

INTRAOBSERVER AND INTEROBSERVER RELIABILITY OF THE LENKE CLASSIFICATION AMONG SPINE SURGEONS

CONFIABILIDADE INTRA OBSERVADOR E INTEROBSERVADOR DA CLASSIFICAÇÃO DE LENKE ENTRE CIRURGIÕES DE COLUNA

CONFIABILIDAD INTRA OBSERVADOR E INTEROBSERVADOR DE LA CLASIFICACIÓN DE LENKE ENTRE CIRUJANOS DE COLUMNA

GEORGE VITURINO NEVES SILVA,^{1,2} MAURÍCIO SANTOS GUSMÃO,^{1,2} RONY BRITO FERNANDES^{1,2}

1. Santa Casa de Misericórdia da Bahia, Salvador, BA, Brazil.
2. Hospital Geral do Estado, Salvador, BA, Brazil.

ABSTRACT

Objective: The objective of this study was to analyze the intraobserver and interobserver reliability of the Lenke classification among spine surgeons from the city of Salvador, Bahia. **Methods:** Preoperative imaging (front, profile and lateral inclinations) examinations of 20 patients at the Outpatient Clinic of the of Santa Izabel Hospital Orthopedic Department, Salvador, Bahia, who had been diagnosed with adolescent idiopathic scoliosis, were selected to be evaluated by 15 spine surgeons two times at an interval of 30 days, for analysis of the intraobserver and interobserver reliability of the Lenke classification. The project was first submitted for ethical analysis to the Institutional Review Board of the Santa Izabel Hospital - Santa Casa de Misericórdia da Bahia / Prof. Dr. Celso Figueirôa and approved with voucher number 002650/2019. All the participants signed the Informed Consent Form (ICF). **Results:** Analyzing the concordance using the Kappa index, interobserver reproducibilities of 0.755, 0.525 and 0.840 were obtained for the type of curve and the lumbar and sagittal modifiers, respectively, while the intraobserver reliabilities for the same parameters were 0.921, 0.370 and 0.929. **Conclusion:** For the study population, the reliability of Lenke's classification was moderate to almost perfect. **Level of evidence III; Interobserver and intraobserver reliability.**

Keywords: Scoliosis; Classification; Reproducibility of results; Observer variation.

RESUMO

Objetivo: O objetivo do presente trabalho consiste em analisar a confiabilidade intraobservador e interobservador da classificação de Lenke entre cirurgiões de coluna da cidade de Salvador, Bahia. **Métodos:** Foram selecionados exames de imagem pré-operatórios (frente, perfil e inclinações laterais) de 20 pacientes acompanhados no Ambulatório de Coluna do Departamento de Ortopedia do Hospital Santa Izabel, Salvador, Bahia, com diagnóstico de escoliose idiopática do adolescente, para serem avaliados por 15 cirurgiões de coluna, em dois momentos, com intervalo de 30 dias, para análise da confiabilidade intraobservador e interobservador da Classificação de Lenke. O projeto foi, antes de tudo, submetido a análise de ética no CEP Hospital Santa Izabel – Santa Casa de Misericórdia da Bahia/Prof. Dr. Celso Figueirôa e aprovado com número de comprovante 002650/2019. Todos os participantes assinaram o Termo de Livre Consentimento Esclarecido (TCLE). **Resultados:** Analisando-se a concordância por meio do índice Kappa, obteve-se uma reprodutibilidade interobservador de 0,755, 0,525 e 0,840, respectivamente, para o tipo de curva, modificador lombar e sagital, já a confiabilidade intraobservador é de 0,921, 0,370 e 0,929, respectivamente para o tipo de curva, modificador lombar e modificador sagital. **Conclusão:** Para a população em estudo, a confiabilidade da classificação de Lenke é de moderada a quase perfeita. **Nível de evidência III; Reprodutibilidade interobservador e intraobservador.**

Descritores: Escoliose; Classificação; Reprodutibilidade dos resultados; Variações dependentes do observador.

