REPRODUCIBILITY OF THE CLASSIFICATION OF PROXIMAL HUMERAL OSSIFICATION IN IDIOPATHIC SCOLIOSIS

REPRODUTIBILIDADE DA CLASSIFICAÇÃO DA OSSIFICAÇÃO DO ÚMERO PROXIMAL NA ESCOLIOSE IDIOPÁTICA

REPRODUCIBILIDAD DE LA CLASIFICACIÓN DE LA OSIFICACIÓN DEL HÚMERO PROXIMAL EN LA ESCOLIOSIS IDIOPÁTICA

Antonio Reinaldo de Sousa Filho[1], Rodrigo Castro de Medeiros[1], George Braga Muniz[1], Marcelo Paulo Melo de Souza[1], André Flavio Freire Pereira[1], Luciano Temporal Borges Cabral[1], Marcus André Costa Ferreira[1], Túlio Albuquerque de Moura Rangel[1]

1. Hospital Getúlio Vargas, Department of Orthopedics and Traumatology, Spine Surgery Group, Recife, PE, Brazil.

ABSTRACT

Objective: To evaluate the intra- and interobserver reproducibility of the classification of proximal humeral ossification in patients with idiopathic scoliosis. Methods: Observational, descriptive, cross-sectional study. Radiographs (anteroposterior) of 29 patients diagnosed with idiopathic scoliosis were chosen and investigated at the Spine Surgery Outpatient Clinic of the Department of Orthopedics and Traumatology of the Hospital Getúlio Vargas, Recife, Pernambuco. These radiographs were evaluated and classified by 9 spinal orthopedic surgeons at two moments, with a minimum interval of 30 days, for analysis of the intraobserver and interobserver reproducibility of the classification of proximal humeral ossification used as a predictor of peak growth velocity. Results: Intraobserver reproducibility achieved moderate to substantial concordance with the kappa index. In the analysis of interobserver reproducibility using the Fleiss' kappa method, a reasonable to moderate concordance was observed, except for in stage 5, in which concordance ranged from substantial to almost perfect. Conclusion: In this study, the reproducibility of the Li classification in most of the stages was reasonable to moderate for the interobserver assessment and moderate to substantial for the intraobserver assessment, which differs from the current literature. ***Level of evidence III; Diagnostic studies – Investigation of a diagnostic test.***

Keywords: Scoliosis; Growth; Reproducibility of Results.

RESUMO

*Objetivo: Avaliar a reprodutibilidade intra e interobservadores da classificação de ossificação do úmero proximal em pacientes portadores de escoliose idiopática. Métodos: Estudo observacional, descritivo e transversal. Foram selecionadas radiografias (anteroposteriores) de 29 pacientes com diagnóstico de escoliose idiopática acompanhados no Ambulatório de Cirurgia da Coluna do Departamento de Ortopedia do Hospital Getúlio Vargas, Recife, Pernambuco. Essas radiografias foram avaliadas e classificadas por nove cirurgiões ortopedistas de coluna em dois momentos, com intervalo mínimo de 30 dias, para análise da reprodutibilidade intraobservador e interobservador da classificação de ossificação do úmero proximal utilizada como preditor do pico de velocidade de crescimento. Resultados: A reprodutibilidade intraobservador com o índice Kappa atingiu concordância moderada a substancial. Já na análise da reprodutibilidade interobservador com o método de kappa de Fleiss constatou-se concordância de razoável a moderada, exceto no estágio 5, no qual a concordância variou de substancial a quase perfeita. Conclusões: Neste estudo, a reprodutibilidade da classificação de Li, na maioria dos estágios foi razoável a moderada na avaliação interobservador e moderada à substancial na avaliação intraobservador, o que diverge da literatura atual. **Nível de evidência III; Estudos diagnósticos – Investigação de um exame para diagnóstico.***

Descritores: Escoliose; Crescimento; Reprodutibilidade dos Testes.

