**ROBERT J. NEE,** PT, PhD[1,2] • **BILL VICENZINO,** PT, PhD[1] • **GWENDOLEN A. JULL,** PT, PhD[1]
**JOSHUA A. CLELAND,** PT, PhD[3] • **MICHEL W. COPPIETERS,** PT, PhD[1]

# Baseline Characteristics of Patients With Nerve-Related Neck and Arm Pain Predict the Likely Response to Neural Tissue Management

**STUDY DESIGN:** Planned secondary analysis of a randomized controlled trial comparing neural tissue management (NTM) to advice to remain active.

**OBJECTIVE:** To develop a model that predicts the likelihood of patient-reported improvement following NTM.

**BACKGROUND:** Matching patients to an intervention they are likely to benefit from potentially improves outcomes. However, baseline characteristics that predict patients' responses to NTM are unknown.

**METHODS:** Data came from 60 consecutive adults who had nontraumatic, nerve-related neck and unilateral arm pain for at least 4 weeks. Participants were assigned to a group that received NTM (n = 40), which involved brief education, manual therapy, and nerve gliding exercises for 4 treatments over 2 weeks, or to a group that was given advice to remain active (n = 20), which involved instruction to continue their usual activities. The participants' global rating of change at a 3- to 4-week follow-up defined improvement. Penalized regression of NTM data identified the best prediction model. A medical nomogram was created for prediction model scoring. Post hoc analysis determined whether the model predicted a specific response to NTM.

**RESULTS:** Absence of neuropathic pain qualities, older age, and smaller deficits in median nerve neurodynamic test range of motion predicted improvement. Prediction model cutoffs increased the likelihood of improvement from 53% to 90% (95% confidence interval: 56%, 98%) or decreased the likelihood of improvement to 9% (95% confidence interval: 1%, 42%). The model did not predict the outcomes of the advice to remain active group.

**CONCLUSION:** Baseline characteristics of patients with nerve-related neck and arm pain predicted the likelihood of improvement with NTM. Model performance needs to be validated in a new sample using different comparison interventions and longer follow-up. Australian New Zealand Clinical Trials Registry (ACTRN 12610000446066).

**LEVEL OF EVIDENCE:** Prognosis, level 2b–.
*J Orthop Sports Phys Ther 2013;43(6):379-391. Epub 30 April 2013. doi:10.2519/jospt.2013.4490*

**KEY WORDS:** *cervical radicular pain, clinical prediction rule, medical nomogram, neurodynamic treatment, penalized regression*

Nearly 35% of patients who seek physical therapy for spinal disorders report neck pain.[17,36] Neck and arm pain is more common than neck pain alone and is associated with higher levels of self-reported disability.[15] Sensitized neural tissues provide one potential mechanism for neck and arm pain to coexist.[6]

Researchers have been encouraged to identify characteristics that predict patients' responses to different physical therapy interventions.[22,57] One physical therapy intervention advocated for patients with nerve-related neck and arm pain is neural tissue management (NTM), also referred to as neurodynamic treatment.[7,8,20] NTM uses specific combinations of active or passive movements that aim to reduce nerve mechanosensitivity and restore symptom-free function.[14] We have shown that NTM provides clinically important benefits beyond advice to remain active (ARA) in the short term for patients with nerve-related neck and arm pain.[51] However, characteristics that predict patients' responses to NTM are unknown. One research-based strategy to identify these characteristics is to develop a prediction model.[44]

Study design, statistical methods for creating the model, and the format of model presentation are important to consider during the initial development of a prediction model.[59] The prediction model's purpose dictates the best study design for its development.[4] A treatment-related pre-

[1]Division of Physiotherapy and National Health and Medical Research Council Centre of Clinical Research Excellence in Spinal Pain, Injury and Health, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia. [2]Department of Physical Therapy, Arizona School of Health Sciences, A.T. Still University, Mesa, AZ. [3]Department of Physical Therapy, College of Graduate and Professional Studies, Franklin Pierce University, Concord, NH. Robert J. Nee was funded by an Endeavour International Postgraduate Research Scholarship from the Australian Government and a Research Scholarship from The University of Queensland. The study was approved by the Medical Research Ethics Committee, The University of Queensland, Brisbane, Australia. The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Dr Robert J. Nee, Department of Physical Therapy, Arizona School of Health Sciences, A.T. Still University, 5850 East Still Circle, Mesa, AZ 85206. E-mail: robertjnee@gmail.com 

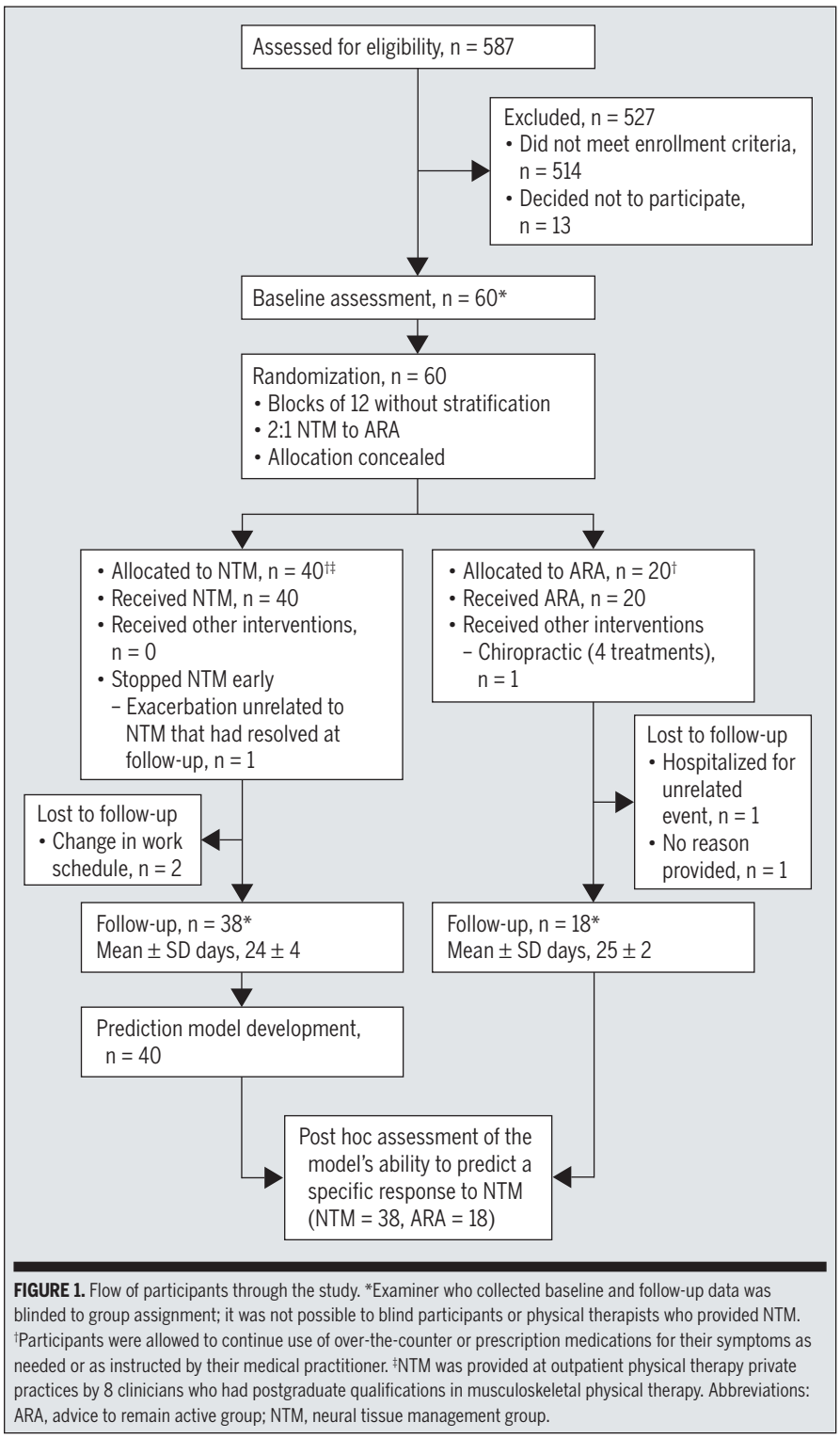



**FIGURE 1.** Flow of participants through the study. *Examiner who collected baseline and follow-up data was blinded to group assignment; it was not possible to blind participants or physical therapists who provided NTM. †Participants were allowed to continue use of over-the-counter or prescription medications for their symptoms as needed or as instructed by their medical practitioner. ‡NTM was provided at outpatient physical therapy private practices by 8 clinicians who had postgraduate qualifications in musculoskeletal physical therapy. Abbreviations: ARA, advice to remain active group; NTM, neural tissue management group.

