b>&lt;0.001</b>
Neck pain	10 (1.1)	0 (0)	
Radiculopathy	379 (41.7)	104 (51.7)	
Myelopathy	130 (14.3)	41 (20.4)	
Neurogenic claudication	221 (24.3)	37 (18.4)	
Others	9 (1.0)	3 (1.5)	
Patients with NRS>4 for pain on postoperative day 1 (%; 95% CI)	56.5 (53.2-59.7)	51.7 (44.8-58.6)	0.22
Body mass index in kg/m <sup>2</sup> (mean ± SD)	29.2 ± 5.6	30.4 ± 7.2	<b>0.013</b>
Smoker (n, %)	174 (19.1)	52 (25.8)	<b>0.031</b>
Daily opioid medication (n, %)	277 (32.2)	92 (45.7)	<b>0.0003</b>
Preoperative neck or back pain measured by NRS (mean ± SD)	6.6 ± 2.5	6.0 ± 2.4	<b>0.002</b>
Depression on patient health questionnaire-9 (mean ± SD)	9.5 ± 6.2	7.9 ± 4.8	<b>&lt;0.001</b>
Severe preoperative disability on NDI or ODI <sup>a</sup> (n, %)	483 (56.7)	66 (33.3)	<b>&lt;0.001</b>
Surgical approach (n, %)			
Any anterior	202 (22.2)	55 (27.4)	0.14
Any posterior	669 (73.5)	134 (66.7)	
Any anterior and posterior	39 (4.3)	12 (6.0)	
Number of motion segment operation (n, %)			
1	562 (50.8)	111 (55.2)	0.98
2	260 (28.6)	53 (26.4)	
≥3	188 (20.7)	37 (18.4)	
Fusion surgery (n, %)	646 (60.0)	105 (52.2)	<b>0.043</b>
Minimally invasive surgery (n, %)	222 (24.4)	40 (19.9)	0.17
Revision surgery (n, %)	131 (14.4)	20 (10.0)	0.10

CAPPS- Calgary Postoperative Pain after Spine Surgery, SD- standard deviation, ODI- Oswestry disability, NDI- Neck disability index, NRS-Numeric rating scale.

<sup>a</sup>Neck Disability Index (NDI) ≥50 or Oswestry Disability Index (ODI) >40.

poor pain control consumed more oral MEDs than those who had good pain control (111.8 mg vs. 44.1 mg,  $p < 0.001$ ) in the first 24 hours after surgery. When stratified by the CAPPS score, there was a clear stepwise increase in the total oral MEDs consumed: 55.6 mg for low-risk ( $n = 59$ ), 76.7 mg for high-risk ( $n = 107$ ), and 126.5 mg for extreme-risk groups ( $n = 35$ ,  $p = 0.018$ ). A statistically significant difference in oral MEDs was observed between the low- and extreme-risk groups ( $p = 0.041$ ), but not between the low- and high-risk groups ( $p = 0.14$ ).

Several baseline patient and surgical characteristics seen in this study (validation cohort) differed from those seen in the sample used to develop the CAPPS score (development cohort) (Table 2). The validation cohort comprised of a higher proportion of patients with spinal stenosis (52.7% vs. 39.6%) and less with spondylolisthesis (12.4% vs. 24.1%). More patients in the validation cohort complained of radiculopathic (51.7% vs. 41.7%) and myelopathic (20.4 vs. 14.3%) symptoms, and a lower proportion reported symptoms of neurogenic claudication (18.4% vs.

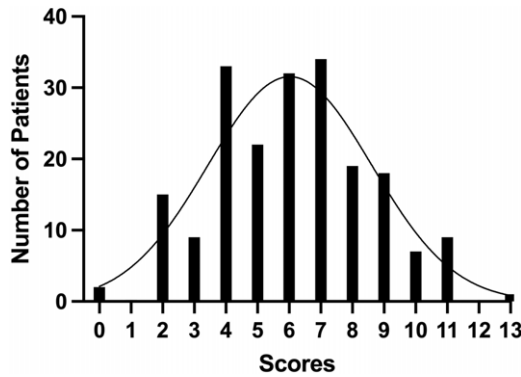


Figure 2: Distribution of CAPPs scores in the validation cohort (n = 201).