RESUMEN

Objetivo: El objetivo del presente trabajo consiste en analizar la confiabilidad intraobservador e interobservador de la clasificación de Lenke entre cirujanos de columna de la ciudad de Salvador, Bahia. **Métodos:** Fueron seleccionados exámenes de imagen preoperatorios (frente, perfil e inclinaciones laterales) de 20 pacientes acompañados en el Ambulatorio de Columna del Departamento de Ortopedia del Hospital Santa Izabel, Salvador, Bahia, con diagnóstico de escoliosis idiopática del adolescente, para ser evaluados por 15 cirujanos de columna, en dos momentos, con intervalo de 30 días, para análisis de la confiabilidad intraobservador e interobservador de la Clasificación de Lenke. El proyecto fue, antes que nada, sometido a análisis de ética en el CEP Hospital Santa Izabel – Santa Casa de Misericordia de Bahia/Prof. Dr. Celso Figueirôa y aprobado con número de comprobante 002650/2019. Todos los participantes firmaron el Término de Libre Consentimiento Esclarecido (TCLE). **Resultados:** Analizándose la concordancia por medio del índice Kappa, se obtuvo una reproductibilidad interobservador de 0,755, 0,525 e 0,840, respectivamente, para el tipo de curva, modificador lumbar y sagital, ya la confiabilidad intraobservador es de 0,921, 0,370 e 0,929, respectivamente para el tipo de curva, modificador lumbar y modificador sagital. **Conclusión:** Para la población en estudio, la confiabilidad de la clasificación de Lenke es de moderada a casi perfecta. **Nivel de evidencia III; Reprodutibilidad interobservador e intraobservador.**

Descriptorios: Escoliosis; Clasificación; Reproducibilidad de los resultados, Variaciones dependientes del observador.

Study conducted at the Spine Clinic of the Orthopedics Department of the Hospital Santa Izabel, Salvador, BA, Brazil.
Correspondence: Av. Luis Viana, 6312, CD Manhattan, Patamares, Salvador, BA, Brasil. 41680-400. gvmd1984@gmail.com



INTRODUCTION

Idiopathic scoliosis is the main cause of spinal deformity. It is a three-dimensional deformity of the spine in which there is deformity in the coronal plane of more than 10°, hypo- or hyperkyphosis in the sagittal plane, and rotation in the axial plane. About 80% of all cases of scoliosis are idiopathic. In adolescents the prevalence is less than 3% for curves close to 10°, which decreases to around 0.5% for curves of 20° and to between 0.1% and 0.3% for curves greater than or equal to 40°. The prevalence of smaller curves is similar between the sexes, but curves greater than 10° are more common in girls, with a 4:1 ratio.

Although it was described by Hippocrates, it was Galen who coined the term scoliosis (along with kyphosis and lordosis), and the first classification system was described in 1948 by J. R. Cobb in his article "Outline for the Study of Scoliosis". With the discovery of X-ray, the classification systems became increasingly qualitative, with a better understanding of the geometry of spinal deformity. Cobb's description of the concepts of major, minor, and structural curves continues to influence surgeons today. In 1983, H. A. King developed a classification method dividing scoliotic curves into five types (I to V), which was the first classification system to enable surgical planning. King also was the first to describe the concept of the stable vertebra, which is the vertebra best bisected by the central sacral line.

Studies conducted after King's work found that his classification system did not have good reproducibility among surgeons, did not take the sagittal plane into account, and did not consider all curve patterns, for example, leaving out double and triple major curves.^{1,2,3,4} In order to solve this problem, Lenke developed a practical two-dimensional classification system that considers the sagittal plane. The Lenke classification takes not only the thoracic curves into account, as does King's classification, but also the thoracolumbar/lumbar curves, allowing a more objective analysis of the deformity, even capable of better guiding treatment.

According to Lenke, each patient with scoliosis has three curves: proximal thoracic, main thoracic, and thoracolumbar/lumbar. The curve with the greatest Cobb angle is called the main curve and is

necessarily rigid. The two other lesser curves may be flexible or rigid, depending on the degree of correction in the lateral inclinations: curves that correct to a value less than 25° are flexible and those that correct to a value greater than 25° are rigid. The thoracic proximal, main thoracic, and thoracolumbar/lumbar curves are divided into six types of curves, types 1 to 6 (Figure 1).