diction model should be developed from a randomized controlled trial, because it is important to show that the model identifies patients who respond specifically to the treatment, rather than patients who simply have a favorable natural history.[4,26,41,53]

Statistical methods for creating the model should maximize the chance that the model will maintain its predictive performance when applied to future patients.[28,59] Consideration of all potentially important predictors during initial model development frequently requires analyzing a large number of clinical variables.[55] Stepwise logistic regression requires at least 5 patients in the limiting sample size for each variable analyzed as a potential predictor for the model.[28,64] The limiting sample size is the number of patients who experience the least frequent outcome after the prediction model treatment (eg, "improved" or "not improved").[28] When analyzing a large number of clinical variables during initial model development, it is usually not feasible to recruit enough patients to obtain the required limiting sample size. Analyzing too many variables for the limiting sample size increases the risk that the derived model will not perform well when applied to future patients.[13,28] Univariate screening of all clinical variables to decrease the number entered into stepwise logistic regression may not reduce this risk, because univariate screening may eliminate variables that have predictive value when included in a stepwise logistic regression.[28,60] Therefore, when developing a prediction model from a large number of clinical variables relative to the limiting sample size, penalized regression may be a better option, because it may be more likely than univariate screening and stepwise logistic regression to produce a model that maintains its predictive performance in future patients.[28,45,61]

The format of model presentation also affects predictive performance in future patients. Physical therapy treatment–related prediction models usually dichotomize continuous predictor variables according to "optimal" cutoffs derived from receiver-operating-characteristic (ROC) curves (eg, Flynn et al,[21] Puentedura et al[54]). The common argument for dichotomizing continuous variables is to make it easier for clinicians to understand and use the final model.[46] However, dichotomizing continuous variables according to optimal cutoffs from ROC

curves or other data-dependent techniques can be problematic.[40] Although the cutoffs may be optimal for the sample used during initial model development, they may not be optimal for the larger population of patients on whom the model will be applied.[2,56] Furthermore, similar to univariate screening, information lost by dichotomizing continuous variables prior to stepwise logistic regression can change which variables are selected for the final model and reduce predictive performance in future patients.[56,58] A better approach may be to keep continuous variables in the model and to generate a scoring system by which the total score would provide an estimate of the patient's chance for improvement with the prediction model treatment.[56] Cutoffs for making clinical decisions are applied to the prediction model's total score.[56] Graphical presentation of the model in a medical nomogram[24,38] allows for easy scoring of continuous variables and facilitates maintenance of the model's performance in future patients, because it preserves all predictive information from continuous variables in the model.[2,24,38,40,56,58]

The aim of this study was to develop a model that predicts whether patients with nerve-related neck and arm pain are likely to improve after NTM. The chosen study methods were those that permitted an assessment of the model's ability to predict a specific response to NTM and increased the chance that the model's predictive performance will be maintained when applied to future patients.

## METHODS

### Design Overview

This study was a planned secondary analysis of a randomized controlled trial that compared NTM to ARA (**FIGURE 1**).[52] Although the methods and results of the trial have been reported in detail elsewhere,[51,52] information relevant to the development and interpretation of the prediction model is presented here. Participant-reported improvement at a 3- to 4-week follow-up was the primary outcome of the trial and was the outcome used to develop the prediction model. The model was created by analyzing data from participants in the NTM group. Consecutive participants who met all enrollment criteria were randomly assigned to NTM or ARA in a 2:1 ratio to increase the amount of NTM data available for prediction model development. Of the 40 participants assigned to NTM, 21 (53%) were classified as improved at follow-up.[51] The participants who were assigned to the group that received ARA (instruction to continue their usual activities) (n = 20) provided data on the natural history of nerve-related neck and arm pain needed to assess whether the model predicted a specific response to NTM. To improve the chance that the derived model will maintain its predictive performance when applied to future patients, we (*a*) identified the best examination items for the model by analyzing NTM data with penalized regression, using the least absolute shrinkage and selection operator (LASSO)[61]; and (*b*) created a medical nomogram for prediction model scoring.[24,38] The model's ability to predict a specific response to NTM was assessed by a post hoc analysis of data from participants in both the NTM and ARA groups.[42] Ethical approval was obtained from the Medical Research Ethics Committee of the University of Queensland, and all participants provided informed consent prior to data collection. The trial was registered with the Australian New Zealand Clinical Trials Registry (ACTRN 12610000446066).

**TABLE 1** Baseline Characteristics of Participants

| Characteristic | All (n = 60) | NTM (n = 40) | ARA (n = 20) |
|---|---|---|---|
| Gender (female), n (%) | 38 (63) | 26 (65) | 12 (60) |
| Age, y* | 47 ± 9 | 47 ± 8 | 48 ± 9 |
| Body mass index, kg/m$^2$* | 26.7 ± 4.4 | 27.3 ± 4.7 | 25.7 ± 3.7 |
| Duration of symptoms, wk† | 26 (12-77) | 32 (15-104) | 18 (8-39) |
| Symptoms distal to elbow, n (%) | 46 (77) | 29 (73) | 17 (85) |
| Arm symptoms worst, n (%) | 20 (33) | 11 (28) | 9 (45) |
| Reported numbness or tingling, n (%) | 32 (53) | 20 (50) | 12 (60) |
| Using medication for symptoms, n (%) | 27 (45) | 23 (58) | 4 (20) |
| Neck pain previous 24 h*‡§ | 4.2 ± 2.0 | 4.3 ± 1.7 | 4.1 ± 2.4 |
| Arm pain previous 24 h*‡§ | 4.0 ± 1.6 | 4.0 ± 1.6 | 4.1 ± 1.6 |
| Neck Disability Index (0-50)*§‖ | 12.5 ± 4.4 | 12.7 ± 4.2 | 12.1 ± 4.7 |

*Abbreviations: ARA, advice to remain active group; NTM, neural tissue management group.*
**Values are mean ± SD.*
*†Values are median (interquartile range).*
*‡Scored on a numeric pain rating scale (0-10), with 0 as no pain and 10 the worst pain imaginable.*
*§Outcome measure for validating the global rating of change criterion for participant-reported improvement.*
*‖Higher scores represent higher levels of self-reported disability.*

### Participants

Sixty participants recruited from the general community qualified for the trial and provided data for developing the prediction model and assessing its ability to predict a specific response to NTM (**TABLE 1**, **FIGURE 1**).[51] To enter the trial, participants had to be 18 to 60 years of age and to have neck and unilateral arm pain that spread below the deltoid tuberosity but was not related to trauma. Symptoms had to have been present for at least 4 weeks and preceded by a pain-free period of 4 weeks or longer.[16] Participants rated the average intensity of their neck and arm pain during the previous week on separate 11-point numeric pain rating scales.[35] The mean of these pain ratings had to be