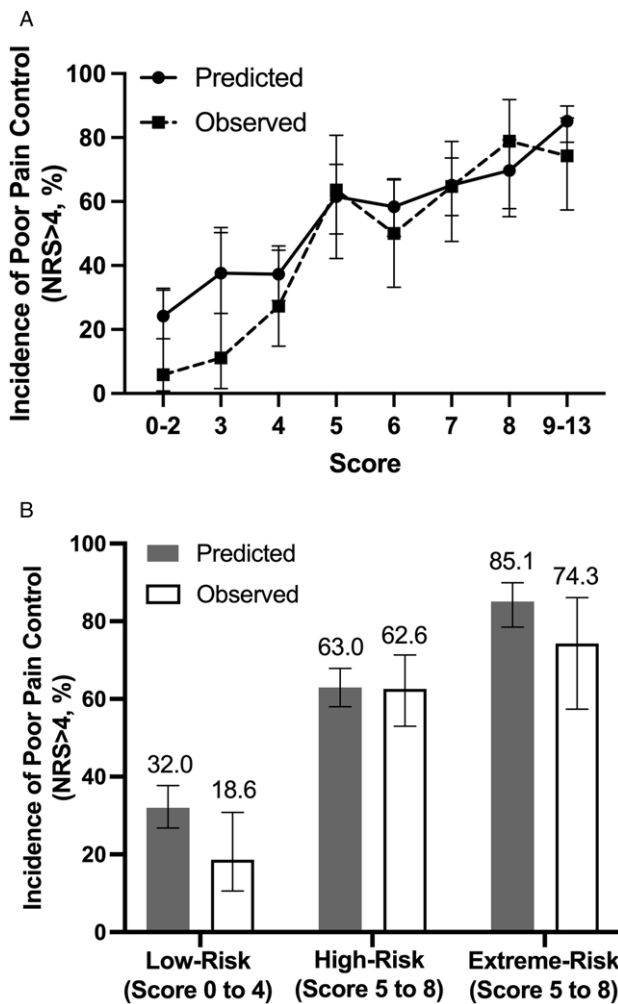


Figure 3: Predicted probability vs. observed incidence of poor pain control (numeric rating scale for pain >4) after elective spine surgery. A) Probabilities across the eight-tier CAPPs score and B) three-tier CAPPs score. Error bars represent 95% confidence intervals.

24.3%). Patients in the validation cohort were more likely to be smokers (25.8% vs. 19.1%,  $p = 0.031$ ), used opioid medications daily before surgery (45.7% vs. 32.2%,  $p = 0.003$ ), have higher BMI (30.4 vs. 29.2,  $p = 0.013$ ), had lower baseline PHQ-9 depression scores (7.9 vs. 9.5,  $p < 0.001$ ), and fewer had severe preoperative ODI and NDI scores (33.3% vs. 56.7%,  $p < 0.001$ ). Fewer patients underwent fusion surgery (52.2% vs. 60.0%,  $p = 0.043$ ).

Table 3: Discrimination of the Calgary Postoperative Pain After Spine Surgery (CAPPs) Score.

Eight-Tier CAPPs Score	Area Under the Curve (95% CI)		
	0.74 (0.68-0.81)		
Three-Tier CAPPs Score	% Misclassification	Sensitivity (95% CI)	Specificity (95% CI)
Low vs. Extreme Risk	21.2	70.3 (53.0-84.1)	84.2 (72.1-92.5)
Low vs. High Risk	30.7	85.9 (76.2-92.7)	54.5 (43.6-65.2)
Low vs. Extreme or High Risk	29.8	89.4 (81.9-94.6)	49.5 (39.2-59.8)

Low-Risk= scores 0 to 4, High-Risk= scores 5 to 8, Extreme-Risk= scores 9 to 13; CI= confidence intervals

Predictive Performance of the CAPPs Score

A histogram depicting CAPPs scores in the validation cohort is shown in Figure 2. The scores followed a normal distribution (skewness 0.11, kurtosis 2.7, and Shapiro-Wilk  $p = 0.94$ ). There was a stepwise increase in the observed incidence of poor pain control from 5.9% for scores 0–2 to 74.3% for scores 9–13 (Figure 3). Calibration was assessed by graphically comparing the predicted and observed incidence of poor pain control after spine surgery. The predicted probability matched the observed incidence of poor pain control for all tiers of the eight- and three-tier CAPPs score with overlapping 95% CIs (Figure 3).