RESUMEN

*Objetivo: Evaluar la reproducibilidad intra e interobservador de la clasificación de osificación del húmero proximal en pacientes con escoliosis idiopática. Métodos: Estudio observacional, descriptivo y transversal. Se seleccionaron radiografías (anteroposteriores) de 29 pacientes con diagnóstico de escoliosis idiopática en la Clínica de Cuidados Ambulatorios de Cirugía de Columna del Departamento de Ortopedia del Hospital Getúlio Vargas, Recife, Pernambuco. Estas radiografías fueron evaluadas y clasificadas por nueve cirujanos ortopedistas de columna, en dos momentos, con un intervalo mínimo de 30 días, para analizar la reproducibilidad intraobservador e interobservador de la clasificación de osificación del húmero proximal utilizada como predictor de la velocidad máxima de crecimiento. Resultados: La reproducibilidad intraobservador con el índice Kappa alcanzó concordancia de moderada a sustancial de la reproducibilidad intraobservador. En el análisis de la reproducibilidad interobservador con el método de kappa de Fleiss, se encontró concordancia de razonable a moderada, excepto en la etapa 5, en que la concordancia varió de sustancial a casi perfecta. Conclusiones: En este estudio, la reproducibilidad de la clasificación de Li en la mayoría de las etapas fue de razonable a moderada en la evaluación interobservador y de moderada a sustancial en la evaluación intraobservador, lo que difiere de la literatura actual. **Nivel de Evidencia III; Estudios de Diagnóstico – Investigación de una prueba de diagnóstico.***

Descriptores: Escoliosis; Crecimiento; Reproducibilidad de los Resultados.

INTRODUCTION

Idiopathic scoliosis (IS) accounts for about 80% of all scolioses and is the main cause of spinal deformity in pediatric patients. Its diagnosis requires a deviation of greater than 10º in the coronal plane, measured by the Cobb technique. The etiology remains undefined and is probably multifactorial.[1,2]

Risk factors for the progression of the deformity are well defined and include sex, magnitude and location of the curve, and skeletal maturity. During the peak of the velocity of the somatic growth curve, patients are at greater risk for progression of the curve. Thus, the indication of different treatment modalities (observation, orthosis or surgery) requires an adequate assessment of skeletal maturity in order to predict the remaining growth.[3,4]

The Scoliosis Research Society[5] recommends that skeletally immature patients, with curves between 25° and 40°, be treated using a brace in order to prevent deformity progression. Surgical treatment is chosen for skeletally immature patients with curves greater than 45° or skeletally mature patients with curves $\geq$ 50°, since these two groups present a prognosis of curve progression in adulthood.

The accuracy of the risk of scoliotic deformity progression as determined through an evaluation of skeletal maturity has been the subject of several studies in the literature. Among them, the following can be highlighted: a) the Risser method, which evaluates the progression of the apophyseal ossification of the iliac crest;[6] b) the classification of ossifications of the distal radius and ulna described by Luk et al.;[7] c) the Greulich-Pyle atlas of ossification of the hands and wrists that, while it is more accurate, is difficult to apply in clinical practice;[8] and d) ossification of the ischial crest, a method that in the study by Landin et al.,[9] was shown to correspond to the Risser sign.

In 2008, Sanders and other authors described a bone age evaluation method, based on the Tanner-Whitehouse and Greulich-Pyle methods, in which radiographs of the hand and wrist are evaluated.[10,11] Although this classification is considered the gold standard in terms of the accuracy of forecasting remaining growth,[12] its practical applicability is also difficult. Thus, the Risser method remains the most commonly used to date for assessing skeletal maturity.[13]

In 2018, Li et al.,[3] proposed a new way to evaluate skeletal maturity that is alleged to be able to predict peak growth velocity. These authors developed a radiographic classification centered around the evaluation of 2 distinct parameters evidenced during the evolution of the ossification of the proximal humerus: a) the morphology of the lateral edge of the proximal epiphysis of the humerus and its geometric correlation with the lateral metaphyseal cortex of the humerus, and b) the evolution of the closure of the proximal humeral physis.

The latter classification mentioned above consists of 5 evolutionary stages (Figure 1). In stage 1, the ossification of the lateral epiphysis is incomplete, such that the lateral margin is straightened and oriented in an oblique position relative to the lateral metaphyseal cortex, which allows it to appear as a radiolucent triangular area. In stage 2, the lateral margin of the proximal epiphysis of the humerus is rounded but not yet touching a line that represents the proximal projection of the lateral metaphyseal cortex. In stages 3 to 5, the lateral edge of the proximal epiphysis of the humerus touches the line that represents the proximal projection of the lateral metaphyseal cortex, and the three stages are distinguished by the closure of the proximal humeral physis. In stage 3, the lateral half of the physis is totally open, with no obvious fusion. In stage 4, the lateral half of the physis is partially fused. In stage 5, there is complete fusion of the entire proximal physis of the humerus.