**TABLE 2** Examination Items Measured at Baseline for Potential Inclusion in the Prediction Model

| Patient Interview and Self-Report Items | Abbreviation | Type | CATREG Scaling Level |
|---|---|---|---|
| 1. Age | AGE | Continuous | Numeric |
| 2. Female | FEMALE | Dichotomous | Ordinal |
| 3. Symptoms distal to elbow | SXDISTELB | Dichotomous | Ordinal |
| 4. "Most bothersome" symptoms located in the arm | ARMWORST | Dichotomous | Ordinal |
| 5. Numeric pain rating for neck pain during previous 24 h | NECKPAIN | Continuous | Numeric |
| 6. Numeric pain rating for arm pain during previous 24 h | ARMPAIN | Continuous | Numeric |
| 7. Participant reports numbness or tingling | NUMTIN | Dichotomous | Ordinal |
| 8. Turning toward symptomatic arm aggravates symptoms | AGIPTURN | Dichotomous | Ordinal |
| 9. Turning away from symptomatic arm aggravates symptoms | AGCONTURN | Dichotomous | Ordinal |
| 10. Looking up aggravates symptoms | AGLKUP | Dichotomous | Ordinal |
| 11. Looking down aggravates symptoms | AGLKDOWN | Dichotomous | Ordinal |
| 12. Driving aggravates symptoms | AGDRIVE | Dichotomous | Ordinal |
| 13. Reading aggravates symptoms | AGREAD | Dichotomous | Ordinal |
| 14. Computer use aggravates symptoms | AGCOMPUTE | Dichotomous | Ordinal |
| 15. Arm use below head level (not computer) aggravates symptoms | AGARMLOW | Dichotomous | Ordinal |
| 16. Arm use above head level aggravates symptoms | AGARMHI | Dichotomous | Ordinal |
| 17. Sleeping aggravates symptoms | AGSLEEP | Dichotomous | Ordinal |
| 18. Duration of current episode of symptoms (less than 12 wk, 12-26 wk, greater than 26 wk) | DURATION | Categorical | Ordinal |
| 19. Previous episodes of similar neck and arm symptoms* | PREVSIM | Dichotomous | Ordinal |
| 20. Neck Disability Index score (0-50 points) | NDI | Continuous | Numeric |
| 21. Self-report Leeds Assessment of Neuropathic Symptoms and Signs score of 12 or greater | S-LANSS | Dichotomous | Ordinal |
| 22. Tampa Scale of Kinesiophobia score (17-68 points) | TSK | Continuous | Numeric |

*Table continues on page 383.*

at least 3/10 to be included in the trial.

Participants' symptoms had to be reproduced by a mechanical provocation test for the cervical nerve roots and median nerve, known as the median nerve upper-limb neurodynamic test (ULNT-$1_{MEDIAN}$).[7,19] Importantly, contralateral neck sidebending or releasing wrist extension at the end of ULNT$1_{MEDIAN}$ had to change these symptoms.[7,19] This ULNT-$1_{MEDIAN}$ response suggested that participants' symptoms were at least partly related to increased nerve mechanosensitivity in the neck and upper limb.[7,25,50] Participants with 2 or more abnormal neurological signs (decreased strength, reflex, or sensation) at the same nerve root level (C5-T1) were excluded.[65] It has been suggested that these 2 enrollment criteria would select participants who would be considered appropriate candidates for NTM.[7,20,25]

Participants were also excluded if they had bilateral arm symptoms, symptoms or signs suggestive of cervical myelopathy, physical therapy for neck and arm pain within the previous 6 weeks, previous neck or upper-limb surgery, or medical red flags[9] that suggested serious pathology. Self-report items required that the participants read and speak English proficiently.

## Neural Tissue Management

The prediction model treatment was administered for 4 treatments over 2 weeks. It included a brief educational component that aimed to reduce any apprehension that participants might have had about NTM, manual therapy applied to the cervical spine (contralateral cervical lateral glide[20]) and shoulder girdle (shoulder girdle mobilization with active craniocervical flexion to elongate the posterior cervical spine[20,52]), and a home program of nerve gliding exercises for the cervical nerve roots and median nerve that involved different combinations of neck, elbow, and wrist movements in varying angles of shoulder abduction. The manual therapy techniques and nerve gliding exercises were prescribed so as not to provoke the participants' symptoms and have been recommended for reducing mechanosensitivity in the cervical nerve roots and median nerve.[7,14,20] The NTM protocol has been described in detail elsewhere.[52] No participants had to stop NTM early because of an adverse response to treatment.[51] Two participants assigned to NTM who did not complete follow-up were classified as not improved for the purpose of prediction model development (**FIGURE 1**).

## Outcome Predicted by the Model

The outcome predicted by the model was participant-reported improvement on a 15-point global rating of change (GROC) scale at a 3- to 4-week follow-up. The

**TABLE 2** Examination Items Measured at Baseline for Potential Inclusion in the Prediction Model (continued)

| Physical Examination Items | Abbreviation | Type | CATREG Scaling Level |
|---|---|---|---|
| 23. Active glenohumeral abduction ROM with wrist extended† | WEGHA | Continuous | Numeric |
| 24. Painful deficit in HBB (symptomatic arm relative to asymptomatic arm) | HBBPDEF | Dichotomous | Ordinal |
| 25. Active neck flexion ROM | FLEX | Continuous | Numeric |
| 26. Active neck extension ROM | EXTEND | Continuous | Numeric |
| 27. Active neck rotation ROM toward symptomatic arm | IPSIROTN | Continuous | Numeric |
| 28. Active neck rotation ROM away from symptomatic arm | CONROTN | Continuous | Numeric |
| 29. Active neck lateral flexion ROM toward symptomatic arm | IPSILF | Continuous | Numeric |
| 30. Active neck lateral flexion ROM away from symptomatic arm | CONLF | Continuous | Numeric |
| 31. Active neck movements provoke or increase arm symptoms | ACTNECK | Dichotomous | Ordinal |
| 32. Spurling test toward side of symptomatic arm provokes symptoms | SPURLING | Dichotomous | Ordinal |
| 33. Valsalva maneuver provokes symptoms | VALSALVA | Dichotomous | Ordinal |
| 34. Supine neck distraction test alleviates symptoms | SUPDIST | Dichotomous | Ordinal |
| 35. Deficit in elbow extension ROM during $ULNT1_{MEDIAN}$ (symptomatic arm relative to asymptomatic arm) | ULNTDEF | Continuous | Numeric |
| 36. Positive radial nerve neurodynamic test on symptomatic arm† | ULNTRAD | Dichotomous | Ordinal |
| 37. Positive ulnar nerve neurodynamic test on symptomatic arm† | ULNTULN | Dichotomous | Ordinal |
| 38. Number of sites sensitive to nerve palpation (brachial plexus, median nerve upper arm, median nerve cubital fossa, radial nerve upper arm, ulnar nerve cubital tunnel) | NRVPALP | Continuous | Numeric |
| 39. CCLG with forearm on abdomen provokes symptoms* | CCLG | Dichotomous | Ordinal |
| 40. CCLG with arm in 45° of abduction with elbow extended provokes symptoms | CCLGAB | Dichotomous | Ordinal |
| 41. Spring testing of first rib on side of symptomatic arm provokes arm symptoms | RIB | Dichotomous | Ordinal |
| 42. Central posterior/anterior pressures C4 to T3 provoke arm symptoms | CENPA | Dichotomous | Ordinal |
| 43. Unilateral posterior/anterior pressures C4 to T3 on side of symptomatic arm provoke arm symptoms | UNIPA | Dichotomous | Ordinal |

*Abbreviations: CATREG, categorical regression; CCLG, contralateral cervical lateral glide; HBB, hand behind back; ROM, range of motion; $ULNT1_{MEDIAN}$, median nerve upper-limb neurodynamic test.*
**Eliminated from CATREG least absolute shrinkage and selection operator analysis because of a zero cell in the 2-way table with improvement after neural tissue management.*
*†Eliminated from CATREG least absolute shrinkage and selection operator analysis to reduce the total number of examination items to 38. We felt that "deficit in elbow extension ROM during $ULNT1_{MEDIAN}$" (ULNTDEF) and "number of sites sensitive to nerve palpation" (NRVPALP) were more important for assessing the construct of nerve mechanosensitivity.*

scale spans from –7 ("a very great deal worse") to 0 ("no change") to +7 ("a very great deal better").[34] A GROC of +4 or greater (at least "moderately better") classified a participant as improved. Due to concerns that participants might not remember their baseline status accurately when scoring their GROC at follow-up,[37] neck pain, arm pain, and participant-reported function were assessed at baseline and follow-up, to validate that, regardless of group assignment, participants who were improved differed from those who were not improved. Neck pain and arm pain outcomes were the mean of the numeric pain ratings for the participants' current, highest, and lowest levels of pain during the previous 24 hours.[11] Participant-reported function was assessed with the Neck Disability Index[62] and the Patient-Specific Functional Scale (PSFS).[66]

## Examination Items for the Prediction Model

Examination items for prediction model development were assessed on each participant at baseline by the examiner who collected outcome data (**TABLE 2**). The rationale for item selection and available information on reliability have been reported previously.[52]

## Statistical Analysis

Statistical analysis involved (*a*) validation of the GROC criterion for improvement, (*b*) prediction model development, and (*c*) a post hoc assessment of the model's ability to predict a specific response to NTM.[42,52] The alpha level was .05 unless stated otherwise.

