Pain Drawing Scoring Is Not Improved by Inclusion of Patient-Reported Pain Sensation

Neal W. Sanders, PhD,*† N. Horace Mann, III, PhD;‡ and Dan M. Spengler, MD*

Study Design. This is a retrospective study of 250 patients who describe low back pain with pain drawings. A computer application using artificial neural networks was designed to analyze pain drawings and evaluate the contribution of patient sensation to drawing classification.

Objective. The primary goal of this study was to assess the contribution of patient recorded sensation marks in classifying pain drawings into one of five broadly defined categories. The hypothesis was that including patient sensation would improve classification.

Summary of Background Data. With no perfect diagnostic test for patients with low back pain, many approaches have been proposed and are used. One common diagnostic tool is the pain drawing. Several quantitative methods have been proposed to score the drawings. Some methods use pain sensation in the scoring; however, the contribution of pain sensation has not been defined.

Methods. A custom computer application classified the pain drawing. Data consisted of 250 pain drawings from patients with low back pain.

Results. Patient recorded pain sensation is not necessary in computer-based scoring of pain drawings.

Conclusion. Patient-reported pain sensation does not improve classification when quantitatively scoring pain drawings.

Key words: low back pain, pain measurement, artificial intelligence, neural networks computer. Spine 2006;31:2735–2741

A patient completed pain drawing is often used by clinicians to obtain a quick assessment of the patient’s pain patterns, and it is used in conjunction with other assessment tools as an aid in developing a treatment plan.1,2 These drawings are often assessed by a clinician “at a glance” for immediate clinical utility; however, there are also more formal methods of scoring the drawings.3–10

The use of pain drawings has evolved over time from a simple tool to facilitate communication regarding pain between the patient and clinician to a more complex tool that is being used for evaluating psychometrics and pain classification. In the earliest uses of pain drawings, the patient marked pain locations and the clinician added pain sensation to the drawing as the patient discussed their pain.2 More recently, the patient has also indicated their pain sensation (“burning,” “aching,” “stabbing,” etc.) on the drawing.9 This pain sensation information is often used for evaluation, especially as related to poor psychometrics. The value of the pain drawing as an indicator of poor psychometrics has been inconclusively debated; however, many clinicians consider the pain drawing as a valuable tool in the assessment of patients with low back pain.4,9,11–15

Although pain sensation is incorporated in many scoring methods, the contribution of pain sensation to scoring pain drawings has not been studied. The most common method for patient coding of pain sensation is to mark the drawing using a specific symbol to indicate various types of pain. These symbols indicate a type of pain sensation, and their location indicates pain patterns. In studies to examine the reliability of a patient to repeat a pain drawing, it was found that the patient can reliably repeat the location of their pain; however, they were less reliably able to repeat their pain sensation.8 Their finding of pain sensation being less reliably repeated than pain location would be expected given that they only included patients who stated that their pain location had not changed, but did not ask about pain sensation changes. Other studies, using pain questionnaires, have concluded that pain sensations seem less useful than other pain descriptors (e.g., onset, locus, and severity) likely because sensation is too subjective; however, they also stated that the subjectiveness characteristics might assist in differentiating certain conditions.16

Until recently, scoring methods of pain drawings were manual and many required significant effort to calculate a score. Newer scoring methods are now being designed using automated computer scoring.5,6,10,17 Most of these automated methods consider spatial-anatomic distribution of pain markings; none has incorporated pain sensation. The addition of pain sensation information to a spatial-anatomic scoring algorithm will increase the information available to the algorithm to more closely match that used by a clinician when considering the drawings “at a glance” or the manual scoring methods that use pain sensation.

It is hypothesized that including pain sensation information in computer-based scoring algorithms will improve the accuracy of existing nonsensation scoring algorithms. This study was undertaken to determine how to best represent pain sensation such that it can be de-
ected by a computer and then to evaluate the contribution of pain sensation in the scoring.

**Materials and Methods**

This study was designed to assess the role of pain sensation when scoring pain drawings. The study required creating two models for scoring. One model included the pain sensation information and the other model excluded the pain sensation information. Both models included the spatial-anatomic distribution of pain.