Discrimination was assessed using the area under the receiver operating characteristic curve (AUC) and percentage misclassification. The eight-tier CAPPs score exhibited an AUC of 0.74 [95% CI = 0.68–0.81], suggesting the score was able to discriminate those with poor (NRS >4) from those with good pain control (NRS ≤4) (Table 2). Using the three-tier CAPPs score, the percentage misclassification between the low- (scores 0–4) and extreme-risk categories (scores 9–13) was 21.2% (sensitivity 70.3%, specificity 84.2%), and between low- and high-risk categories (scores 5–8) was 30.7% (sensitivity 85.9% and specificity 54.5%) (Table 3). When the CAPPs score was dichotomized between the low-risk group (scores 0–4) and high- or extreme-risk groups (scores >4), the percentage misclassification was 29.8% (sensitivity 89.4% and specificity 49.5%) (Table 3).

Discussion

Pain after spine surgery is universal and expected. However, poorly controlled pain can lead to patient dissatisfaction and poor outcomes. In this study, the CAPPs score was validated on an independent sample with noted differences from the development cohort (Table 2), further strengthening the score’s generalizability. The CAPPs score can be used in the preoperative setting to accurately identify patients at risk of experiencing poorly controlled pain, to facilitate patient education, and to inform the development of clinical care pathways to improve postoperative acute pain outcomes.

In our original study,<sup>11</sup> the CAPPs score was developed using a random 70% sample of the collected data and internally validated using the remaining 30%. Randomly splitting the data merely creates two comparable datasets, which could lead to inflated initial

estimates of predictive performance that are overoptimistic.<sup>15</sup> Validation on an independent sample evaluates the score's stability, reproducibility, and generalizability.<sup>12</sup> In this validation study, the CAPPs score was well calibrated for patients with scores between 5 and 8 (high-risk group) with predicted and observed incidence of poor pain control of 63.0% and 62.6%, respectively. However, the CAPPs score appeared to overestimate the probability of poorly controlled pain for patients who scored between 0 and 4 (low-risk group) and 9 to 13 (extreme-risk group), although these differences were not statistically significant. These differences may be explained by the smaller number of patients in these groups (low-risk group:  $n = 57$  [29.4%]; extreme-risk group:  $n = 35$  [17.4%]), and thus more random error in the estimated incidence of poorly controlled pain. Other reasons for the observed discrepancies could relate to differences in baseline patient characteristics, and the distribution of prognostic values between the validation and the development cohorts.<sup>18</sup> Collection of data immediately before surgery as opposed to during preoperative consultation (as had been done for the development cohort) may have also contributed to the differences observed. Nevertheless, the overestimation of poor pain control is preferable to underestimation.

For discrimination, the AUC for the eight-tier CAPPs score was 0.74. This means 74% of randomly selected patients in the validation cohort who had poor pain control (NRS > 4) after surgery had a higher score compared to that of a patient who experienced good pain control (NRS ≤ 4). The discriminatory performance of the eight-tier CAPPs score in this study (AUC = 0.74 [95% CI = 0.68–0.81]) was similar to that seen in development cohort (AUC = 0.73 [95% CI = 0.69–0.76]).<sup>11</sup> The percentage misclassification of the three-tier CAPPs score in the validation cohort was lower compared to the development cohort (21.2% vs. 29.9% in the low vs. extreme-risk group, and 30.7% vs. 38.0% in the low- vs. high-risk group).<sup>11</sup> There is no consensus on what constitutes an adequate AUC or percentage misclassification for the prediction of poor postoperative pain. To provide context, a study by Tighe et al (2015)<sup>19</sup> used machine learning algorithms to predict moderate to severe pain on postoperative day 1 in 8071 surgical patients using 796 clinical variables. The machine learning algorithms yielded an AUC of 0.70 and percentage misclassification of 35%.<sup>(19)</sup> When the CAPPs score was dichotomized to low- vs. high- or extreme-risk, the sensitivity and specificity were 89.4% and 49.5%, respectively. As such, patients who score between 0 and 4 are unlikely to develop poorly controlled pain after spine surgery. Notably, unlike in this study, most studies for validation of prediction models yield worse prognostic discrimination.<sup>20</sup> The comparable AUC and improved percentage misclassification of the eight-tier and three-tier CAPPs score further strengthens the independent validation.