**Validation of the GROC Criterion for Improvement** Separate analyses of covariance determined whether follow-up scores for neck pain, arm pain, and the Neck Disability Index were significantly different between participants who were improved and not improved, regardless of group assignment. Follow-up scores in each analysis of covariance were adjusted by using the baseline score as the covariate.[63] Because PSFS activities were different for

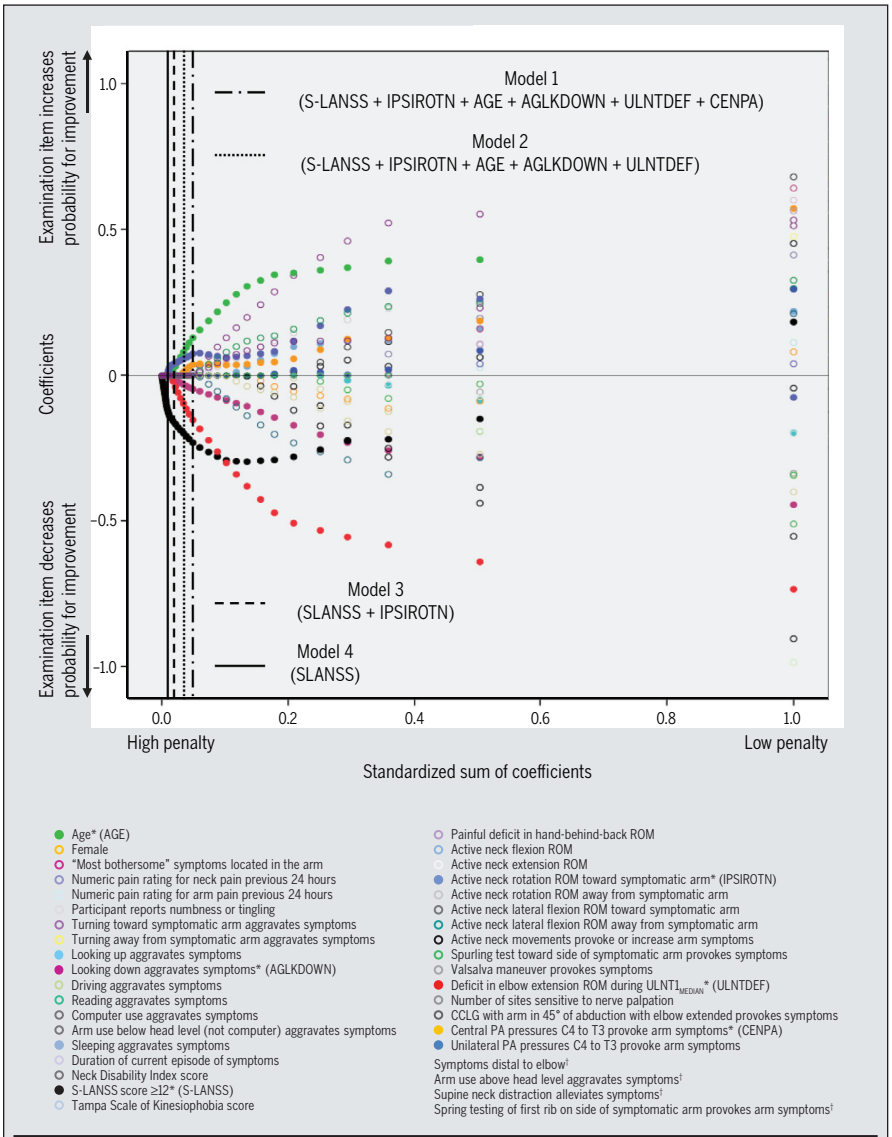



**FIGURE 2.** Regression coefficient paths from the CATREG LASSO analysis. Each vertically aligned set of symbols represents a potential model solution. Regression coefficients for less important examination items shrink to zero (middle horizontal line) more quickly as the LASSO penalty increases from right to left. When regression coefficients for 1 or more examination items shrink to zero, the subsequent model solution contains the remaining items with nonzero coefficients. The vertical lines mark those potential model solutions with 6 or fewer examination items. *Included in 1 or more of the potential model solutions with 6 or fewer examination items. †Removed from the figure because examination item was not part of any potential model solution. Abbreviations: CATREG, categorical regression; CCLG, contralateral cervical lateral glide; LASSO, least absolute shrinkage and selection operator; PA, posterior/anterior; ROM, range of motion; S-LANSS, self-report Leeds Assessment of Neuropathic Symptoms and Signs; $ULNT1_{MEDIAN}$, median nerve upper-limb neurodynamic test.

each participant, PSFS change scores were analyzed with an unpaired *t* test.

**Prediction Model Development** Data from participants who were randomized to the NTM group were analyzed for prediction model development. Dichotomous examination items with zero cells in their 2-way table with improvement after NTM were removed from the analysis because a zero cell prevents accurate calculation of that item's regression coefficient.[32] All remaining items were entered into a LASSO analysis using the categorical regression (CATREG) option in SPSS Statistics Version 19 (IBM Corporation, Armonk, NY). The LASSO applies a statistical penalty that shrinks the size of each examination item's regression coefficient, and items with coefficients that shrink to zero are not selected for the final model.[61] Given the large number of examination items that were analyzed as potential predictors, penalized regression with the LASSO was necessary, because obtaining the sample required for stepwise logistic regression was not feasible.[28,45,52,61,64] CATREG requires researchers to designate a scaling level for each examination item. When nonlinear relationships with the outcome are of interest, the scaling level should be less restrictive than the measurement level of the examination item (eg, ordinal scaling level for a continuous item).[29] Because we were not interested in nonlinear relationships, categorical and dichotomous items were scaled at the ordinal level and continuous items at the numeric level (**TABLE 2**).

The CATREG LASSO algorithm produces a series of potential model solutions based on minimizing the expected prediction error, which is the projected amount of error in the model's estimate of a patient's chance for improvement after NTM when the model is applied to a new sample of patients.[59] The expected prediction error for each potential model solution was estimated with the "0.632 bootstrap" ("leave-one-out bootstrap") from 200 bootstrap samples.[18]

LASSO regression coefficients are usually biased toward zero.[31] One way to correct this bias is to recalculate the coefficients for each potential model solution without the LASSO penalty.[31] Therefore, potential model solutions with 6 or fewer examination items were analyzed separately in CATREG without the LASSO penalty to select the final model.[29,30] This choice was based on previously proposed treatment-related prediction models that contained as many as 6 items.[10] If the unpenalized CATREG coefficient for 1 or more examination items within a model was not statistically significant ($P>.05$), a revised version of the model

without those items was analyzed.[32,59] The unpenalized CATREG estimate of each model's expected prediction error was based on the 0.632 bootstrap from 200 bootstrap samples. The goal was to find the model with the lowest expected prediction error and the smallest number of examination items. Performance of the final model was summarized by Nagelkerke $R^2$,[49] the area under the ROC curve,[59] and the Hosmer-Lemeshow goodness-of-fit test.[32]

A medical nomogram for the final model, based on unpenalized logistic regression coefficients, was created in R with the regression modeling strategies package Version 3.3-1.[27] Separate ROC curves determined the best prediction model cutoffs for identifying participants who were and were not likely to improve after NTM.[48] Sensitivity, specificity, and likelihood ratios with 95% confidence intervals (CIs) quantified the clinical value of each cutoff.[33]

**Assessment of the Model's Ability to Predict a Specific Response to NTM** A post hoc analysis proposed by Kent et al[42] assessed whether the model's cutoff for improvement predicted a specific response to NTM. The differences in the frequency of participant-reported improvement between the NTM and ARA groups were evaluated in participants who met the prediction model cutoff for improvement and in those who did not. To evaluate the hypothesis that the model predicted a specific response to NTM, the difference in the frequency of participant-reported improvement between the NTM and ARA groups should be larger among participants who met the prediction model cutoff for improvement.[26,42] The post hoc analysis only included participants who completed follow-up (**FIGURE 1**).