In addition to the type of curve, the Lenke classification includes a lumbar modifier and a thoracic sagittal modifier. To define the lumbar modifier, pelvic obliquity must be corrected and the central vertical line of the sacrum be drawn in the coronal radiograph. If this line passes between the pedicles of the apical lumbar vertebra, it receives modifier A. If it only tangents the pedicle, it receives modifier B. If it passes laterally to the pedicle of the apical lumbar vertebra, it receives modifier C. The thoracic sagittal modifier is calculated according to the value of kyphosis between T5 and T12: the modifier is "-" if the value is less than 10°, the modifier is "N" if the value is between 10° and 40°, and the modifier is "+" if the value is greater than 40° (Figure 1).

In his study, Lenke demonstrated the excellent reliability of his classification system and reported interobserver and intraobserver Kappa values for the type of curve of 0.740 and 0.893 for an independent group of seven spine surgeons.¹ According to the same author, the estimated mean Kappa value for intraobserver reliability among spine surgeons using the King classification was 0.62 and the interobserver reliability obtained a Kappa value of 0.49.¹ Thus, it is concluded that the Lenke classification is more reproducible and more reliable than the King classification.

The objective of this study is to evaluate the intraobserver and interobserver reliability of the Lenke classification among spine surgeons in the city of Salvador.

METHODS

We selected the preoperative imaging examinations (front, profile, and lateral inclinations) of 20 patients attended at the outpatient spine clinic of the Department of Orthopedics of the Hospital Santa Izabel,

