The intra-rater reliability of a revised 3-point grading system for accessory joint mobilizations

Jennifer Ward^a, Clair Hebron^a and Nicola J. Petty^b

^aSchool of Health Sciences, University of Brighton, Eastbourne, UK; ^bSchool of Health Sciences, Centre for Health Research, University of Brighton, Eastbourne, UK

ABSTRACT

Objectives: Joint mobilizations are often quantified using a 4-point grading system based on the physiotherapist's detection of resistance. It is suggested that the initial resistance to joint mobilizations is imperceptible to physiotherapists, but that at some point through range becomes perceptible, a point termed R1. Grades of mobilization traditionally hinge around this concept and are performed either before or after R1. Physiotherapists, however, show poor reliability in applying grades of mobilization. The definition of R1 is ambiguous and dependent on the skills of the individual physiotherapist. The aim of this study is to test a revised grading system where R1 is considered at the beginning of range, and the entire range, as perceived by the physiotherapists maximum force application, is divided into three, creating 3 grades of mobilization.

Method: Thirty-two post-registration physiotherapists and nineteen pre-registration students assessed end of range (point R2) and then applied 3 grades of AP mobilizations, over the talus, in an asymptomatic models ankle. Vertical forces were recorded through a force platform. Intra-class Correlation Coefficients, Standard Error of Measurement, and Minimal Detectable Change were calculated to explore intra-rater reliability on intra-day and inter-day testing. T-tests determined group differences.

Results: Intra-rater reliability was excellent for intra-day testing (ICC 0.96–0.97), and inter-day testing (ICC 0.85–0.93). No statistical difference was found between pre- and post-registration groups.

Discussion: Standardizing the definition of grades of mobilization, by moving R1 to the beginning of range and separating grades into thirds, results in excellent intra-rater reliability on intra-day and inter-day tests.

Level of Evidence: 3b.

KEYWORDS

Joint mobilization;
movement diagram; grades
of mobilization; intra-rater
reliability

Introduction

Accessory joint mobilizations involve the application of passive rhythmic oscillatory forces to the soft tissues overlying the joint [1] and are often quantified using a subjective grading system. Margarey [2] and Maitland [3] advocated a 4-point grading system based on the physiotherapists' assessment of resistance to movement. They suggested that although resistance would be encountered the moment that joint movement begun, the initial resistance would be so minor that it would be imperceptible. The point at which resistance became perceptible was termed R1, and became the point around which grades of movement were defined. Margarey [2] and Maitland [3] suggested that grades I and II movements be applied in resistance free range, before R1. Grades III and IV were applied into resistance. Margarey [2] highlighted that the amount of resistance into which the movement be performed might vary, and could be depicted using one or several + or – symbols.

Maitland [1] and Magarey [2] developed a movement diagram to aid communication between physiotherapists and to be used as a teaching tool. It depicts the behavior of resistance, spasm, and pain through the available range of movement and can be used to document a grade of movement used in treatment. The physiotherapist identified on the x-axis the point in range where resistance was first felt (R1). Where resistance limited movement, and the physiotherapist was not prepared to apply any more force, (R2) was depicted as a thick black line at the end of available range. The line between R1 and R2 depicted the quality of the resistance felt. Lee and Evans [4] suggested that the resistance curve documented on a movement diagram could be considered analogous with a force displacement curve (Figure 1) where R1 was suggested to occur at the point between toe and linear resistance.[4–6]

Petty et al. [7] sought to more accurately define R1 by examining 30 force displacement graphs from spinal

Figure 1. Relationship of movement diagram (ABCD) to load displacement curve (Reproduced with kind permission from Lee and Evans [4]).

and peripheral joints. While in most graphs a distinct toe region was found, they failed to find a clear point of inflexion where toe became linear resistance, and this challenged the concept of R1. Petty et al. [7] thus proposed R1 would be more accurately considered to occur at the start of movement, point A, on the movement diagram.

