Article published online: 2021-12-13

Avaliação da confiabilidade e reprodutibilidade da classificação de Roussouly para os tipos de lordose lombar

Evaluation of the Reliability and Reproducibility of the Roussouly Classification for Lumbar Lordosis Types

Camila Oda Yamazato[1] Gustavo Ribeiro[2] Fabio Chaud de Paula[3] Ramon Oliveira Soares[3] Paulo Santa Cruz[4] Michel Kanas[4]

[1] Departamento de Ortopedia e Traumatologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brasil
[2] Hospital Israelita Albert Einstein, São Paulo, SP, Brasil
[3] Complexo Hospitalar Universitário Professor Edgar Santos, São Paulo, SP, Brasil
[4] Ambulatório de Coluna do Centro de Traumatologia do Esporte, Departamento de Ortopedia e Traumatologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brasil

Endereço para correspondência Camila Oda Yamazato, Departamento de Ortopedia e Traumatologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, Rua Botucatu, 740, Térreo, São Paulo, SP, 04023-900, Brasil (e-mail: camila.yamazato@gmail.com).

Rev Bras Ortop 2022;57(2):321–326.

Resumo

Objetivo Determinar a confiabilidade e reprodutibilidade intra- e interavaliadores da classificação dos tipos de lordose lombar de Roussouly.

Métodos Foram utilizadas 104 radiografias panorâmicas da coluna vertebral em incidência de perfil, de banco de dados, de indivíduos do sexo masculino com idade entre 18 e 40 anos. Utilizando o software Surgimap (Nemaris Inc., Nova York, NY, EUA), seis examinadores com diferentes níveis de experiência aferiram os ângulos espino-pélvicos e classificaram o tipo de lordose de acordo com a classificação de Roussouly. Após um intervalo de 1 mês, as mensurações foram realizadas novamente, sendo calculadas as concordâncias intra- e interavaliadores através do teste Kappa de Fleiss.

Resultados O estudo demonstrou evidências positivas em relação à reprodutibilidade da classificação de Roussouly, com teste Kappa de Fleiss para concordância intra-avaliador entre razoável à quase perfeita (0,307–0,827), e interavaliador entre moderada (0,43) e razoável (0,369). Os examinadores mais experientes apresentaram maior concordância interavaliador, variando entre substancial (0,619) e moderada (0,439).

Conclusão A classificação de Roussouly, demonstrou boa confiabilidade e reprodutibilidade, tendo em vista que a concordância intra- e interavaliadores foi considerada no mínimo razoável, podendo atingir os níveis de substanciais à quase perfeitos em

Palavras-chave
- curvaturas da coluna vertebral
- equilíbrio postural
- lordose/classificação

recebido
28 de Setembro de 2020
aceito
01 de Dezembro de 2020
Publicado on-line
Dezembro 13, 2021

DOI https://doi.org/10.1055/s-0041-1729581.
ISSN 0102-3616.

© 2021. Sociedade Brasileira de Ortopedia e Traumatologia. All rights reserved.
This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Thieme Revinter Publicações Ltda., Rua do Matoso 170, Rio de Janeiro, RJ, CEP 20270-135, Brazil

algumas situações. Os avaliadores com maior nível de experiência apresentaram maior concordância intra- e interavaliadores.

Abstract

Objective The present study aims to determine the intra- and inter-rater reliability and reproducibility of the Roussouly classification for lumbar lordosis types.

Methods A database of 104 panoramic, lateral radiographs of the spine of male individuals aged between 18 and 40 years old was used. Six examiners with different expertise levels measured spinopelvic angles and classified lordosis types according to the Roussouly classification using the Surgimap software (Nemaris Inc., New York, NY, USA). After a 1-month interval, the measurements were repeated, and the intra- and inter-rater agreement were calculated using the Fleiss Kappa test.

Results The study revealed positive evidence regarding the reproducibility of the Roussouly classification, with reasonable to virtually perfect (0.307–0.827) intra-rater agreement, and moderate (0.43) to reasonable (0.369) inter-rater agreement according to the Fleiss kappa test. The most experienced examiners showed greater inter-rater agreement, ranging from substantial (0.619) to moderate (0.439).

Conclusion The Roussouly classification demonstrated good reliability and reproducibility, with intra- and inter-rater agreements at least reasonable, and reaching substantial to virtually perfect levels in some situations. Evaluators with highest expertise levels showed greater intra and inter-rater agreement.