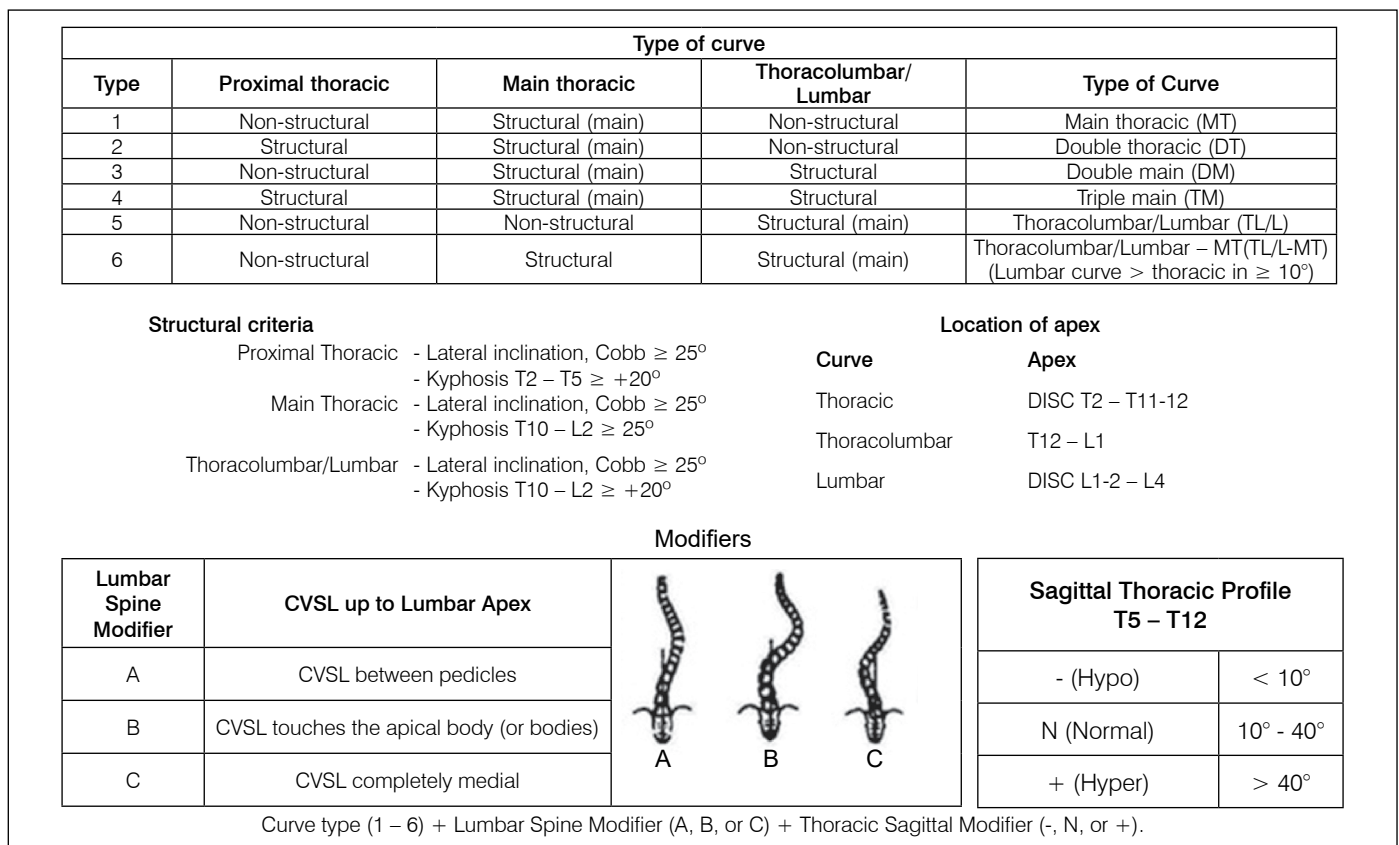


Figure 1. Lenke Classification.

Salvador, Bahia, and diagnosed with adolescent idiopathic scoliosis. The study inclusion criteria were patients indicated for surgical treatment whose radiographs were of good quality and from which the surgeons could calculate the Cobb angle and other lines of reference for classification. Fifteen spine surgeons, who practice in the city of Salvador-BA, participated in this study and evaluated the radiographs in two sessions 30 days apart. The study was conducted between 01/29/2019 and 03/13/2019 and the data were analyzed statistically using the SPSS (Statistical Package for Social Sciences), version 13.0 for Windows program. The responses were used to calculate the intraobserver and interobserver reliability using the Kappa index. Results with $p < 0.05$ were considered significant. To interpret the Kappa index, we considered values between 0.81 and 1.0 as almost perfect concordance, values between 0.61 and 0.8 as strong concordance, between 0.41 and 0.6, as moderate, between 0.21 and 0.4, as reasonable, between 0 and 0.2, as weak, and less than zero as insignificant.⁵ The project was first submitted for analysis to the Institutional Review Board of the Hospital Santa Izabel – Santa Casa de Misericórdia da Bahia/Prof. Dr. Celso Figueirôa and approved with voucher number 002650/2019. All the participants signed the Informed Consent Form (ICF).

RESULTS

A 46% response rate was achieved ($N = 7$ of 15) up until the deadline set by the investigator. Kappa values were estimated for intraobserver and interobserver concordance for each type of curve and lumbar and sagittal modifiers.

For the type of curve, the overall estimated Kappa statistic was equal to 0.75 among the observers, which can be considered to represent strong reliability among the interviewees. The estimated 95% confidence interval is given as 0.79917012 to 0.71111392. The statistical test results indicated rejection of the null hypothesis at a significance level of 0.05 ($p < 0.0001$). This means that there was concordance among the evaluators (Table 1). The intraobserver reliability for the Lenke classification curve type was 0.921, almost statistically perfect.

The estimated Kappa statistic for the lumbar modifier was equal to 0.525 for interobserver reliability. This value is considered to represent a moderate concordance among the observers (Table 2). The estimated 95% confidence interval is given as 0.59204116 to 0.45849977. The result of the statistical test indicated the rejection of the null hypothesis at a significance level of 0.05 ($p < 0.0001$), characterizing the existence of concordance among the evaluators. The mean concordance value for intraobserver variation was 0.370, or statistically reasonable.

For the thoracic sagittal modifier, the overall estimated Kappa statistic for interobserver reliability was equal to 0.840. This value is considered representative of almost perfect concordance among the observers (Table 3). The estimated 95% confidence level is given as 0.93426984 to 0.74695321. The result of the statistical test indicated the rejection of the null hypothesis at a significance level of 0.05 ($p < 0.0001$). For intraobserver reliability, the estimated Kappa value was 0.929, indicating almost perfect concordance among the observers.

DISCUSSION

The Lenke classification is currently the model with the best reproducibility and that best assists in planning the surgical treatment for

Table 1. Interobserver reliability of the Lenke classification for type of curve.

Type of curve	Kappa	p value	Degree of concordance
1	0.875274	< 0.0001	Almost perfect
2	0.642914	< 0.0001	Strong
3	0.623655	< 0.0001	Strong
4	0.638242	< 0.0001	Strong
5	0.844444	< 0.0001	Almost perfect
6	0.665798	< 0.0001	Strong

Table 2. Interobserver reliability of the Lenke classification for the lumbar modifier.