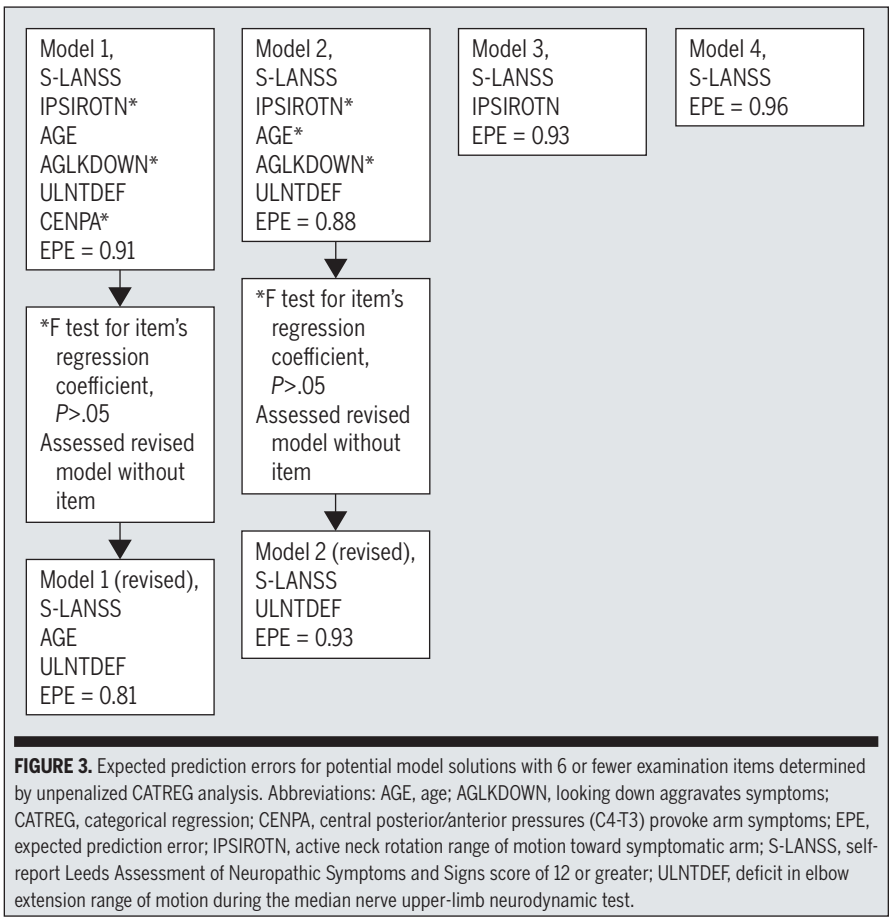



**FIGURE 3.** Expected prediction errors for potential model solutions with 6 or fewer examination items determined by unpenalized CATREG analysis. Abbreviations: AGE, age; AGLKDOWN, looking down aggravates symptoms; CATREG, categorical regression; CENPA, central posterior/anterior pressures (C4-T3) provoke arm symptoms; EPE, expected prediction error; IPSIROTN, active neck rotation range of motion toward symptomatic arm; S-LANSS, self-report Leeds Assessment of Neuropathic Symptoms and Signs score of 12 or greater; ULNTDEF, deficit in elbow extension range of motion during the median nerve upper-limb neurodynamic test.

# RESULTS

## Validation of the GROC Criterion for Improvement

Results supported the validity of the GROC criterion for participant-reported improvement. Mean differences (95% CI) showed that, regardless of group assignment, participants who improved (GROC of +4 or greater) had lower follow-up scores for neck pain (difference, −2.4; 95% CI: −1.5, −3.2; $P<.001$), arm pain (difference, −2.5; 95% CI: −1.7, −3.2; $P<.001$), and the Neck Disability Index (difference, −5.7; 95% CI: −3.6, −7.8; $P<.001$), and higher PSFS change scores (difference, 2.7; 95% CI: 1.8, 3.6; $P<.001$).

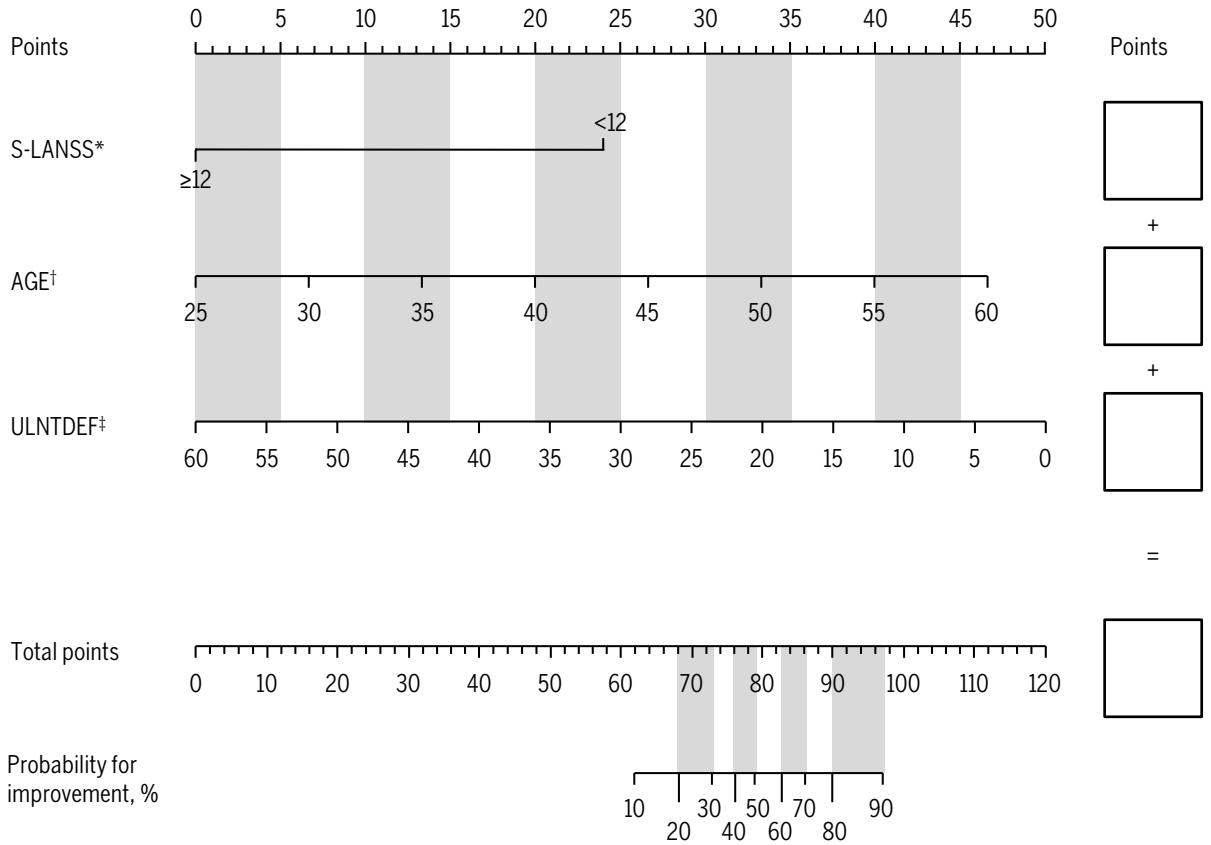
## Prediction Model

In CATREG, the number of valid cases must be greater than the number of potential predictors plus 1. Because 40 participants were assigned to NTM, only 38 of the 43 examination items could be entered into the CATREG LASSO analysis. Two items were eliminated because of zero cells in their 2-way tables with improvement after NTM (**TABLE 2**). Because several examination items assessed the construct of nerve mechanosensitivity, 3 items that we considered less important for assessing this construct were eliminated to reach the limit of 38 (**TABLE 2**).

The CATREG LASSO analysis revealed 4 potential model solutions with 6 or fewer examination items (**FIGURE 2**). Unpenalized CATREG analysis of these potential model solutions determined that a model containing "self-report Leeds Assessment of Neuropathic Symptoms and Signs score of 12 or greater" (S-LANSS), "age," and "deficit in elbow extension range of motion during $ULNT1_{MEDIAN}$" had the lowest expected prediction error (**FIGURE 3**). S-LANSS less than 12 (absence of neuropathic pain qualities), older age, and smaller deficit in elbow extension range of motion during $ULNT1_{MEDIAN}$ predicted improvement after NTM. The model explained 46% of

Medical nomogram for predicting
improvement after neural tissue management

SCORING: Locate the patient's score for each examination item on the corresponding horizontal scale. Draw a straight line up to the "Points" scale to assign a certain number of points to the patient's score for each item (shaded columns provide orientation to 5-point increments on the "Points" scale). Sum the points for all items and locate this value on the "Total Points" scale. Draw a straight line down to estimate the patient's "Probability for Improvement" after neural tissue management (NTM) (shaded columns provide orientation to increments of 10% on the "Probability for Improvement" scale).