If the concept of R1 occurring some point through range is questioned, the grading system that hinges around the concept of R1 must also be questioned.[1,2] Unsurprisingly, research that has explored physiotherapist's application of grades of movement defined by R1 has shown poor reliability, particularly during inter-rater reliability (ICC 0.03–0.05).[8–10] However, variable forces between physiotherapists during grades of mobilizations are expected as patient factors such as age,[11] disability, location and severity of symptoms,[12] weight, spinal stiffness, range of movement,[13] and physiotherapist factors such as experience, qualification, and mobilization frequency have all been shown to influence forces used.[11–13] Variable forces might therefore represent best clinical practice, as the physiotherapist adapts manual therapy application to individual patient needs and characteristics.

The research on intra-rater reliability using grades of movement defined by R1, has also shown poor reliability. Harms and Bader [8] explored the intra-rater reliability of applying grades I–IV mobilizations, defined by the detection of resistance, on the L3 vertebra. Thirty experienced physiotherapists were recruited and applied intra-day and inter-day force applications. The results showed highly variable forces, particularly for grade I and II mobilizations, which on intra-day testing varied by 63 and 44%, and on inter-day testing varied by 114 and 94%.

Snodgrass et al. [11] found more favorable results for intra-rater reliability. One hundred and sixteen physiotherapists applied 4 grades of mobilization to the C2 and C7 spinal levels. The results showed good intra-rater reliability on intra-day testing, (ICC 0.93). However, there was no standardized definition of mobilization grades given, rather the participants chose their own definitions, based either on the resistance felt,[3] or the range in movement.[14] When asked what grading system they had used, participants provided a wide range of descriptions, with up to 22 different variables [11] often not related to the detection of R1, but rather related to the available range. It is argued that eliminating the ambiguities of detecting the onset of resistance improved the intra-rater reliability.

Silvernail et al. [15] explored the intra-rater reliability of forces applied during accessory joint mobilizations to the patellofemoral joint, in one physiotherapist, completing two repeated tests. They found good reliability for force used during inferior glides of the patella (ICC 0.93–0.95), and medial lateral glides of the patella (ICC 0.75–0.96). Tragord et al. [16] used a similar methodology to explore the intra-rater reliability of medial glides of the patella, in six physiotherapists. They also found good reliability with ICC values between 0.95 and 0.98. However, these studies focused on the application of grade III and IV mobilizations only, which by definition are applied into resistance. The assessment of R1 is therefore not directly relevant and better results could be expected. In addition, Tragord et al. [16] provided participants with further clarification of the force required, by asking them to target 50% of the range of movement, which removes the ambiguities of the traditional grading system used in other studies.

In light of the challenges highlighted in the grading system defined by R1, Petty [17] proposed a revised system, where the onset of resistance, R1 was considered to start at the beginning of range of motion. The physiotherapist perception of the end of range of motion, or the maximum force they were prepared to apply determined R2, and this was expected to vary between physiotherapists. Three grades of movement were then defined, assuming the resistance was linear, occurring within the first, middle, or last third of range, relative to the assessment of R2. The oscillation amplitude was disentangled from the definition of grade, and noted separately as a small or large oscillation. The aim of this study therefore was to test the intra-rater reliability of applying this revised grading system, during AP mobilizations on the talus which have been shown to improve joint stiffness and dorsiflexion ROM following ankle sprain.[18,19] The aim was also to compare physiotherapists with varying levels of experience as there is currently no research on the impact of experience and qualifications on the intra-rater reliability of performing grades of mobilization.

Methods

Participants

Participants were recruited via email and then grouped according to their experience and qualifications. Group 1 consisted of twenty-eight chartered physiotherapists (with a minimum of five years experience), all of whom were attending a postgraduate MSc module in neuromusculoskeletal physiotherapy. In addition, there were four lecturers with MSc qualification in Musculoskeletal Physiotherapy and members of the Musculoskeletal Association of Chartered Physiotherapists (MACP). Group 2 consisted of nineteen pre-registration physiotherapy students. All participants had experience in the use of applying grades of joint mobilizations and were excluded if they had any recent trauma to the upper limb, neck or thoracic spine, or any pain or impairment that limited or prevented their usual performance of joint mobilization techniques.

The demographic data for physiotherapists are illustrated in Table 1.

One male and two female models were recruited for the study, with a mean BMI of 23.98 (range 22.62–25.71) and mean age of 34 (SD 4).

