Reliability of the Lagrange and Rigault classification system of supracondylar humerus extension fractures in children

A. de Gheldere a,⁎, M. Legname a, M. Leyder a, G. Mezzadri b, P.-L. Docquier c

a Brabois Children’s Hospital, Nancy Teaching Medical Center, Pediatric Orthopaedic Surgery Department, rue du Morvan, 54500 Vandœuvre-lès-Nancy, France
b Mothers Children’s Hospital Lyon Teaching Medical Center, Pediatric Orthopaedic Surgery Department, 59, boulevard Pinel 69677 Bron, France
c Saint-Luc Academic Hospitals, Orthopaedic Surgery Department, 10, avenue Hippocrate, 1200 Bruxelles, Belgium

Accepted: 8 March 2010

Summary
Introduction: The Lagrange and Rigault classification was designed to describe extension-type supracondylar fractures of the humerus. It can also help in treatment decision-making.
Hypothesis: The reliability of this classification has not yet been proven. The goal of this study was to assess this system’s intra- and interobserver reliability.
Methods: One hundred supracondylar fracture radiographs were randomly retrieved and reviewed by five different observers on two occasions in a different order. The kappa index was used to calculate the intra- and interobserver reliability.
Results: Intraobserver reliability was 0.76 and interobserver reliability was 0.69.
Discussion: The study shows good intra- and interobserver reliability. The Lagrange and Rigault classification has similar reliability to other supracondylar fracture classifications.

KEYWORDS
Supracondylar fractures; Lagrange and Rigault classification; Reliability; Intra- and interobserver reliability comparison

Introduction

Extension-type supracondylar fractures are the most frequent elbow fractures in children [1,2]. The classification proposed by Lagrange and Rigault in 1962 to describe this fracture [3] remains the most widely used in France and in most French-speaking countries. It is based on the amount of displacement of the distal fragment. Originally composed
of five stages, the fifth is now infrequently used.

This classification also assists in making the decision on the type of treatment to apply depending on the stage. Orthopaedic teams vary widely in terms of indications for the various stages [4]. Stage I (undisplaced fracture) responds well to orthopaedic treatment with a long arm cast. Stage II is treated with the Blount method [5] or surgically. Stage III fractures can sometimes be treated with the Blount method, but most surgeons prefer osteosynthesis, as in stage IV. Stage V is rarer, but less stable given the more proximal location of the fracture line [6]. It requires surgical stabilization [6].

The quality of the classification system resides in its intra- and interobserver reliability. Yet the reliability of the Lagrange and Rigault classification has never been evaluated. The objective of this study was to measure the intra- and interobserver reliability of this classification so as to validate its continued use.

**Material and methods**

One hundred randomly selected radiographic files of extension-type supracondylar fractures were obtained retrospectively from two different hospitals. Each radiographic file included an AP and a lateral X-ray. The images were taken in emergency situations with the consequential disadvantages in terms of image quality.

Five different observers (two orthopaedic surgeons, one senior registrar, and two junior registrars) participated in the study. The first three observers had more than 6 years of experience. Each observer had a diagram outlining the Lagrange and Rigault classification (Fig. 1). The 100 radiographic files of the fracture were provided to each observer, who classified them according to their stage (I, II, III, IV, or V). After a minimum 15 days, the same files were submitted in a different order to the same observers, who classified them once again. Each classification was made without knowledge of the responses of the other observers or the responses given the first time by the same observer.

Intra- and interobserver reliability was calculated using SPSS.15.0 for Windows (SPSS Inc., Chicago, IL, USA). The kappa value measures the proportion of the number of times the observers gave the same response, modified to take into account random concordances. This variable measures the proportion of responses for which the observers are in agreement, taking into account the proportion related to chance [7]. The maximum kappa value (1.00) indicates perfect agreement between two observers. The minimum kappa value (0) represents agreement between two observers entirely due to chance [8]. Table 1 shows the reliability value of the kappa variable as described by Landis and Koch [8].

