A Validated Rating Scale for Hyperkinetic Facial Lines

Eugene J. Kim, MD; Jay B. Reeck, MD; Corey S. Maas, MD

Objective: To test the reliability of a simple rating system to describe hyperkinetic facial lines.

Methods: A rated numeric kinetic line scale was developed and presented to 11 postresidency physicians specializing in aesthetic facial care. These physicians independently reviewed photographs of 20 patients, first at rest, then with activation of the frontalis, corrugator, and orbicularis oculi muscles. Kappa statistics for multiple raters were used to assess interobserver reliability.

Results: The nonweighted \( \kappa \) values were between 0.4 and 0.8 for the frontalis, corrugator, and orbicularis muscle groups. This represents moderate to substantial observer agreement and is highly significant for each muscle group.

Conclusions: A new rating scale for hyperkinetic facial lines accounts for facial appearance at rest and with expression. It is easily used and has interobserver reliability. As the only objective and validated scale for hyperkinetic facial lines, this rated numeric kinetic line scale is recommended for the evaluation of pretreatment and posttreatment results in patients undergoing therapy for this problem. Moreover, an alternative scale rating resting and kinetic lines as independent variables is also being developed. Both must be considered to evaluate treatment outcomes when using neurotoxins.

Arch Facial Plast Surg. 2004;6:253-256
Hyperkinetic facial lines could provide a method for treatment stratification of patients. The ideal rating system for hyperkinetic facial lines would be easy to use, noninvasive, allow for repeated evaluations, and show interobserver reliability. In addition, this rating system would account for the dynamic nature of hyperkinetic facial lines as subjects are evaluated at rest and with effort of the facial musculature. We developed a rating system for hyperkinetic facial lines and tested its reliability.

**METHODS**

The Rated Numeric Kinetic Line Scale (RNKLS) was developed to rate hyperkinetic facial lines in the context of facial expression. Dermal lines or lines of relaxed skin tension are excepted from the scale as they are not secondary to hyperkinetic function. A PowerPoint (Microsoft Corp, Redmond, Wash) slide presentation was created for each of the three muscle groups: frontalis, corrugator, and orbicularis oculi. Each presentation explained the rating system, showed photographic examples of each score (RNKLS 0-4), and presented a series of 20 subjects at rest and with expression (Figures 1, 2, and 3). Eleven postresidency physicians specializing in aesthetic facial care reviewed the presentations and independently assigned RNKLS scores to each series of 20 photographs.

Interobserver reliability was analyzed using a nonweighted $\kappa$ statistic for each of the three muscle groups. Nonweighted $\kappa$ statistics measure interobserver agreement and represent a standard method for evaluating reliability among multiple raters. More specifically, they test whether there is more agreement between observers than can be expected by chance.

**RESULTS**

Nonweighted $\kappa$ statistics for the 11 observers’ hyperkinetic facial lines caused by the frontalis, corrugator, and orbicularis oculi muscles are shown in the following tabulation:

<table>
<thead>
<tr>
<th>Muscle Group</th>
<th>$\kappa$ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frontalis</td>
<td>0.64</td>
</tr>
<tr>
<td>Corrugator</td>
<td>0.52</td>
</tr>
<tr>
<td>Orbicularis oculi</td>
<td>0.43</td>
</tr>
</tbody>
</table>

As $\kappa$ scores approach 1.0, the degree of interobserver agreement increases. Although arbitrary, $\kappa$ statistics are labeled according to ranges originally described by Landis and Koch. They are poor (<0.00), slight (0.00-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80), and almost perfect (0.81-1.00). The frontalis ratings (0.64) showed substantial interobserver agreement. The corrugator (0.52) and orbicularis (0.43) both yielded moderate levels of agreement. All 3 series were highly statistically significant ($P < .001$).

**COMMENT**

The RNKLS has demonstrated moderate to substantial reliability among observers, with high statistical significance. In this series, the rating for the frontalis group showed relatively higher interobserver agreement than the ratings for the corrugator and the orbicularis oculi groups. This finding may indicate that this site is easier to categorize within the RNKLS.

---

**Table**

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No wrinkles</td>
</tr>
<tr>
<td>1</td>
<td>Wrinkles not present at rest, fine lines with facial expression</td>
</tr>
<tr>
<td>2</td>
<td>Wrinkles not present at rest, deep lines with facial expression</td>
</tr>
<tr>
<td>3</td>
<td>Fine wrinkles present at rest, deeper with facial expression</td>
</tr>
<tr>
<td>4</td>
<td>Deep wrinkles at rest, deep furrows with facial expression</td>
</tr>
</tbody>
</table>

| Figure 1. Assessment of frontalis hyperkinetic facial lines with the Rated Numeric Kinetic Line Scale (RNKLS): A, RNKLS 0; B, RNKLS 1; C, RNKLS 2; D, RNKLS 3; E, RNKLS 4. |
It should be stressed that one of the advantages of the RNKLS is to allow evaluation of patients both at rest and with expression. Since hyperkinetic facial lines by definition are a function of muscle contraction, the ability of the RNKLS to take this into account makes it a more useful instrument. By easily learning how to use the scale with a simple set of instructions and examples, the observers demonstrated its ease of use. Objective measures such as computer-assisted measurements or optical profilometry are more tedious to achieve, more expensive, and more difficult to learn.

We believe that the RNKLS is a valuable tool for the assessment of hyperkinetic facial lines. It has potential as a highly useful rating system. We used the RNKLS successfully in the study of botulinum toxin type B (MyoBloc) and in our pilot study. The former showed the RNKLS to correlate with age in patients treated with botulinum toxin type B. In short, the RNKLS is valuable to assess outcome measures, especially in prospective studies on hyperkinetic facial lines and their treatment. However, a scale used for outcome measures of treatments using neurotoxins should take into account resting and dynamic facial lines, and one is being developed.

Accepted for publication April 14, 2004.
This work was presented at the Eighth International Symposium of Facial Plastic Surgery; May 3, 2002, New York, NY.
We thank Charles E. McCulloch, PhD, for the statistical analysis.
Correspondence: Corey S. Maas, MD, The Maas Clinic, 2400 Clay St, San Francisco, CA 94115 (drmaas@drmaas.com).
REFERENCES