Also according to Li et al.,[12] the transition between stages 2 and 3 precedes the second peak of the somatic growth curve by up to one year in 72% of the cases analyzed. Finally, these same authors note that there is less overlap between the stages of humeral ossification than in the Risser system and that, therefore, their classification is easy to understand and simpler to use, and it does not require additional radiographs in patients with scoliosis.

In light of the advantages described above for the outpatient applicability of the classification of the ossification of the humeral head as a predictor of peak growth velocity, an assessment of its intra- and interobserver reproducibility is essential.

Thus, the objective of this study is to evaluate the intra- and interobserver reproducibility of the classification of proximal humeral ossification in patients with idiopathic scoliosis.

METHODS

This was an observational, descriptive, cross-sectional study in which preoperative, panoramic, anteroposterior, spinal radiographs of patients with idiopathic scoliosis being treated at the Hospital Getúlio Vargas in Recife/PE during the period comprising 2018 and 2019 were evaluated. The radiographic images were collected from the medical files of the spine surgery group at the hospital.

Initially, 186 radiographs of different patients were catalogued. The inclusion criterion used during the radiograph selection process was the possibility of adequately and simultaneously analyzing the scoliotic spinal deformities and the proximal portion of the left humerus. The choice of laterality was in accordance with that used in the work of the author of the classification in order to standardize the analysis. Radiographs that did not show non-idiopathic scoliosis or in which the radiographic technique was considered to have inadequate visibility to evaluate the degree of proximal humeral ossification were excluded from the study. Epidemiological data related to age and sex were also collected.

The selection process permitted the inclusion of 29 radiographs, in which it was possible to observe all the different phases of the spectrum of classification of the ossification of the proximal humerus proposed by Li et al.[3] All the images were compiled in PowerPoint to be presented at two different moments, with a minimum interval of 30 days between them, to a group of 10 spine surgeons.

All the participants are orthopedists and have a broad range of years of experience in the specialty. All the participants were previously introduced to the classification system of Li et al.[3] Then, standardized cards were provided for marking the objective responses (selection of only one stage for the classification of each of 29 sequential cases). The response time was left free according to individual needs. The evaluation was performed separately by each of the participants. Finally, between the first and second presentations, the cases were reordered blindly and only the author knew the correlation of the presentation of the images between the two moments.

Radiographic standardization for analysis of the left shoulder, adequate radiographic quality, providing a standardized card for objective responses, presenting the images to the participants separately with no time limit, and the interval between the two evaluations were meant to avoid mitigation of possible bias and enable better evaluation of the intra- and interobserver reproducibility for the classification being analyzed.

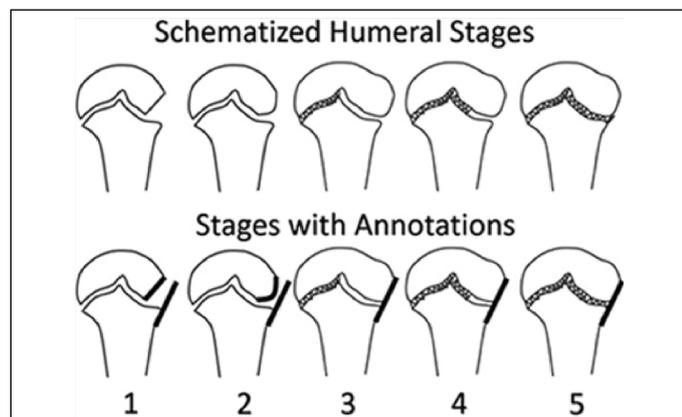The collected data were compiled in Excel tables and analyzed