By identifying patients at higher risk for developing poorly controlled pain (using the validated CAPPs score), a preoperative and perioperative pain management strategy based on a patient's risk factors can be designed to decrease the reliance on opioids and improve pain. For example, opioid deprescribing programs for patients taking daily opioid medications,<sup>21</sup> pre-habilitation programs for patients who score high on the PHQ-9 depression score,<sup>22</sup> and early involvement of acute pain services teams for those with non-modifiable risk factors (e.g., younger age, female sex, long segment, and fusion operations).<sup>13</sup> More invasive or time-consuming analgesic techniques such as intrathecal morphine or erector spinae plane block could be reserved for patients categorized in the extreme-risk group for developing poorly controlled pain. The CAPPs score could also be used to appropriately

select patients for clinical trials aimed at improving pain control and provide appropriate adjustment of confounders in observational studies.

The recent Enhanced Recovery After Surgery (ERAS<sup>®</sup>) Society consensus guideline on the perioperative management of lumbar spinal fusion identified the CAPPs score as a tool that can be used to improve preoperative patient education on expected pain control after surgery.<sup>23</sup> The incorporation of the CAPPs score in future ERAS programs for spinal surgery may lead to an individualized approach in the prevention and management of postoperative pain leading to improved patient recovery and reduction in overall surgical stress.<sup>23–25</sup> This paradigm shift in management could lead to improvements in postoperative pain intensity and patient satisfaction, and a reduction in postoperative opioid use and other pain-related complications such as the development of persistent postsurgical pain.

### Strengths and Limitations

The main strength of this study was the prospective nature in which baseline patient and prognostic variables were collected. This study attained the *a priori* determined number of subjects with the outcome of interest to adequately validate the CAPPs score. Postoperative pain assessments were performed by clinical care nurses without knowledge of the research objectives, and scoring was performed by independent research coordinators minimizing the risk for misclassification bias.

However, the findings from this present report should be interpreted in the context of the study design. The validation of the CAPPs score was performed on a temporally distinct population, but at the same institution where the prediction score was developed. This may limit the generalizability and overinflate the predictive performance. However, the predictive performance remained robust even with the differences in baseline patient characteristics and prognostic values (i.e., BMI, smoking history, baseline severe preoperative disability measured by ODI and NDI, principal pathology, baseline daily opioid medication use, and fusion surgery) compared to the development cohort. This provides confidence that the prediction score will perform well at different centers with different patient characteristics. Future studies should be conducted to improve the accuracy of the CAPPs score by exploring inclusion of additional preoperative and perioperative risk factors.

### Conclusions

Poor postoperative pain control is common following elective spine surgery. The CAPPs score, based on seven easily obtainable and reliable prognostic variables, has been validated on a prospective, independent sample. It can effectively predict the likelihood of experiencing poor postoperative pain control after elective spine surgery.

**Acknowledgment.** We thank Ish Bains MSc and Mehr Malhorta BHSc for their assistance in the ethics application and data collection for this study.

**Previous Presentations.** Parts of this work was presented at the Congress of Neurological Surgeons conference in Austin, Texas in 2021.

**Funding.** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Conflict of Interest.** The authors declare no competing interests.

**Statement of Authorship.** MMHY conceptualized the study, drafted the initial draft of the manuscript, performed data collection and data analyses, interpreted the study results, and critically reviewed the final draft of the manuscript.

JRC conceptualized the study, interpreted the study results, critically reviewed the final draft of the manuscript, and provided study supervision

JC performed data analysis, interpreted study results, and critically reviewed the final draft of the manuscript.

SC conceptualized the study, interpreted study results, critically reviewed the final draft of the manuscript, and provided study supervision.