*S-LANSS, total score on the Self-report Leeds Assessment of Neuropathic Symptoms and Signs scale. Patient reminded to complete the questionnaire for the worst area of symptoms.

†AGE, patient's report of age.

‡ULNTDEF, deficit in elbow extension range of motion during the median nerve neurodynamic test ($ULNT1_{MEDIAN}$). The symptomatic limb is tested first. After stabilizing the shoulder girdle to prevent elevation during testing, each movement (shoulder abduction, wrist extension, forearm supination, and shoulder external rotation) is performed to a point just before the provocation of any symptoms (or to end range if no symptoms are provoked). The elbow is then extended to the onset of symptoms and the angle is measured with a standard goniometer (180° = full extension). The process is repeated on the asymptomatic limb, but shoulder, wrist, and forearm angles are matched to the angles reached on the symptomatic limb. Elbow extension on the asymptomatic limb is performed to the onset of a sensory response. ULNTDEF = (elbow extension angle asymptomatic limb) – (elbow extension angle symptomatic limb).

**FIGURE 4.** Medical nomogram for the prediction model.

**TABLE 3** Accuracy Statistics for Prediction Model Cutoffs to Identify Participants Who Were Likely and Not Likely to Improve After Neural Tissue Management

| Likely to Improve | | Not Likely to Improve | |
|---|---|---|---|
| Prediction model score ⩾89, n | | Prediction model score ⩽71, n | |
| Improved | 9 | Improved | 1 |
| Not improved | 1 | Not improved | 10 |
| Prediction model score <89, n | | Prediction model score >71, n | |
| Improved | 12 | Improved | 20 |
| Not improved | 18 | Not improved | 9 |
| Sensitivity* | 0.43 (0.25, 0.64) | Sensitivity* | 0.05 (0.01, 0.23) |
| Specificity* | 0.95 (0.75, 0.99) | Specificity* | 0.47 (0.27, 0.68) |
| Likelihood ratio for improvement* | 8.14 (1.14, 58.42) | Likelihood ratio for improvement* | 0.09 (0.01, 0.64) |

**Values in parentheses are 95% confidence intervals.*

**TABLE 4** Post Hoc Assessment of the Model's Ability to Predict a Specific Response to Neural Tissue Management: Complete Case Analysis

| | Prediction Model Score ⩾89 | Prediction Model Score <89 |
|---|---|---|
| NTM, n | | |
| Improved | 9 | 12 |
| Not improved | 1 | 16 |
| ARA, n | | |
| Improved | 0 | 1 |
| Not improved | 5 | 12 |
| Frequency of improvement, NTM | 0.90 | 0.43 |
| Frequency of improvement, ARA | 0.00 | 0.08 |
| Difference in frequency of improvement favoring NTM* | 0.90 (0.60, 1.00) | 0.35 (0.12, 0.59) |

*Abbreviations: ARA, advice to remain active group; NTM, neural tissue management group.*
**Values in parentheses are 95% confidence intervals.*

the variance in outcome (Nagelkerke $R^2$ = 0.46) and discriminated participants who improved from those who did not (area under the ROC curve, 0.85; 95% CI: 0.72, 0.98). The Hosmer-Lemeshow goodness-of-fit test was removed from the analysis of model performance, because the NTM group had only 40 participants. This small sample meant that the Hosmer-Lemeshow goodness-of-fit test did not have enough statistical power to determine whether there was a meaningful difference between observed outcomes and outcomes predicted by the model.[32]

The medical nomogram for the final model is shown in **FIGURE 4**. Prediction model scores of 89 or more points (120 maximum) identified participants who were likely to improve after NTM (**TABLE 3**). Scores of 71 or less identified participants who were not likely to improve (**TABLE 3**). Given a pretest probability for improvement of 53%,[51] the posttest probability increased to 90% (95% CI: 56%, 98%) for participants with scores of 89 or greater and decreased to 9% (95% CI: 1%, 42%) for participants with scores of 71 or less.

### The Model's Ability to Predict a Specific Response to NTM

The post hoc analysis supported the hypothesis that the model's cutoff of 89 or more points predicted a specific response to NTM. Among participants with prediction model scores of 89 or more points, the difference in the frequency of participant-reported improvement between the NTM and ARA groups was 90% (95% CI: 60%, 100%), whereas the difference in the frequency of improvement between the NTM and ARA groups for participants with prediction model scores of less than 89 was only 35% (95% CI: 12%, 59%) (**TABLE 4**).

## DISCUSSION

Patients who met our definition of nerve-related neck and arm pain had a 53% chance of improving with NTM at a 3- to 4-week follow-up.[51] Because this is equivalent to flipping a coin, a tool that could help clinicians identify those patients who may or may not be good candidates for NTM would be useful. A model that combines neuropathic pain qualities (S-LANSS), age, and deficit in elbow extension range of motion during $ULNT1_{MEDIAN}$ may help clinicians with this identification process, because the model's cutoffs signify a clinically important increase or decrease in a patient's chance for improvement with NTM in the short term. Patients who score 89 or more points on the model appear to be good candidates for NTM, because their chance for improvement increases to 90% (95% CI: 56%, 98%). Those who score 71 or fewer points should probably receive an intervention other than NTM initially, because their chance for improving with neurodynamic treatment decreases to 9% (95% CI: 1%, 42%). The model may help clinicians decide whether to use NTM in approximately 50% of patients who meet our definition of nerve-related neck and arm pain, because 31 of 60 participants had prediction model scores of 89 or greater or of 71 or less. The other 50% of patients who scored less than 89 or greater than 71 may still benefit from NTM, but the model cannot assist clinicians' decision-making process

for these individuals. Future studies need to determine whether broadening the definition of nerve-related neck and arm pain or modifying the prediction model will enable the model to assist clinicians' decision-making process for a larger proportion of patients.

It is unlikely that the model simply predicts the natural history of the patient's pain. The previously reported results of the randomized trial showed that the outcome predicted by the model (participant-reported improvement) represents an NTM treatment effect.[51] More importantly, the post hoc analysis indicated that the model did not predict outcome in the ARA group.[26,42] None of the 5 participants who received ARA and scored 89 or more points on the model were improved at follow-up, and the 1 participant who did improve with ARA scored less than 89 points (**TABLE 4**).

It should be emphasized that, consistent with the questionnaire's instructions, participants were reminded to focus specifically on the area of their worst symptoms when completing the S-LANSS.[5] For example, a participant whose worst symptoms were in the neck or scapular region completed the S-LANSS based on these symptoms, not arm symptoms. This distinction is important, because recent evidence suggests that questionnaire-based classifications of neuropathic pain may differ between local spine symptoms and associated limb symptoms.[3] The S-LANSS must be administered with this reminder when using this prediction model.

Although our results suggest that the model is useful, an assessment of usefulness should also consider whether the individual predictors in the model make sense clinically (face validity).[44] S-LANSS scores of 12 or greater (presence of neuropathic pain qualities) and larger deficits in $ULNT1_{MEDIAN}$ range of motion were associated with a lower likelihood of improving with the nonprovocative NTM program. Both predictors suggest a higher level of nervous system sensitivity. When nervous system sensitivity is relatively high, NTM techniques that are still nonprovocative may need to be so gentle that the dosage could not provide enough of a mechanical or neurophysiological stimulus to produce a treatment effect. Therefore, S-LANSS scores of 12 or greater and larger deficits in $ULNT1_{MEDIAN}$ range of motion predicting a lower likelihood of improvement with the nonprovocative NTM program seem clinically reasonable. These predictors suggest that for patients with a highly sensitive nervous system, a different approach may be required initially so that the treatment dosage delivers an adequate stimulus to produce beneficial effects without aggravating highly sensitized neural tissues.