Figure 1. Classification of the ossification of the proximal humerus. The sequence from 1 to 5 represents the milestones in the progression of this phenomenon. Reproduced from Li et al.[3]

using the Statistical Package for Social Sciences (SPSS) program. The results of categorical variables were presented in a table with their respective absolute and relative frequencies. To verify intra- and interobserver concordance, the data were submitted to statistical treatment using the kappa (intraobservers) and Fleiss' kappa (interobservers) tests.[14] For interpretation of the kappa index, values between 0.81 and 1.0 were considered to show almost perfect concordance; values between 0.61 and 0.8, substantial concordance; between 0.41 and 0.6, moderate concordance; between 0.21 and 0.4, fair concordance, between 0 and 0.2, weak concordance; and less than zero, insignificant concordance. The interpretation of the results was based on the analysis of Landis and Koch.[14]

This study was submitted to and approved by the Institutional Review Board of the Hospital Agamenon Magalhães - Recife/PE (CAAE: 31450820.8.0000.5197). All the participants signed the Informed Consent Form (ICF).

RESULTS

Of the 29 radiographs analyzed, 21 (72.4%) belonged to female patients. The mean age of the cases was 13.6 $\pm$ 3.9 years (Table 1).

Only one of the 10 orthopedic surgeons invited to participate in the study did not respond to the invitation within the timeframe stipulated by the researcher. Of the remaining nine, three (33.3%) had between 1 and 5 years of spinal surgery experience, while 6 (66.7%) had more than 10 years of experience.

In the intraobserver analysis (Table 2), five observers obtained substantial concordance while 2 others obtained only moderate concordance. It was not possible to calculate the kappa from the results of the other two observers because there was no symmetry between the responses from the two evaluation moments. The overall interobserver concordance was moderate both in the first assessment, k=0.564, and in the second assessment, k=0.489 (p < 0.001). In the analysis of the two moments together, moderate reproducibility, k = 0.531 (p < 0.001), was also obtained (Table 3).

Stage 5 of Li et al.[3] was the only one to present almost perfect interobserver concordance in the first assessment and the only one to present substantial concordance in the second moment and in the overall assessment. In stages 2 and 3 there was fair concordance in all the assessment scenarios (p < 0.001) (Table 3).

DISCUSSION

Idiopathic scoliosis affects around 2 to 3% of the population under 16 years of age. There is a higher incidence of females among patients in whom the onset occurs after seven years of age.

Table 1. Epidemiological data on the sex and age of patients with idiopathic scoliosis who underwent radiographic assessment of the shoulder.

Variables	n	%
Sex		
Male	8	27.6
Female	21	72.4

Table 2. Intraobserver assessment between the two moments of application of the Li et al.[3] classification.

Comparison	Kappa	p-value	Degree of Con-cordance
A1 x A2	0,551	<0,001	Moderate
B1 x B2	0,776	<0,001	Substantial
C1 x C2	*	*	*
D1 x D2	0,767	<0,001	Substantial
E1 x E2	0,688	<0,001	Substantial
F1 x F2	0,723	<0,001	Substantial
G1 x G2	*	*	*
H1 x H2	0,485	<0,001	Moderate
I1 X I2	0,774	<0,001	Substantial

*Not possible to calculate.

Table 3. Interobserver concordance stratified by the Li et al.[3] stages for the 1st and 2nd mo-ments, presented separately by stage and as overall concordance, evaluated using the Fleiss' kappa method[14]

Comparison	Kappa	p-value	Degree of Con-cordance
Assessment 1			
Overall	0,564	< 0,001	Moderate
Stage			
1	0,513	< 0,001	Moderate
2	0,390	< 0,001	Fair
3	0,259	< 0,001	Fair
4	0,659	< 0,001	Substantial
5	0,815	< 0,001	Almost perfect
Assessment 1			
Overall	0,489	< 0,001	Moderate
Stage			
1	0,489	< 0,001	Moderate
2	0,334	< 0,001	Fair
3	0,384	< 0,001	Fair
4	0,439	< 0,001	Moderate
5	0,702	< 0,001	Substantial
Assessment 1 and 2 together			
Overall	0,531	< 0,001	Moderate
Stage			
1	0,512	< 0,001	Moderate
2	0,367	< 0,001	Fair
3	0,342	< 0,001	Fair
4	0,551	< 0,001	Moderate
5	0,758	< 0,001	Substantial

In the present study the mean age of the patients evaluated was 13.60 $\pm$ 3.99 years, 72.4% of them being female, which agrees with the literature.[1,15,16]

In their original 2018 article, Li et al.,[3] analyzed anteroposterior radiographs of the proximal humerus of patients with their upper limbs at the sides of their body and their hands turned towards the front (anatomical position). The published intraobserver and interobserver concordance results were substantial, with the kappa test value of 0.80 and 0.78, respectively. The calculation of the intraclass correlation coefficient was excellent (ICC $\geq$ 0.75).