**Study Data.** The data for the study were 250 deidentified patient pain drawings that were selected from the cases of an orthopedic surgeon. The inclusion criteria were a nonconflicting diagnosis based on a retrospective review of the complete patient case, after the patient’s treatment was completed, by an orthopedic surgeon. The case review information included all relevant clinic notes, image findings, surgical reports, and treatments; however, it did not include the pain drawing. The drawings were completed by English-speaking patients on their first outpatient clinic visit between 1979 and 1989. These drawings have been used in several previous studies.6,10,18,19

The pain drawings were completed on standard 8.5” by 11” white paper, which contained a gender neutral anterior and posterior view outline of a body (Figure 1). The drawing also included directions for the patient indicating how to represent six pain sensations. The directions for the colors and marks to represent the sensations were:

- Numbness (yellow) —
- Pins and Needles (purple) ooo
- Burning (red) xxx
- Aching (blue) +++
- Stabbing (green) ///
- Other (brown) ***

Because of practical issues related to maintaining and distributing six color markers, the patient used only the sensation marks to indicate their pain sensation. The colors (used by the
computer to detect type of sensation) representing the sensation were added later by a trained research assistant.

Variables. The dependent variable for the study models was a diagnostic category classification. The diagnostic categories were selected for previous studies of these pain drawings and are general enough to cover most of the cases seen in a low back pain clinic while being specific enough for clinical usefulness. The five diagnostic categories were:

1. Benign Back Pain Disorder (BD). These disorders have in common a harmless self-limiting course and usually cannot be related to any specific or aggravating factors.
2. Herniation of the Nucleus Pulposus (HNP). HNP is a rupture or displacement of material through the soft tissue of the intervertebral disc.
3. Spinal Stenosis (SS). This disorder is a narrowing of the spinal canal and foramen, whether it is of congenital, developmental, or degenerative origin.
4. Serious Underlying Disorders (UD). Underlying disorders encompasses abnormalities, which present symptoms of orthopedic back pain but are indicative of some other type of disorder, such as metastatic cancer.
5. Psychogenic Regional Pain Disturbance (PSY). This category is for patients who have symptoms that appear to be behavioral in origin. These symptoms may appear alone or concurrently with an organic problem.

Each of the 250 drawings, 50 in each category, was classified into one of these categories by the orthopaedic surgeon based on their review of the patient case after the course of treatment was completed.

The independent variables were based on the pain drawing’s pain markings. The drawing was segmented into regions that corresponded to low back-related dermatomes and other gross anatomic regions as well as nonbody regions of the drawing. For each region, the percentage of that region with any pain mark was the pain variable for that region. In addition, for each region, the percentage of each pain sensation mark in that region was computed and used as a variable for the pain sensation model.

Methods. Custom software was developed to digitize the pain drawings, determine the pain markings and sensations in each region, and create a model for scoring the drawings.

Image processing. Using a custom software program, the paper drawings were scanned into a computer and stored in a database. The software also aligned the drawing, counted individual pixels in each defined region, determined the pain sensation represented by each pixel, and stored the individual pain sensation counts (based on pixel color) per region in the database.

Model. An artificial neural network was chosen to model the drawing scoring. Previous research has been successful using neural networks to score pain drawings. In addition, this provided an automated scoring method that did not require the statistical assumptions or other mathematical formalization of traditional statistics.

An artificial neural network was designed for evaluating various configurations of the scoring model. Neural networks may be designed using various algorithms; this network used a back-propagation algorithm because it is well suited for classification. The inputs to the neural network were the independent variables and the outputs represented each of the five classifications. The inputs were a percentage of the selected region that contained a pain (or pain sensation) mark. Each
output allowed a range from zero to one and represented the strength, or certainty, of that classification.

A neural network is “trained” to create the desired model. This training is a well-defined mathematical process of iteratively presenting the neural network with inputs and the desired output for that input. During training, the back-propagation algorithm adjusts the model until it is able to correctly compute the outputs for a set of inputs. When a network is trained, an input set is applied and the output is computed. There are no well-accepted rules for how much data are needed to properly train an artificial neural network. The complexity of a network can be estimated by its number of inputs and outputs. As the complexity increases, the quantity of data needed to properly train it also increases. To minimize the potential impact of low data quantity, the complexity of the neural network was minimized by reducing the number of input variables. From the most granular set of 115 regions, various combinations were tested to find those yielding the best scoring performance. The empirically derived goal was to decrease the number of input variables to less than 200 when pain sensations were used. The desired number of regions was less than 30 due to using 7 variables per region (six pain sensation variables and a total pain variable for each region).