The relationship between younger age and a lower likelihood of improvement with NTM is more difficult to interpret. A prospective cohort study on physical therapy treatment of nerve-related neck and arm pain showed the opposite relationship; younger age was associated with improvement in the short term.[12] It seems unlikely that there was something specific in the NTM treatment used in the present study or in its participants that would explain the difference between our prediction model and the results of this prospective cohort study. While the relationship between younger age and a lower likelihood of improvement is difficult to explain, age was a necessary predictor for creating the model with the lowest expected prediction error (**FIGURE 3**). This situation highlights the concept that items that predict whether an outcome occurs do not necessarily explain how the prediction model treatment achieves its effects.[39,43] It is important to recognize that the aim of the model was to predict the likelihood of improvement with NTM rather than to identify explanatory mechanisms for NTM treatment effects.

Creating the model by analyzing only NTM data requires further discussion. The ideal approach to creating a treatment-related prediction model from randomized controlled trial data is to include data from both the treatment group (NTM) and control group (ARA) in the analysis.[43] This allows for testing of the interactions between the treatment group assignment and each examination item considered for inclusion in the model. An interaction is present when the strength of the relationship between an examination item and the outcome is different for each treatment group.[39] Only examination items with significant interactions are selected for the final model. This approach is preferred because it establishes that each examination item in the final model predicts a specific response to NTM.[42,43] It was not possible to analyze both NTM and ARA data when creating our prediction model, because only 1 participant who received ARA was improved at follow-up. This meant that there were not enough data for the contrasts necessary to identify valid interactions between each examination item and treatment-group assignment.[39] We anticipated this problem when planning the study,[52] because previous randomized controlled trial data showed that the short-term natural history of nerve-related neck and arm pain was relatively stable.[1] Therefore, we created the prediction model by analyzing only NTM data and assessed its ability to predict a specific response to NTM through post hoc analysis. This is considered a viable approach to generating an initial hypothesis for a treatment-related prediction model.[42] However, it is unclear whether this approach would have any detrimental impact on predictive performance in future patients.

The results of this study need to be interpreted with 3 additional considerations. First, the sample was smaller than planned (60 participants instead of 84) because of the study's limited recruitment time (2 years).[51] Some examination items, therefore, could not be entered into the CATREG LASSO analysis. This might also have contributed to the wide 95% CIs for the likelihood ratios associated with the prediction model cutoffs and the assessment of the model's ability to predict a specific response to NTM. Second, this was the initial proposal of the predic-

tion model. Although the LASSO analysis and preservation of continuous examination item information in the medical nomogram may increase the chance that the model will maintain its predictive performance in future patients, this performance needs to be confirmed in a new sample before the model can be recommended for widespread clinical use.[47] Third, the assessment of the model's ability to predict a specific response to NTM is directly related to the comparison intervention used in the randomized trial, the time frame of follow-up, and the threshold that defines improvement.[41] Patients who meet this study's enrollment criteria and have prediction model scores of 89 or greater are likely to do much better with NTM than with ARA in the short term. However, it is unclear whether these patients would do better with NTM than with other interventions and whether the model would still predict a specific response to NTM over a longer follow-up or with different thresholds for improvement. Future randomized trials are needed to answer these questions.

In addition to re-evaluating model performance, future randomized trials will provide insight into the potential impact that the model may have on clinical decision making. If NTM treatment effects remain relatively large compared to different interventions, NTM could be considered as "best practice" for most patients with nerve-related neck and arm pain, and selecting patients according to the prediction model may not be necessary.[23] However, if NTM treatment effects are similar to other interventions, selecting patients according to the prediction model (if validated) may help to maximize patient outcomes.[23]

## CONCLUSION

When treating patients who meet our definition of nerve-related neck and arm pain with NTM, absence of neuropathic pain qualities (S-LANSS scores less than 12), older age, and smaller deficits in $ULNT1_{MEDIAN}$ range of motion predicted an increase in the chance for improvement at a 3- to 4-week follow-up. Conversely, presence of neuropathic pain qualities (S-LANSS scores of 12 or greater), younger age, and larger deficits in $ULNT1_{MEDIAN}$ range of motion predicted a decrease in the chance for improvement. Validation of prediction model performance in a new sample is needed, as is an assessment of the model's ability to predict a specific response to NTM when using different comparison interventions and a longer follow-up. ◉

### KEY POINTS

**FINDINGS:** Absence of neuropathic pain qualities, older age, and smaller deficits in $ULNT1_{MEDIAN}$ range of motion predicted whether patients with nerve-related neck and arm pain were likely to improve with NTM at a 3- to 4-week follow-up. Prediction model cutoffs increased a patient's chance for improvement from 53% to 90% or decreased it to 9%. Post hoc analysis suggested that the model predicted a specific response to NTM.

**IMPLICATIONS:** This prediction model is a research-based tool that provides preliminary evidence that baseline characteristics may be able to identify patients with nerve-related neck and arm pain who are and are not a good match for NTM.

**CAUTION:** Model performance needs to be re-evaluated in a validation study. The model's ability to predict a specific response to NTM when using different comparison interventions and longer follow-up is unknown.

ACKNOWLEDGEMENTS: *The authors thank the participants and physical therapists involved in the trial and Benjamin Soon Tze Chin and Lieszel Plumbe for assistance with randomization.*

## REFERENCES


1. Allison GT, Nagy BM, Hall T. A randomized clinical trial of manual therapy for cervicobrachial pain syndrome – a pilot study. *Man Ther*. 2002;7:95-102.
2. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst*. 1994;86:829-835.
3. Attal N, Perrot S, Fermanian J, Bouhassira D. The neuropathic components of chronic low back pain: a prospective multicenter study using the DN4 questionnaire. *J Pain*. 2011;12:1080-1087. http://dx.doi.org/10.1016/j.jpain.2011.05.006
4. Beattie P, Nelson R. Clinical prediction rules: what are they and what do they tell us? *Aust J Physiother*. 2006;52:157-163.
5. Bennett MI, Smith BH, Torrance N, Potter J. The S-LANSS score for identifying pain of predominantly neuropathic origin: validation for use in clinical and postal research. *J Pain*. 2005;6:149-158. http://dx.doi.org/10.1016/j.jpain.2004.11.007
6. Bogduk N. The anatomy and pathophysiology of neck pain. *Phys Med Rehabil Clin N Am*. 2011;22:367-382. http://dx.doi.org/10.1016/j.pmr.2011.03.008
7. Butler DS. *The Sensitive Nervous System*. Adelaide, Australia: Noigroup Publications; 2000.
8. Childs JD, Cleland JA, Elliott JM, et al. Neck pain: clinical practice guidelines linked to the International Classification of Functioning, Disability, and Health from the Orthopaedic Section of the American Physical Therapy Association. *J Orthop Sports Phys Ther*. 2008;38:A1-A34. http://dx.doi.org/10.2519/jospt.2008.0303
9. Childs JD, Fritz JM, Piva SR, Whitman JM. Proposal of a classification system for patients with neck pain. *J Orthop Sports Phys Ther*. 2004;34:686-700. http://dx.doi.org/10.2519/jospt.2004.1451
10. Cleland JA, Childs JD, Fritz JM, Whitman JM, Eberhart SL. Development of a clinical prediction rule for guiding treatment of a subgroup of patients with neck pain: use of thoracic spine manipulation, exercise, and patient education. *Phys Ther*. 2007;87:9-23. http://dx.doi.org/10.2522/ptj.20060155
11. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil*. 2008;89:69-74. http://dx.doi.org/10.1016/j.apmr.2007.08.126
12. Cleland JA, Fritz JM, Whitman JM, Heath R. Predictors of short-term outcome in people with a clinical diagnosis of cervical radiculopathy. *Phys Ther*. 2007;87:1619-1632. http://dx.doi.org/10.2522/ptj.20060287
13. Concato J, Feinstein AR, Holford TR. The risk