## References

- Gerbershagen HJ, Aduckathil S, van Wijck AJ, Peelen LM, Kalkman CJ, Meissner W. Pain intensity on the first day after surgery: a prospective cohort study comparing 179 surgical procedures. *Anesthesiology*. 2013;118:934–44.
- Sommer M, De Rijke JM, Van Kleef M, et al. The prevalence of postoperative pain in a sample of 1490 surgical inpatients. *Eur J Anaesthesiol*. 2008;25:267–74.
- Grob D, Mannion AF. The patient's perspective on complications after spine surgery. *Eur Spine J*. 2009;18 Suppl, 3:380–5.
- Hah JM, Cramer E, Hilmoe H, et al. Factors associated with acute pain estimation, postoperative pain resolution, opioid cessation, and recovery: secondary analysis of a randomized clinical trial. *JAMA Netw Open*. 2019;2:e190168–e.
- Gerbershagen HJ, Pogatzki-Zahn E, Aduckathil S, et al. Procedure-specific risk factor analysis for the development of severe postoperative pain. *Anesthesiology*. 2014;120:1237–45.
- Yang MMH, Hartley RL, Leung AA, et al. Preoperative predictors of poor acute postoperative pain control: a systematic review and meta-analysis. *BMJ Open*. 2019;9:e025091.
- Schnabel A, Yahiaoui-Doktor M, Meissner W, Zahn PK, Pogatzki-Zahn EM. Predicting poor postoperative acute pain outcome in adults: an international, multicentre database analysis of risk factors in 50,005 patients. *PAIN Rep*. 2020;5:e831.
- van Boekel RLM, Bronkhorst EM, Vloet L, Steegers MAM, Vissers KCP. Identification of preoperative predictors for acute postsurgical pain and for pain at three months after surgery: a prospective observational study. *Sci Rep*. 2021;11:16459.
- Yang MMH, Riva-Cambrin J. Prediction tools for postoperative pain. *PAIN Rep*. 2021;6:e875.
- Harris EP, MacDonald DB, Boland L, Boet S, Lalu MM, McIsaac DI. Personalized perioperative medicine: a scoping review of personalized assessment and communication of risk before surgery. *Can J Anaesth*. 2019;66:1026–37.
- Yang MMH, Riva-Cambrin J, Cunningham J, et al. Development and validation of a clinical prediction score for poor postoperative pain control following elective spine surgery: Presented at the 2020 AANS/CNS Joint Section on Disorders of the Spine and Peripheral Nerves. *J Neurosurg Spine*. 2021;34:3–12.
- Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multi-variable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162:W1–73.
- Chou R, Gordon DB, de Leon-Casasola OA, et al. Management of Postoperative Pain: a clinical practice guideline from the American Pain Society, the American Society of Regional Anesthesia and Pain Medicine, and the American Society of Anesthesiologists' committee on regional anesthesia, executive committee, and administrative council. *J Pain*. 2016;17:131–57.
- Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med*. 2016;35:214–26.
- Grant SW, Collins GS, Nashef SA. Statistical Primer: developing and validating a risk prediction model. *Eur J Cardiothorac Surg*. 2018;54:203–8.
- Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–31.
- Shaikh S. Measures derived from a 2 x 2 table for an accuracy of a diagnostic test. *J Biomet Biostat*. 2011;2:128.
- Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009;338:b605–b605.
- Tighe PJ, Harle CA, Hurley RW, Aytug H, Boezaart AP, Fillingim RB. Teaching a machine to feel postoperative pain: combining high-dimensional clinical data with machine learning algorithms to forecast acute postoperative pain. *Pain Med*. 2015;16:1386–401.
- Siontis GC, Zoulaki I, Castaldi PJ, Ioannidis JP. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. *J Clin Epidemiol*. 2015;68:25–34.
- Hassamal S, Haglund M, Wittnebel K, Danovitch I. A preoperative interdisciplinary biopsychosocial opioid reduction program in patients on chronic opioid analgesia prior to spine surgery: a preliminary report and case series. *Scand J Pain*. 2016;13:27–31.
- Abid Azam M, Weinrib AZ, Montbriand J, et al. Acceptance and Commitment Therapy to manage pain and opioid use after major surgery: preliminary outcomes from the Toronto General Hospital Transitional Pain Service. *Can J Pain*. 2017;1:37–49.
- Debono B, Wainwright TW, Wang MY, et al. Consensus statement for perioperative care in lumbar spinal fusion: Enhanced Recovery After Surgery (ERAS®) Society recommendations. *Spine J*. 2021;21:729–52.
- Ghogawala Z. Editorial. Postoperative pain control after spinal surgery: an opportunity. *J Neurosurg Spine*. 2020, 1–2.
- Editorial Berrington NR. The opioid crisis: an opportunity to alter morbidity through the implementation of enhanced recovery after surgery protocols during spinal surgery? *Neurosurg Focus*. 2019;46:E13.