Keywords
- spinal curvatures
- postural balance
- lordosis/classification

Introdução

Nas últimas 2 décadas, o estudo dos ângulos espinopélvicos e do equilíbrio sagital vem sendo cada vez mais relevante na cirurgia da coluna, principalmente para correção de deformidades do adulto. Foi demonstrado que essas medidas podem variar de acordo com idade, etnia e biotipo em pacientes assintomáticos; assim como em pacientes sintomáticos, de acordo com a etiologia do desequilíbrio sagital.[1–3]

As deformidades da coluna no adulto destoam clínica e radiologicamente da do adolescente, devido à associação a processos degenerativos; apresentando diferentes padrões e história natural de progressão. Ademais, o desequilíbrio da coluna no plano sagital também pode ser sequela de fraturas ou complicação pós-operatória.[4]

Foi encontrada uma correlação inversamente proporcional entre o grau de lordose lombar e dor; além da associação entre parâmetros espinopélvicos e tipo de lordose lombar, com desenvolvimento de degeneração discal, sobrecarga facetaria, espondilolistese, dor lombar crônica, hérnia discal e incapacidade funcional.[5–10]

Para compensar o desequilíbrio sagital, o corpo lança mão de mecanismos como: aumento da cifose/lordose de segmentos adjacentes, hiperextensão do tronco, anteversão ou retroversão da pelve, flexão dos joelhos e extensão dos tornozelos.[11,12] Esses mecanismos, assim como parâmetros anatômicos da coluna e o alinhamento sagital do indivíduo, devem ser levados em conta na indicação e no planejamento de cirurgias, por influenciarem no prognóstico pós-operatório.[13]

Tendo em vista a importância dessa avaliação, Roussouly et al.[14] propuseram uma classificação que divide a coluna lombar em 4 tipos, de acordo com o ápice da lordose e o ângulo de inclinação sacral. Esta classificação vem sendo utilizada tanto em pesquisas quanto na prática clínica, desde sua criação. No entanto, faltam trabalhos que comprovem sua validação.

O objetivo do estudo foi avaliar a confiabilidade e reprodutibilidade da classificação dos tipos de lordose lombar de Roussouly et al.,[14] e verificar se há influência do nível de experiência na concordância interavaliadores.

Materiais e Métodos

Foram utilizadas 104 radiografias panorâmicas da coluna vertebral em incidência de perfil de indivíduos do sexo masculino com idade entre 18 e 40 anos.

As radiografias pertenciam a um banco de dados, e já haviam sido utilizadas em estudos prévios para outras avaliações. Devido à impossibilidade de contato com esses indivíduos e a anonimização prévia dos mesmos, o estudo obteve dispensa do termo de consentimento livre esclarecido, sendo aprovado pelo Comitê de Ética em Pesquisa, conforme protocolo nº 3.828.093.

Todas as radiografias foram realizadas no mesmo aparelho e, durante o exame, era solicitado ao paciente que permanecesse em pé, com o tronco ereto, membros superiores apoiados em suporte, ombros com 30° de flexão, cotovelos levemente fletidos e os joelhos estendidos. As radiografias panorâmicas de perfil deveriam abranger desde a base do crânio até a região proximal do fêmur, sendo excluídas radiografias de má qualidade que impossibilitavam as mensurações.

 © 2021. Sociedade Brasileira de Ortopedia e Traumatologia. All rights reserved.

Fig. 1 Exemplo de aferição dos ângulos espinopélvicos e eixo vertical sagital (EVS), com o uso do software Surgimap.

Por meio do software Surgimap versão 2.3.1.5 (Nemaris Inc., Nova York, NY, EUA), 6 avaliadores, sendo 2 cirurgiões de coluna com > 5 anos de experiência (A1 e A2), 2 residentes em Cirurgia da Coluna (B1 e B2), e 2 residentes de Ortopedia e Traumatologia (C1 e C2), aferiram os ângulos espinopélvicos e o eixo vertical sagital (►**Figura 1**). Utilizando estes dados, o tipo de lordose lombar foi classificado de acordo com Roussouly et al.[15] (►**Figura 2**). Após um intervalo de 1 mês, as mensurações foram realizadas novamente, pelos mesmos avaliadores, para avaliação da concordância intra- e interavaliador.

Métodos estatísticos

Inicialmente, foi realizada uma análise descritiva dos resultados para a obtenção de gráficos e tabelas de frequência, com o intuito de caracterizar os participantes da pesquisa. Para a descrição dos resultados, foram utilizadas frequência absoluta e porcentagem para as variáveis categóricas. Foram construídos gráficos para avaliar a frequência das variáveis de interesse.