Modifier	Kappa	P value	Degree of concordance
A	0.523809	< 0.0005	Moderate
B	0.281838	0.0345	Reasonable
C	0.694677	< 0.0005	Strong

Table 3. Interobserver reliability of the Lenke classification for the sagittal modifier.

Modifier	Kappa	p value	Degree of concordance
-	-0.007194	0.9694	Insignificant
N	0.839040	0.0048	Almost perfect
+	0.871666	< 0.0001	Almost perfect

scoliosis. This is due to a better understanding of the scoliotic curve, considering it a two-dimensional deformity. King's classification only took the coronal plane and thoracic curve into account. Lenke realized that the criteria established by King were insufficient for good surgical planning, suggesting that a lack of knowledge about the lumbar and sagittal modifiers was a major cause of postoperative decompensation when the King classification was used.

In their original article, Lenke et al.¹ published interobserver and intraobserver Kappa values for the type of curve of 0.92 and 0.83, respectively, for the five developers of the classification and of 0.740 and 0.893 for a group of seven independent spine surgeons. In the independent group, the mean interobserver and intraobserver Kappa values were 0.80 and 0.84, respectively, for the lumbar modifier and 0.93 and 0.97 for the sagittal modifier. These modifiers were all in a range above 0.75, except for interobserver reliability in the independent group, whose Kappa value dropped below this level (Kappa = 0.74).

After the publication by Lenke et al.,¹ later studies demonstrated the good intraobserver and interobserver reliability of the new classification system. Ogon et al.² reported mean Kappa values for interobserver and intraobserver reliability of 0.62 and 0.73, respectively, concluding that the Lenke classification is more reliable than the older King classification. In addition, Niemeyer et al.³ concluded that the greater the experience and the level of orthopedic training, the better the interobserver and intraobserver reliability.

Thus, it is assumed that most studies indicate good reliability and reproducibility for the Lenke classification, so that spine surgeons can describe the different types of curve patterns and exchange information using the same language. The objective of this study was to analyze if the Lenke classification also had good reliability in a specific sample of spine surgeons in the city of Salvador-BA.

In the population investigated in this study, interobserver reliability values of 0.755, 0.525, and 0.840 were obtained for the type of curve, lumbar modifier, and sagittal modifier, respectively. A Kappa value between 0.525 and 0.755 is considered as a reference for moderate to strong concordance. The Kappa values obtained for intraobserver reliability were of almost perfect concordance for the type of curve, and sagittal modifier (0.921 and 0.929, respectively), while the lumbar modifier had moderate concordance (Kappa = 0.370).

In general, the mean Kappa values obtained in this study are close to the values found in other studies published to date. This suggests that for the study population the Lenke classification is a reliable and reproducible communication tool among professionals involved in the treatment of Adolescent Idiopathic Scoliosis.

CONCLUSION

For this study, it is estimated that the Lenke classification has moderate to almost perfect concordance within the investigated population.

All authors declare no potential conflict of interest related to this article.

CONTRIBUTION OF THE AUTHORS: Each author made significant individual contributions to this manuscript. GVNS: writing, revision, data collection, and statistical analysis; MSG: data analysis, writing, final approval of the manuscript to be published; RBF: data analysis, writing, final approval of the manuscript to be published.

REFERENCES

1. Lenke LG, Betz RR, Harms J, Bridwell KH, Clements DH, Lowe TG, et al. Adolescent idiopathic scoliosis: a new classification to determine extent of spinal arthrodesis. *J Bone Joint Surg Am.* 2001;83-A(8):1169-81.
2. Ogon M, Giesinger K, Behensky H, Wimmer C, Nogler M, Bach CM, et al. Interobserver and intraobserver reliability of Lenke's new scoliosis classification system. *Spine (Phila Pa 1976).* 2002;27(8):858-62.
3. Niemeyer T, Wolf A, Kluba S, Halm HF, Dietz K, Kluba T. Interobserver and intraobserver agreement of Lenke and King classifications for idiopathic scoliosis and the influence of level of professional training. *Spine (Phila Pa 1976).* 2006;31(18):2103-7.
4. Richards BS, Sucato DJ, Konigsberg DE, Ouellet JA. Comparison of reliability between the Lenke and King classification systems for adolescent idiopathic scoliosis using radiographs that were not premeasured. *Spine (Phila Pa 1976).* 2003;28(11):1148-56.
5. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-74.