The network was trained by splitting the data into a training set of 40 randomly selected drawings per classification and a test set consisting of the remaining 10 drawings. When the network was trained, the model was completed and the test data were applied. For each test input, the output with the greatest value was compared with the drawing’s classification to determine the correctness of the model for that test input. After storing the test results, the training and test set was resampled and the training and testing of the model was repeated for 120 trials (the number of trials was determined empirically based on observing the stability of the results). The average percent correct classification in each of the five diagnostic categories was computed from the 120 trials. Others have used similar resampling techniques to reduce small-sample bias and to estimate performance variations.

**Results**

Two patient pain drawing scoring models were developed to evaluate the impact of including pain sensation information in the scoring model. The classification accuracy of the model incorporating pain sensation was approximately the same as the nonsensation model.

The pain drawings were characterized by examining the pain sensation in each classification and by the distribution of pain marks within each classification (Figures 2, 3). Patients classified as psychogenic regional pain disturbance had the greatest quantity of marks in each sensation. Patients classified as benign back pain disorder had over 50% of their marks in the posterior lower back region.

The complexity of the model was reduced by combining the spatial-anatomic regions that were used as the input variables. Sixteen combinations of the base 115 regions were tested for their ability to represent the pain drawing for scoring (Table 1). These tests were per-
The pain sensation model was tested using 10 spatial-anatomic regions. This model had 70 inputs. Each region was represented by each of the 6 pain sensations and 1 sum of pain marks. The sum of pain marks was used to help emphasize to the model the overall spatial pain pattern. This pain sensation model had a sensitivity of 43.0% in classifying the drawings. Of the individual classifications, benign back pain disorders classified the best and spinal stenosis classified the worst (Figure 5).

The nonpain sensation model classified across all categories at 43.6% sensitivity and the pain sensation model classified at 43.0% across all categories and are not statistically different ($P < 0.05$). The individual categories sensitivities varied only slightly between the models. The spinal stenosis category was the only category with sensitivity below random chance.

**Discussion**

This study was designed to determine the contribution of pain sensation in computer-based pain drawing scoring algorithms. It was hypothesized that including pain sensation information as an input to the scoring algorithm would improve the ability of the resulting model to correctly classify the pain drawing when compared with a model that did not include pain sensation information.

This study was limited to 250 previously studied and classified patient pain drawings. This limitation was due to the difficulty in obtaining properly classified drawings. Because adding 6 pain sensations to the scoring algo-

---

**Table 1. Noteworthy Spatial-Anatomic Region Groupings That Were Tested When Reducing the Complexity of the Pain Drawing Scoring Model**

<table>
<thead>
<tr>
<th>No. of Regions</th>
<th>Description of Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Posterior lower back (L1–L5, S0–S2)</td>
</tr>
<tr>
<td></td>
<td>Posterior left leg (all area on posterior of left leg and foot)</td>
</tr>
<tr>
<td></td>
<td>Posterior right leg (all area on posterior of right leg and foot)</td>
</tr>
<tr>
<td></td>
<td>Posterior upper back (C5–C8, T1–T12)</td>
</tr>
<tr>
<td></td>
<td>Posterior arms (all area on posterior of both arms and hand)</td>
</tr>
<tr>
<td></td>
<td>Anterior chest (C4, C5, T1–T12)</td>
</tr>
<tr>
<td></td>
<td>Anterior arms (all area on anterior of both arms and hand)</td>
</tr>
<tr>
<td></td>
<td>Anterior right leg (all area on anterior of right leg and foot)</td>
</tr>
<tr>
<td></td>
<td>Anterior left leg (all area on anterior of left leg and foot)</td>
</tr>
<tr>
<td></td>
<td>Other regions (including outside of body)</td>
</tr>
<tr>
<td>17</td>
<td>Same as 10 above, but with posterior lower back broken into 8 dermatomal regions</td>
</tr>
<tr>
<td>25</td>
<td>Same as 17 above, but with each posterior right and left leg broken into 5 regions each</td>
</tr>
<tr>
<td>29</td>
<td>Same as 17 above, but with each anterior right and left leg broken into 7 regions each</td>
</tr>
<tr>
<td>37</td>
<td>Same as 17 above, but with both anterior and posterior legs broken into regions as described above</td>
</tr>
<tr>
<td>77</td>
<td>This includes dermatomal regions for all parts of the body as well as regions outside the body</td>
</tr>
</tbody>
</table>