Para verificar a concordância intra- e interavaliador da Classificação de Roussouly, foi utilizado o teste Kappa de Fleiss (1981), uma generalização do teste Kappa, apropriado para situações nas quais diversos avaliadores avaliam a mesma amostra em uma escala com várias categorias, como é o caso da Classificação de Roussouly, que pode variar de tipo 1 a 4. O coeficiente de concordância Kappa possui valores que variam de + 1 (concordância perfeita) a 0 (concordância igual a esperada ao acaso) a - 1 (discordância completa).[16]

De acordo com Landis e Koch[17], o coeficiente Kappa foi classificado como demonstrado na ►**Tabela 1**.

Todas as análises foram realizadas com o auxílio do ambiente estatístico R versão 3.3.1 (R Foundation, Viena, Áustria), e o nível de significância foi fixado em 5%.[18]

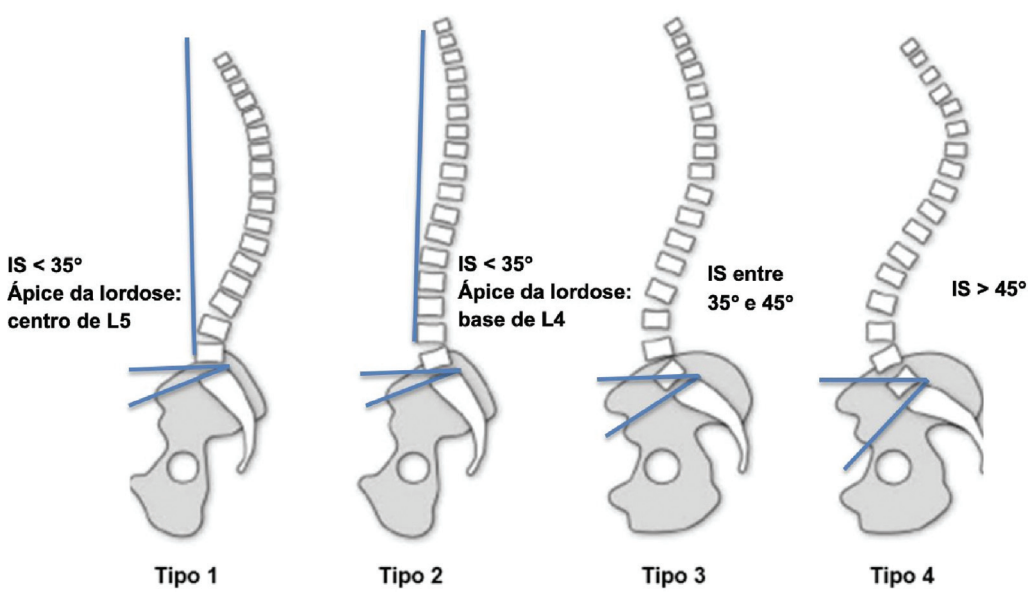

Fig. 2 Classificação dos tipos de lordose lombar de acordo com Roussouly et al. **Tipo 1**: A inclinação sacral (IS) é < 35°, o ápice da lordose lombar está localizado no centro do corpo vertebral L5. **Tipo 2**: A IS é < 35°, o ápice da lordose lombar está localizado na base do corpo vertebral L4. **Tipo 3: A** IS é entre 35° e 45°. **Tipo 4:** A IS é > 45°. (Adaptado de Roussouly et al.[15])

 © 2021. Sociedade Brasileira de Ortopedia e Traumatologia. All rights reserved.

Tabela 1 Classificação do coeficiente Kappa de acordo com Landis et al.[17]

Kappa	Força de concordância
< 0,00	Pobre
0,00–0,20	Fraca
0,21–0,40	Razoável
0,41–0,60	Moderada
0,61–0,80	Substancial
0,81–1,00	Quase perfeita

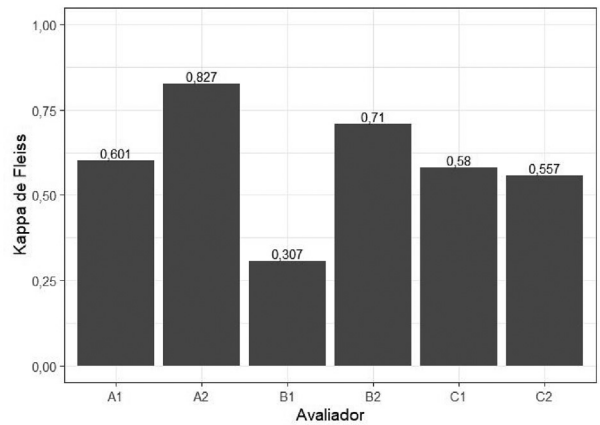

Fig. 3 Coeficiente Kappa de Fleiss para a concordância intra-avaliador.