**Results**

The kappa values are summarized in Table 2. The mean kappa index of intraobserver reliability was 0.76. For the three most experienced observers, the proportion of X-rays that were classified identically during the two evaluations were submitted in a different order to the same observers, who classified them once again. Each classification was made without knowledge of the responses of the other observers or the responses given the first time by the same observer.

### Table 1

<table>
<thead>
<tr>
<th>Kappa</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.21</td>
<td>Very poor</td>
</tr>
<tr>
<td>0.21—0.40</td>
<td>Poor</td>
</tr>
<tr>
<td>0.41—0.60</td>
<td>Fair</td>
</tr>
<tr>
<td>0.61—0.80</td>
<td>Good</td>
</tr>
<tr>
<td>0.81—1.00</td>
<td>Very good</td>
</tr>
</tbody>
</table>

The quality of the classification system resides in its intra- and interobserver reliability. Yet the reliability of the Lagrange and Rigault classification has never been evaluated. The objective of this study was to measure the intra- and interobserver reliability of this classification so as to validate its continued use.

**Material and methods**

One hundred randomly selected radiographic files of extension-type supracondylar fractures were obtained retrospectively from two different hospitals. Each radiographic file included an AP and a lateral X-ray. The images were taken in emergency situations with the consequential disadvantages in terms of image quality.

Five different observers (two orthopaedic surgeons, one senior registrar, and two junior registrars) participated in the study. The first three observers had more than 6 years of experience. Each observer had a diagram outlining the Lagrange and Rigault classification (Fig. 1). The 100 radiographic files of the fracture were provided to each observer, who classified them according to their stage (I, II, III, IV, or V). After a minimum 15 days, the same files were submitted in a different order to the same observers, who classified them once again. Each classification was made without knowledge of the responses of the other observers or the responses given the first time by the same observer.

Intra- and interobserver reliability was calculated using SPSS.15.0 for Windows (SPSS Inc., Chicago, IL, USA). The kappa value measures the proportion of the number of times the observers gave the same response, modified to take into account random concordances. This variable measures the proportion of responses for which the observers are in agreement, taking into account the proportion related to chance [7]. The maximum kappa value (1.00) indicates perfect agreement between two observers. The minimum kappa value (0) represents agreement between two observers entirely due to chance [8]. Table 1 shows the reliability value of the kappa variable as described by Landis and Koch [8].

**Results**

The kappa values are summarized in Table 2. The mean kappa index of intraobserver reliability was 0.76. For the three most experienced observers, the proportion of X-rays that were classified identically during the two evaluations were submitted in a different order to the same observers, who classified them once again. Each classification was made without knowledge of the responses of the other observers or the responses given the first time by the same observer.

Intra- and interobserver reliability was calculated using SPSS.15.0 for Windows (SPSS Inc., Chicago, IL, USA). The kappa value measures the proportion of the number of times the observers gave the same response, modified to take into account random concordances. This variable measures the proportion of responses for which the observers are in agreement, taking into account the proportion related to chance [7]. The maximum kappa value (1.00) indicates perfect agreement between two observers. The minimum kappa value (0) represents agreement between two observers entirely due to chance [8]. Table 1 shows the reliability value of the kappa variable as described by Landis and Koch [8].

**Results**

The kappa values are summarized in Table 2. The mean kappa index of intraobserver reliability was 0.76. For the three most experienced observers, the proportion of X-rays that were classified identically during the two evaluations were submitted in a different order to the same observers, who classified them once again. Each classification was made without knowledge of the responses of the other observers or the responses given the first time by the same observer.
Table 2  Kappa values expressing intra- and interobserver reliability (a = first observation; b = second observation; bold = intraobserver reliability).

<table>
<thead>
<tr>
<th></th>
<th>Obs 1.b</th>
<th>Obs 2.b</th>
<th>Obs 3.b</th>
<th>Obs 4.b</th>
<th>Obs 5.b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs 1.a</td>
<td>0.695</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Obs 2.a</td>
<td>0.665</td>
<td>0.777</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Obs 3.a</td>
<td>0.632</td>
<td>0.686</td>
<td>0.730</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Obs 4.a</td>
<td>0.667</td>
<td>0.749</td>
<td>0.633</td>
<td>0.747</td>
<td>—</td>
</tr>
<tr>
<td>Obs 5.a</td>
<td>0.675</td>
<td>0.731</td>
<td>0.644</td>
<td>0.776</td>
<td>0.827</td>
</tr>
</tbody>
</table>

Table 3  Comparison of studies reporting intra- and interobserver reliability for different fracture classification systems.