of determining risk with multivariable models. *Ann Intern Med*. 1993;118:201-210.
14. Coppieters MW, Butler DS. Do 'sliders' slide and 'tensioners' tension? An analysis of neurodynamic techniques and considerations regarding their application. *Man Ther*. 2008;13:213-221. http://dx.doi.org/10.1016/j.math.2006.12.008
15. Daffner SD, Hilibrand AS, Hanscom BS, Brislin BT, Vaccaro AR, Albert TJ. Impact of neck and arm pain on overall health status. *Spine (Phila Pa 1976)*. 2003;28:2030-2035. http://dx.doi.org/10.1097/01.BRS.0000083325.27357.39
16. de Vet HC, Heymans MW, Dunn KM, et al. Episodes of low back pain: a proposal for uniform definitions to be used in research. *Spine (Phila Pa 1976)*. 2002;27:2409-2416. http://dx.doi.org/10.1097/01.BRS.0000030307.34002.BE
17. Di Fabio RP, Boissonnault W. Physical therapy and health-related outcomes for patients with common orthopaedic diagnoses. *J Orthop Sports Phys Ther*. 1998;27:219-230.
18. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J Am Stat Assoc*. 1983;78:316-331. http://dx.doi.org/10.1080/01621459.1983.10477973
19. Elvey RL. Physical evaluation of the peripheral nervous system in disorders of pain and dysfunction. *J Hand Ther*. 1997;10:122-129.
20. Elvey RL. Treatment of arm pain associated with abnormal brachial plexus tension. *Aust J Physiother*. 1986;32:225-230.
21. Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine (Phila Pa 1976)*. 2002;27:2835-2843. http://dx.doi.org/10.1097/01.BRS.0000035681.33747.8D
22. Foster NE, Dziedzic KS, van der Windt DA, Fritz JM, Hay EM. Research priorities for non-pharmacological therapies for common musculoskeletal problems: nationally and internationally agreed recommendations. *BMC Musculoskelet Disord*. 2009;10:3. http://dx.doi.org/10.1186/1471-2474-10-3
23. Fritz JM. Clinical prediction rules in physical therapy: coming of age? *J Orthop Sports Phys Ther*. 2009;39:159-161. http://dx.doi.org/10.2519/jospt.2009.0110
24. Fu AZ, Cantor SB, Kattan MW. Use of nomograms for personalized decision-analytic recommendations. *Med Decis Making*. 2010;30:267-274. http://dx.doi.org/10.1177/0272989X09342278
25. Hall TM, Elvey RL. Nerve trunk pain: physical diagnosis and treatment. *Man Ther*. 1999;4:63-73. http://dx.doi.org/10.1054/math.1999.0172
26. Hancock M, Herbert RD, Maher CG. A guide to interpretation of studies investigating subgroups of responders to physical therapy interventions. *Phys Ther*. 2009;89:698-704. http://dx.doi.org/10.2522/ptj.20080351
27. Harrell FE, Jr. RMS: regression modeling strategies version 3.3-1. Available at: http://CRAN.R-project.org/package=rms. Accessed August 1, 2011.
28. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361-387. http://dx.doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4
29. Hartmann A, Van Der Kooij AJ, Zeeck A. Exploring nonlinear relations: models of clinical decision making by regression with optimal scaling. *Psychother Res*. 2009;19:482-492. http://dx.doi.org/10.1080/10503300902905939
30. Hartmann A, Zeeck A, van der Kooij AJ. Severity of bulimia nervosa. Measurement and classification into health or pathology. *Psychopathology*. 2009;42:22-31. http://dx.doi.org/10.1159/000173700
31. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York, NY: Springer; 2009.
32. Hosmer D, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York, NY: Wiley-Interscience; 2000.
33. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*. 1994;271:703-707.
34. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10:407-415.
35. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain*. 1994;58:387-392.
36. Jette DU, Jette AM. Physical therapy and health outcomes in patients with spinal impairments. *Phys Ther*. 1996;76:930-941; discussion 942-945.
37. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther*. 2009;17:163-170.
38. Kattan MW, Marasco J. What is a real nomogram? *Semin Oncol*. 2010;37:23-26. http://dx.doi.org/10.1053/j.seminoncol.2009.12.003
39. Katz MH. Multivariable analysis: a primer for readers of medical research. *Ann Intern Med*. 2003;138:644-650.
40. Keegan PE, Matthews JN, Lunec J, Neal DE. Statistical problems with 'optimal' thresholds in studies of new prognostic factors in urology. *BJU Int*. 2000;85:392-397.
41. Kent P, Hancock M, Petersen DH, Mjøsund HL. Clinimetrics corner: choosing appropriate study designs for particular questions about treatment subgroups. *J Man Manip Ther*. 2010;18:147-152. http://dx.doi.org/10.1179/106698110X12640740712419
42. Kent P, Keating JL, Leboeuf-Yde C. Research methods for subgrouping low back pain. *BMC Med Res Methodol*. 2010;10:62. http://dx.doi.org/10.1186/1471-2288-10-62
43. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry*. 2002;59:877-883.
44. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA*. 1997;277:488-494. http://dx.doi.org/10.1001/jama.1997.03540300056034
45. Ma S, Huang J. Penalized feature selection and classification in bioinformatics. *Brief Bioinform*. 2008;9:392-403. http://dx.doi.org/10.1093/bib/bbn027
46. MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychol Methods*. 2002;7:19-40.
47. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA*. 2000;284:79-84.
48. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med*. 1978;8:283-298.
49. Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika*. 1991;78:691-692. http://dx.doi.org/10.1093/biomet/78.3.691
50. Nee RJ, Jull GA, Vicenzino B, Coppieters MW. The validity of upper-limb neurodynamic tests for detecting peripheral neuropathic pain. *J Orthop Sports Phys Ther*. 2012;42:413-424. http://dx.doi.org/10.2519/jospt.2012.3988
51. Nee RJ, Vicenzino B, Jull GA, Cleland JA, Coppieters MW. Neural tissue management provides immediate clinically relevant benefits without harmful effects for patients with nerve-related neck and arm pain: a randomised trial. *J Physiother*. 2012;58:23-31. http://dx.doi.org/10.1016/S1836-9553(12)70069-3
52. Nee RJ, Vicenzino B, Jull GA, Cleland JA, Coppieters MW. A novel protocol to develop a prediction model that identifies patients with nerve-related neck and arm pain who benefit from the early introduction of neural tissue management. *Contemp Clin Trials*. 2011;32:760-770. http://dx.doi.org/10.1016/j.cct.2011.05.018
53. Pincus T, Miles C, Froud R, Underwood M, Carnes D, Taylor SJ. Methodological criteria for the assessment of moderators in systematic reviews of randomised controlled trials: a consensus study. *BMC Med Res Methodol*. 2011;11:14. http://dx.doi.org/10.1186/1471-2288-11-14
54. Puentedura EJ, Cleland JA, Landers MR, Mintken PE, Louw A, Fernández-de-las-Peñas C. Development of a clinical prediction rule to identify patients with neck pain likely to benefit from thrust joint manipulation to the cervical spine. *J Orthop Sports Phys Ther*. 2012;42:577-592. http://dx.doi.org/10.2519/jospt.2012.4243

**55.** Randolph AG, Guyatt GH, Calvin JE, Doig G, Richardson WS. Understanding articles describing clinical prediction tools. Evidence Based Medicine in Critical Care Group. *Crit Care Med*. 1998;26:1603-1612.

**56.** Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*. 2006;25:127-141. http://dx.doi.org/10.1002/sim.2331

**57.** Salt E, Wright C, Kelly S, Dean A. A systematic literature review on the effectiveness of non-invasive therapy for cervicobrachial pain. *Man Ther*. 2011;16:53-65. http://dx.doi.org/10.1016/j.math.2010.09.005

**58.** Schellingerhout JM, Heymans MW, de Vet HC, Koes BW, Verhagen AP. Categorizing continuous variables resulted in different predictors in a prognostic model for nonspecific neck pain. *J Clin Epidemiol*. 2009;62:868-874. http://dx.doi.org/10.1016/j.jclinepi.2008.10.010

**59.** Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York, NY: Springer; 2009.

**60.** Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol*. 1996;49:907-916.

**61.** Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16:385-395. http://dx.doi.org/10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3

**62.** Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther*. 1991;14:409-415.

**63.** Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ*. 2001;323:1123-1124.

**64.** Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol*. 2007;165:710-718. http://dx.doi.org/10.1093/aje/kwk052

**65.** Walker MJ, Boyles RE, Young BA, et al. The effectiveness of manual physical therapy and exercise for mechanical neck pain: a randomized clinical trial. *Spine (Phila Pa 1976)*. 2008;33:2371-2378. http://dx.doi.org/10.1097/BRS.0b013e318183391e

**66.** Westaway MD, Stratford PW, Binkley JM. The Patient-Specific Functional Scale: validation of its use in persons with neck dysfunction. *J Orthop Sports Phys Ther*. 1998;27:331-338.