---

![Figure 4](image-url)  
*Figure 4. The nonpain sensation scoring model’s sensitivity and specificity for each category, across 120 resample trials, when using 10 spatial-anatomic regions.*
rithm would result in a model that was more complex than one without pain sensations, the number of spatial-anatomic regions used for model inputs had to be reduced. The artificial neural network model should be capable of discounting or combining the inputs through its internal modeling algorithms if the combination of regions proved a better model. The neural network model could have been allowed to provide this capability; however, doing so would have resulted in a more complex model and thus required more data than was available. Although the exact quantity of data needed to properly train an artificial neural network is not known, an insufficiency of data can be observed in a large difference in the scoring sensitivity of the model when the training set data are measured against the test set data. With 10 regions, the difference between the training and test set sensitivity was not large enough to be a concern; thus, the 250 drawings were sufficient to properly train the neural network.

Another study concern was that reducing the number of regions risked degrading the models scoring capability; however, testing to reduce the number of input regions unexpectedly found that as few as 10 regions (Table 1) allowed the model to score as well as 77 regions. The one exception to this was the spinal stenosis category, which scored better when using 77 regions. This may be an indication that spinal stenosis has a more complex pain pattern than the other categories.

The drawings used to construct and test this model were distributed evenly among 5 classifications. This was a forced distribution and the typical clinic distribution of patients in each of the classifications it is not known. However, it is not necessarily preferred to train the neural network model using a true distribution.

These drawings have been previously scored using an artificial neural network model and a subset of them have been scored by physician experts.10,18 The scoring capability of this model compared favorably to both the previous nonsensation neural network model (which used 85 regions) and to the physician experts who subjectively classified the drawings.

No research literature was found that discusses using pain sensation in a computer-based scoring model. Although other studies have indicated that pain sensation may be of less value than pain location, neither specifically examined the contribution of pain sensation in scoring pain drawings.8,16 Pain sensation is widely used in the manual “Ransford scoring” methods4,9; however, the utility of the sensations indicating a psychogenic component of pain is debated.14,15 There may also be a problem with the additional difficulty of the patient understanding the directions. Although all patients in the study spoke English, it is not known if English was their first language. Asking the patient to indicate where they hurt is much simpler than asking them to indicate via a specific type of mark the pain sensation at each location.

This study finds that patient-reported pain sensation is not necessary in computer-based models for scoring pain drawings. This finding greatly reduces the complex-

Figure 5. Pain sensation scoring model’s sensitivity and specificity for each category, across 120 resample trials, when using 10 spatial-anatomic regions.
ity of the process for digitally processing the images as well as the neural network scoring model. It may also eventually allow the patient to indicate only where they are experiencing pain and not require them to also indicate pain sensation. The study results do not suggest that pain sensation is not important to the clinician; however, the results do suggest that, when patient supplied, the pain pattern contains the useful information. In the initial use of pain drawings, the clinician would elicit from the patient the pain sensation at the indicated locations rather than using patient-supplied sensations. Perhaps this clinician-elicited and -interpreted pain sensation information is of greater value in scoring the drawings than the patient-supplied pain sensation information. Currently, this automated scoring is only used for research related to the pain drawings. For clinical purposes, we continue to ask the patient to indicate pain sensation as this facilitates a discussion between the patient and clinician.

The cost for implementing this system is minimal. It requires a computer running Microsoft Windows (Microsoft Corp., Redmond, WA), a flat-bed scanner, and a database. The database functionality may be obtained from freely available databases and the computer may be shared with other tasks. They cost of implementing this system could be as low as a few hundred dollars.

**Future Direction**

There is a need to obtain a much larger set of classified pain drawings along with other patient demographics information and assessment questionnaires. The additional data and information should allow the creation of a more robust model that eventually, along with other information sources, will lead to a decision support tool for patients with low back pain.

### Key Points

- Pain drawings were analyzed for the contribution of pain sensation in classification.
- A customized computer application was used to classify the drawings.
- Patient-reported pain sensation does not improve classification when quantitatively scoring pain drawings.

### Acknowledgments

The authors thank Dr. Mark Brown at the University of Miami for providing the pain drawings used in this study.

### References