<table>
<thead>
<tr>
<th>Study</th>
<th>Fracture site</th>
<th>Classification</th>
<th>Interobserver $\kappa$</th>
<th>Intraobserver $\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thomsen et al. [14]</td>
<td>Ankle</td>
<td>Lauge-Hansen</td>
<td>0.49 and 0.60</td>
<td>0.60–0.70</td>
</tr>
<tr>
<td>Thomsen et al. [14]</td>
<td>Ankle</td>
<td>Weber</td>
<td>0.58 and 0.56</td>
<td>0.60–0.76</td>
</tr>
<tr>
<td>Sidor et al. [15]</td>
<td>Shoulder</td>
<td>Neer</td>
<td>0.48 and 0.52</td>
<td>0.66</td>
</tr>
<tr>
<td>Siebenrock and Gerber [16]</td>
<td>Shoulder</td>
<td>Neer</td>
<td>0.40</td>
<td>0.66</td>
</tr>
<tr>
<td>Barton et al. [9]</td>
<td>Elbow</td>
<td>Gartland</td>
<td>0.74</td>
<td>0.84</td>
</tr>
<tr>
<td>Heal et al. [17]</td>
<td>Elbow</td>
<td>Modified Gartland</td>
<td>0.54</td>
<td>0.77</td>
</tr>
<tr>
<td>Our study</td>
<td>Elbow</td>
<td>Lagrange-Rigault</td>
<td>0.69</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Table 4  Rate of agreement between observers (%).

<table>
<thead>
<tr>
<th>Total agreement</th>
<th>Series 1 (%)</th>
<th>Series 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All five observers</td>
<td>51</td>
<td>54</td>
</tr>
<tr>
<td>At least four observers</td>
<td>85</td>
<td>79</td>
</tr>
<tr>
<td>At least three observers</td>
<td>100</td>
<td>99</td>
</tr>
</tbody>
</table>

was 84%. For the last two, less experienced observers, it was 82%. This difference was not significant.

The mean kappa index of interobserver reliability was 0.69. Agreement was perfect between the five observers in only 51% of the cases in the first assessment and in 54% of the cases in the second. Agreement is therefore independent of the clinician’s experience.

Discussion

The evaluations were carried out 15 days apart for each of the observers. For Barton et al. [9], this time lapse did not influence intraobserver reliability. The intraobserver reliability calculated can be qualified as "good" for four observers and very good for the fifth. The mean kappa index for intraobserver reliability (0.76) was higher than or comparable to other classification systems used in orthopaedics (Table 3).

Interobserver reliability was good between all the observers participating in the present study. The mean kappa index for interobserver reliability (0.69) was lower than the intraobserver reliability, as in most other classification systems. However, the percentage of agreement in the results of the different observers remains fairly low (Table 4).

Differentiating between stages 3 and 4 was the source of most of the disagreements, followed by differentiating between stages 1 and 2, and finally between stages 2 and 3. This may be due to the problems obtaining good-quality images when cases presented severe displacement, with pain. In these cases it is difficult to determine whether contact remains between the fragments. The absence of a difference between experienced and unexperienced observers proves that the classification is easy to assimilate.

The Lagrange and Rigault classification is therefore as reliable as the Gartland classification described in 1959 [10], which has been modified several times over the years and is sometimes a source of more confusion than precision. Initially described in three stages [10], each of the stages is now subdivided into subgroups [11,12] and a fourth stage was described for intraoperative observations [13].

It is therefore easier to use the Lagrange and Rigault classification in daily practice. However, it shows that the facility of reduction and stability of fractures cannot be assessed with this system. It should remain within its descriptive limits allowing the orthopaedic surgeon to orient treatment without dictating it.

Conflict of interest statement

None.

References

Reliability of the Lagrange and Rigault classification