Resultados

A ►**Tabela 2** apresenta a distribuição de frequências do tipo de lordose, atribuída por cada avaliador, nas duas aferições. Nota-se que o tipo 3 foi o mais frequentemente atribuído (> 50%), para todos os avaliadores, com exceção do B1, que classificou a maioria dos casos como tipo 4.

Concordância intra-avaliador

Ao realizar análise da concordância intra-avaliador, foi observado que o avaliador A2 apresentou o melhor nível de concordância, com coeficiente quase perfeito (0,827). O menor nível de concordância nessa análise foi o do avaliador B1, com coeficiente razoável (0,307). Os avaliadores A1 e B2 apresentaram concordâncias substanciais (0,601 e 0,710, respectivamente), ao passo que os avaliadores C1 e C2 apresentaram concordâncias moderadas (0,580 e 0,557). A média da concordância intra-avaliador foi de 0,597, considerada moderada. Todos os valores encontrados apresentaram valor de $p < 0{,}001$ (►**Figura 3** e ►**Tabela 3**).

Concordância interavaliador

Ao realizar a análise da concordância interavaliador, o coeficiente geral obtido na primeira aferição foi de 0,43 (moderada), enquanto na segunda aferição este foi ligeiramente menor, de 0,369 (razoável) (►**Tabela 4**).

Dentre avaliadores com o mesmo nível de experiência, nota-se que houve concordância estatisticamente significativa ($p < 0{,}001$) entre todos os grupos, sendo que os cirurgiões de coluna com > 5 anos de experiência apresentaram o mais alto nível de concordância interavaliador, variando entre substancial na primeira aferição (0,619) e moderada na segunda (0,439). Os residentes em cirurgia da coluna e os residentes em ortopedia e traumatologia apresentaram níveis razoáveis de concordância interavaliador dentro de suas classes (►**Tabela 5**).

Discussão

Uma classificação adequada, além de ditar o tratamento ou fornecer algum prognóstico, também deve ser reprodutível para profissionais com diversos níveis de experiência. A classificação de Roussouly para os tipos de lordose lombar foi criada com intuito de prover uma ferramenta capaz de

Tabela 2 Distribuição de frequências da classificação de Roussouly

Aferição	Avaliador	Tipo de lordose lombar segundo a Classificação de Roussouly			
		1	2	3	4
Aferição 1	A1	9 (9%)	11 (11%)	59 (57%)	25 (24%)
	A2	7 (7%)	12 (12%)	57 (55%)	28 (27%)
	B1	14 (13%)	10 (10%)	23 (22%)	57 (55%)
	B2	8 (8%)	8 (8%)	57 (55%)	31 (30%)
	C1	7 (7%)	8 (8%)	58 (56%)	31 (30%)
	C2	7 (7%)	10 (10%)	59 (57%)	28 (27%)
Aferição 2	A1	6 (6%)	9 (9%)	64 (62%)	25 (24%)
	A2	10 (10%)	9 (9%)	57 (55%)	28 (27%)
	B1	19 (18%)	8 (8%)	30 (29%)	47 (45%)
	B2	10 (10%)	5 (5%)	57 (55%)	32 (31%)
	C1	3 (3%)	8 (8%)	61 (59%)	32 (31%)
	C2	6 (6%)	14 (13%)	62 (60%)	22 (21%)

Avaliadores segundo a experiência: A, Cirurgiões de Coluna; B, Residentes em Cirurgia da Coluna; C, Residentes em Ortopedia e Traumatologia.

 © 2021. Sociedade Brasileira de Ortopedia e Traumatologia. All rights reserved.

Tabela 3 Coeficientes Kappa de Fleiss dispostos na ►**Figura 1**, com intervalos de confiança e valores *p* dos testes aplicados

Avaliador	Kappa de Fleiss	IC95%	*valor-p*
A1	0,601	(0,462–0,740)	$< 0,001^*$
A2	0,827	(0,738–0,915)	$< 0,001^*$
B1	0,307	(0,163–0,452)	$< 0,001^*$
B2	0,710	(0,586–0,833)	$< 0,001^*$
C1	0,580	(0,440–0,720)	$< 0,001^*$
C2	0,557	(0,407–0,708)	$< 0,001^*$

Abreviação: IC, intervalo de confiança.
Nota-se que, em todos os casos, a concordância obtida se mostrou significativa, ao nível de 5% de significância ($p < 0,001$).