In a second study conducted by the same group in 2019, the reliability of the Li et al.,[3] classification was evaluated using the intraclass correlation coefficient. In this second study, the evaluation of the humerus did not follow the radiographic protocol of the first, in which the position of the shoulder/arm was not controlled. Despite this absence of a radiographic standard, no difference was observed in the reliability of the classification. The intraclass coefficient was 0.97 for the intraobserver analysis and 0.92 for the interobserver, remaining with an interpretation of excellent reproducibility.[12]

In the present study, the presentation of the cases to a group of surgeons with heterogeneous experience is, in theory, a factor that could reinforce the daily applicability of the Li et al.[3] classification since it would have the ability to reach everyone. However, the results obtained were different from those published by the group that created this classification, both for the interobserver and intraobserver assessments.

Since the intraobserver concordance results obtained were moderate and substantial, in contrast to the interobserver concordance which, in most cases, ranged from fair to substantial, two analyses are needed: a) within the distribution of the cases presented, the number of patients classified as "stage 5" must have been statistically greater than the others and, thus, have skewed the intraobserver kappa result towards a more robust reliability. In addition, in stage 5 ossification is already complete, making it easier to identify it; b) if, on the one hand, low interobserver concordance in most of the scenarios may reflect inadequate training of the participants

for the application of the classification, on the other hand, it may suggest a divergence in the understanding of the concept of the closing of the growth plate among the participants.

Although the totally open physis progresses towards calcification through vascular invasion in the metaphyseal pole (temporary calcification layer), its "closing/maturation" begins with the deposit of calcium in the hypertrophic zone (longitudinal growth zone), a fact that is visible in the radiograph as "radiodense specks" dispersed in the middle of the radiolucent area of the physis that can lead to the divergence in classification between an open and a non-open physis.[17-19]

Is should also be noted that, in a totally closed physis, a bone bridge has already been created that produces a homogeneous image of trabecular bone that prevents the difficulty of definition. This fact was evidenced in our study by good interobserver concordance only in stage 5 (physis totally closed) and the low concordance exactly in the definition of stages 2 and 3 (moments of the onset of calcification of the physis).[17-19]

Although each specific physis presents a distinct closing pattern, it seems to us that the Risser classification[20,21] is more accepted because it evaluates an apophysis, which would make it easier to show the presence of partial calcification (Risser 1 to 4) and complete closure (Risser 5).

Thus, it may be that adequate training in the application of the Li et al. classification[3] requires better guidance for the physicians in terms of the definition and identification of physis closure.

CONCLUSION

In this study, the reproducibility of the Li et al., classification[3] was, in most of the stages, from fair to moderate in the interobserver assessment and from moderate to substantial in the intraobserver assessment, which diverges from the current literature. The search for a more precise, easy to learn and reproducible method for assessing bone age is important, as it is the basis for the treatment of patients with scoliosis. Therefore, more studies are needed.