Tabela 4 Teste Kappa de Fleiss para a concordância interavaliador

Avaliador	Kappa de Fleiss	IC95%	*valor-p*
Aferição 1	0,430	(0,344–0,516)	$< 0,001$
Aferição 2	0,369	(0,288–0,451)	$< 0,001$

analisar o alinhamento sagital da coluna ao mesmo tempo em que leva em consideração a orientação pélvica, caracterizando o biotipo do indivíduo.[14]

Os tipos 1 e 2 de Roussouly têm menor inclinação sacral (IS) ($< 35^o$) e lordose de menor valor angular, promovendo maior carga na região anterior da coluna, podendo estar mais associados a degeneração discal[3,6,7] e dor lombar baixa crônica.[9] O tipo 3 é o mais frequente em populações assintomáticas, mesmo entre diferentes etnias e faixas etárias.[1,2] Já o tipo 4 é o que possui maior IS ($> 45^o$) e lordose lombar, estando mais relacionado a espondilolistese e sobrecarga facetária.[8]

No estudo realizado, a lordose do tipo 3 foi a mais encontrada pela maioria dos avaliadores, concordando com a frequência encontrada em estudos prévios na população assintomática.[1,2] Este tipo de lordose foi considerado como mais fisiológico por Roussouly.[14]

A aferição dos ângulos espinopélvicos, necessária para a classificação, pode ser realizada manualmente em radiografias panorâmicas da coluna, em perfil, com inclusão da pelve e cabeças femorais, com uso do goniômetro; no entanto, o software Surgimap já foi validado e facilita a aferição.[19]

Apesar da classificação de Roussouly apresentar critérios objetivos na sua descrição, dependendo principalmente de referências mensuráveis e bem definidas, a variação de apenas 1° na IS é capaz de mudar o tipo de lordose, fazendo com que pacientes cujos parâmetros sejam limítrofes aos cortes ($\sim 35^o$ ou 45°) possam apresentar classificações discrepantes entre diferentes observadores ou aferições distintas de um mesmo observador. Além disso, a definição do ápice da lordose também pode ser duvidosa, possibilitando divergências entre os tipos 1 e 2. Desta forma, é válida a hipótese de que a maior ou menor presença de colunas com tais características pode afetar a reprodutibilidade da classificação.

O presente estudo demonstrou boa reprodutibilidade da classificação de Roussouly, à medida que tanto a concordância intra-avaliador quanto a interavaliador apresentaram coeficientes de concordância Kappa de Fleiss no mínimo razoáveis ($> 0,20$), sendo que a concordância intra-avaliador variou de razoável a quase perfeita, e a interavaliador, de razoável a moderada.

O avaliador B1 destacou-se dos demais, apresentando menor concordância intra- e interavaliador, além de ser o único a encontrar maior prevalência de lordose do tipo 4. Tais diferenças podem ser explicadas por alguma divergência na interpretação da classificação, erro técnico de aferição, ou inerente ao fato de aferições com valores próximos poderem ser classificadas como tipos diferentes.

Também foi observado que a experiência parece afetar a reprodutibilidade da classificação, pois os avaliadores mais experientes (A1 e A2) apresentaram maior concordância intra- e interavaliador, o que pode ser justificado pela maior

Tabela 5 Teste Kappa de Fleiss para a concordância interavaliador (par a par)

Medição	Avaliador	A1	A2	B1	B2	C1	C2
Aferição 1	A1	1,000					
	A2	0,619	1,000				
	B1	0,192	0,261	1,000			
	B2	0,488	0,412	0,236	1,000		
	C1	0,565	0,583	0,218	0,434	1,000	
	C2	0,597	0,584	0,196	0,516	0,496	1,000
Aferição 2	A1	1,000					
	A2	0,439	1,000				
	B1	0,222	0,138	1,000			
	B2	0,483	0,458	0,283	1,000		
	C1	0,515	0,539	0,168	0,449	1,000	
	C2	0,440	0,404	0,166	0,496	0,325	1,000

 © 2021. Sociedade Brasileira de Ortopedia e Traumatologia. All rights reserved.

prática com essas aferições e entendimento dos ângulos espinopélvicos que cirurgiões de coluna formados possuem em relação aos residentes em treinamento.

Conclusão

A classificação de Roussouly demonstrou boa confiabilidade e reprodutibilidade, tendo em vista que as concordâncias intra- e interavaliadores foram consideradas no mínimo razoáveis, podendo atingir os níveis de substanciais à quase perfeitas em algumas situações. Avaliadores com maior nível de experiência apresentaram maior concordância intra e interavaliador.

Suporte Financeiro
Não houve suporte financeiro de fontes públicas, comerciais, ou sem fins lucrativos.

Conflito de Interesses
Os autores não têm conflito de interesses a declarar.

Referências

1 Hu P, Yu M, Sun Z, et al. Analysis of global sagittal postural patterns in asymptomatic Chinese adults. Asian Spine J 2016;10 (02):282–288
2 Bassani T, Galbusera F, Luca A, Lovi A, Gallazzi E, Brayda-Bruno M. Physiological variations in the sagittal spine alignment in an asymptomatic elderly population. Spine J 2019;19(11):1840–1849
3 Bae J, Lee SH, Shin SH, Seo JS, Kim KH, Jang JS. Radiological analysis of upper lumbar disc herniation and spinopelvic sagittal alignment. Eur Spine J 2016;25(05):1382–1388
4 Lowe T, Berven SH, Schwab FJ, Bridwell KH. The SRS classification for adult spinal deformity: building on the King/Moe and Lenke classification systems. Spine 2006;31(19, Suppl):S119–S125
5 Schwab FJ, Smith VA, Biserni M, Gamez L, Farcy JP, Pagala M. Adult scoliosis: a quantitative radiographic and clinical analysis. Spine 2002;27(04):387–392
6 Menezes-Reis R, Bonugli GP, Dalto VF, da Silva Herrero CF, Defino HL, Nogueira-Barbosa MH. Association between lumbar spine sagittal alignment and L4-L5 disc degeneration among asymptomatic young adults. Spine 2016;41(18):E1081–E1087
7 Mardare M, Oprea M, Popa I, Zazgyva A, Niculescu M, Poenaru DV. Sagittal balance parameters correlate with spinal conformational type and MRI changes in lumbar degenerative disc disease: results of a retrospective study. Eur J Orthop Surg Traumatol 2016;26(07):735–743
8 Labelle H, Roussouly P, Berthonnaud E, Dimnet J, O'Brien M. The importance of spino-pelvic balance in L5-s1 developmental spondylolisthesis: a review of pertinent radiologic measurements. Spine 2005;30(06, Suppl):S27–S34
9 Chaléat-Valayer E, Mac-Thiong J, Paquet J, et al. Sagittal spino-pelvic alignment in chronic low back pain. Eur Spine J 2011; 20Suppl 5(Suppl 5):634–640
10 Schwab FJ, Blondel B, Bess S, et al.International Spine Study Group (ISSG) Radiographical spinopelvic parameters and disability in the setting of adult spinal deformity: a prospective multicenter analysis. Spine 2013;38(13):E803–E812
11 Barrey C, Roussouly P, Le Huec JC, et al. Compensatory mechanisms contributing to keep the sagittal balance of the spine. Eur Spine J 2013;22 Suppl 6(Suppl 6):S834–841
12 Le Huec JC, Thompson W, Mohsinaly Y, Barrey C, Faundez A. Sagittal balance of the spine. Eur Spine J 2019;28(09):1889–1905
13 Le Huec JC, Faundez A, Dominguez D, Hoffmeyer P, Aunoble S. Evidence showing the relationship between sagittal balance and clinical outcomes in surgical treatment of degenerative spinal diseases: a literature review. Int Orthop 2015;39(01):87–95
14 Roussouly P, Gollogly S, Berthonnaud E, Dimnet J. Classification of the normal variation in the sagittal alignment of the human lumbar spine and pelvis in the standing position. Spine 2005; 30(03):346–353
15 Roussouly P, Berthonnaud E, Dimnet J. [Geometrical and mechanical analysis of lumbar lordosis in an asymptomatic population: proposed classification]. Rev Chir Orthop Repar Appar Mot 2003; 89(07):632–639
16 Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: Wiley; 1981
17 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(01):159–174
18 R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing-Vienna, Austria2015. Disponível em: <http://www.Rproject.org
19 Lafage R, Ferrero E, Henry JK, et al. Validation of a new computer-assisted tool to measure spino-pelvic parameters. Spine J 2015;15 (12):2493–2502

 © 2021. Sociedade Brasileira de Ortopedia e Traumatologia. All rights reserved.