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

REFERENCES

1. Lenke LG, Silva FE, Lehman RA Jr. Idiopathic scoliosis. In: Garfin SR, Eismont FJ, Bell GR, Bono CM, Fischgrund J. Rothman-Simeone and Herkowitzs the spine. Philadelphia: Saunders/Elsevier; 2011. p. 451-68.
2. Penha PJ, Ramos NLJP, Carvalho BKG, Andrade RM, Schmitt ACB, João SMA. Prevalence of adolescent idiopathic scoliosis in the state of Sao Paulo, Brazil. Spine (Phila Pa 1976). 2018;43(24):1710-8. doi: 10.1097/BRS.0000000000002725
3. Li D, Cui JJ, DeVries S, Nicholson AD, Li E, Petit L, et al. Humeral head ossification predicts peak height velocity timing and percentage of growth remaining in children. J Pediatr Orthop. 2018;38(9):e546-50. doi: 10.1097/BPO.0000000000001232.
4. Neal K, Shirley E, Kiebzak G. Maturity indicators and adolescent idiopathic scoliosis. Spine (Phila Pa 1976). 2018;43(7):E406-12. doi: 10.1097/BRS.0000000000002483.
5. Scoliosis Research Society. Dúvidas Comuns sobre Escoliose. 2020 [acesso em: 2020 fev 2020]. Disponível em: https://www.srs.org/portuguese/patient_and_family/scoliosis/.
6. Manring MM, Calhoun J. Joseph C. Risser Sr., 1892–1982. Clin Orthop Relat Res. 2010;468(3):643-5. doi: 10.1007/s11999-009-1095-0
7. Luk KD, Saw LB, Grozman S, Cheung KM, Samartzis D. Assessment of skeletal maturity in scoliosis patients to determine clinical management: a new classification scheme using distal radius and ulna radiographs. Spine J. 2014;14(2):315-25. doi: 10.1016/j.spinee.2013.10.045.
8. Cheung JP, Samartzis D, Cheung PW, Cheung KM, Luk KD. Reliability analysis of the distal radius and ulna classification for assessing skeletal maturity for patients with adolescent idiopathic scoliosis. Global Spine J. 2016;6(2):164-8. doi: 10.1055/s-0035-1557142.
9. Landim E, Cavali PTM, Pasqualini W, Gracitelli LC. Classificação da ossificação da crista do ísquio na maturidade esquelética e sua correlação com o sinal de Risser. Rev Bras Ortop. 1998;33(2):109-10.
10. Dimeglio A, Canavese F. Progression or not progression? How to deal with adolescent idiopathic scoliosis during puberty. J Child Orthop. 2013;7(1):43-9. doi: 10.1007/s11832-012-0463-6

11. Sitoula P, Verma K, Holmes L Jr, Gabos PG, Sanders JO, Yorgova P, et al. Prediction of curve progression in idiopathic scoliosis: validation of the sanders skeletal maturity staging system. Spine (Phila Pa 1976). 2015;40(13):1006-13. doi: 10.1097/BRS.0000000000000952.
12. LI D, Linderman GC, Cui JJ, DeVries S, Nicholson AD, Li E, et al. The proximal humeral ossification system improves assessment of maturity in patients with scoliosis. J Bone Joint Surg Am. 2019;101(20):1868-74. doi: 10.2106/JBJS.19.00296.
13. Yang JH, Bhandarkar AW, Suh SW, Hong JY, Hwang JH, Ham CH. Evaluation of accuracy of plain radiography in determining the Risser stage and identification of common sources of errors. J Orthop Surg Res. 2014;9:101. doi: 10.1186/s13018-014-0101-8.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74
15. Freeman III BL. Escoliose e cifose. In: Canale ST (ed.). Cirurgia ortopédica de Campbell. 10ª ed. Barueri: Manole; 2006. p. 1751-954.
16. Hawary RE, Zaaroor-Regev D, Floman Y, Lonner BS, Alkhalife YI, Betz RR. Brace treatment in adolescent idiopathic scoliosis: risk factors for failure-a literature review. Spine J. 2019;19(12):1917-25. doi: 10.1016/j.spinee.2019.07.008.
17. Diedtz FR, Morcuend JA. Embriologia e desenvolvimento do sistema muscoloesqueletico. In: Morrissy RT, Weinstein SL. Ortopedia pediátrica de Lovell e Winter. Baureri: Manole; 2005. p. 1-35.
18. Diméglio A. Crescimento em ortopedia pediátrica. In: Morrissy RT, Weinstein SL. Ortopedia pediátrica de Lovell e Winter. Baureri: Manole; 2005. p. 36-67.
19. Zember JS, Rosenberg ZS, Kwong S, Kothary SP, Bedoya MA. Normal skeletal maturation and imaging pitfalls in the pediatric shoulder. Radiographics. 2015;35(4):1108-22. doi: 10.1148/rg.2015140254.
20. Vira S, Husain Q, Jalai C, Paul J, Poorman GW, Poorman C, et al. The interobserver and intraobserver reliability of the Sanders Classification versus the Risser Stage. J Pediatr Orthop. 2017;37(4):e246-9. doi: 10.1097/BPO.0000000000000891.
21. Newton PO, Wenger DR. Escolioses idiopática e congênita. In: Morrissy RT, Weinstein SL. Ortopedia pediátrica de Lovell e Winter. Baureri: Manole; 2005. p. 737-806.