Apparatus

Force plates (AMTI OR6-7 Advanced Mechanical Technology Inc., Watertown, MA, USA) were used to measure forces applied to the model. The model lay supine on a lightly padded plinth, which was bolted onto the force platform, so that forces could be indirectly measured as they were transferred through the plinth. Since horizontal forces have been shown to only play a small part in joint accessory mobilizations, only vertical forces were calculated.^[20] A lean bar, positioned next to the side of the plinth, prevented participants from leaning against the plinth and this ensured that all the vertical forces were transferred to the force plate.^[20] The models ankle was positioned at 90 degrees from the shaft of the fibula and stabilized using a wooden block (see Figure 2).

Procedure

For each assessment, the model lay on the same wooden plinth and the same researcher marked the

Figure 2. Picture of the ankle positioned at 90° from the shaft of the fibular.

Notes: The ankle was measured at 90° with a goniometer to ensure consistency. The foot was stabilized against a wooden block, which was clamped onto the plinth. 1 folded towel was placed under the models heel for comfort. The lean bar positioned at the side of the plinth, was stabilized with weights, allowing the physiotherapist to lean against it while applying the mobilizations.

talus. The ankle was positioned at 90 degrees from the shaft of the fibula for all tests. This process was repeated for each model during each data collection period. Only five of the participants were familiar with the revised grading system and used it regularly in clinical practice, therefore all participants were provided with a diagram explaining the revised concept of R1, R2, and the thirds of resistance. To aid understanding, this process was repeated verbally prior to initial testing. Participants were allowed to use any amplitude they chose, as this was not being recorded. They were then offered the opportunity to ask questions or seek clarification concerning any aspect of the task. Participants were allowed time to familiarize themselves with the joint accessory motion at the ankle. The participant placed their hands over the anterior surface of the talus and applied repeated oscillatory AP mobilizations to fully explore the range and judge where point R2 occurred (see Figure 3). A minimum of three oscillations to end range has been shown to adequately precondition the tissues,^[21] therefore participants were asked to assess R2 at least 3 times prior to testing.

Once R2 had been fully explored by the participants, they began oscillating within the first third of resistance and data collection started. After approximately 6 s participants were prompted to oscillate into the middle third of resistance, without removing their hands, and then into the last third of resistance. Participants were also asked to find R2. Participants were free to use amplitude and speed of their choice for each mobilization, as this was not being analyzed. After a break of 5 min, this process was repeated again in the same order. The participants then returned approximately 1 week later to repeat the same process, but this time only once.

Table 1. Physiotherapist participant demographic data.

	M/F	Height (cm)	Weight (kg)	Years of experience
Post-registration (32)	16/16			
Mean (SD)		172 (9.1)	72.4 (13.7)	6.2 (3.8)
Range		157–189	45–95	2–16
Pre-registration (19)	8/11			
Mean (SD)		169.9 (10.3)	69.1 (13.9)	N/A
Range		155–188	47–100	

Figure 3. Picture of the hand hold used to apply AP mobilizations to the talus.

The School of Health Professions, School of Research Ethics and Governance Panel approved the study.

Data analysis

The mobilization forces used were converted into graph format (Figure 4). R2 force was identified as the maximum peak force in the final third of resistance.

The graphs were analyzed to determine the mean peak force during the first, middle, and last third of resistance. The initial peak force of each third was discounted, as Latimer et al. [5] suggested this initial force reading is more variable in nature than the remaining forces. The mean peak force for the first, middle, and last third of resistance was calculated from the 2nd, 3rd, and 4th oscillation.

Normality of data for the mean peak forces for each third and R2 was tested in SPSS using Shapiro–Wilk, which showed significant deviations for most data (p 0.004–0.155). Despite the deviations found, data used to calculate ICCs were not transformed as these tests are considered robust to deviations.[22] The consequence of transforming the data would have been to change the construct of the data, meaning that SEM and MDC could not be calculated in N, therefore limiting the interpretation of the data.[22]

Pooled intra-rater reliability was established using ICC_{3,1} using the single measurement from the mean peak force of the 2nd, 3rd, and 4th oscillation for each third of mobilization and R2. The SEM was calculated, and the minimal detectable change was calculated using the SEM to represent the 95% confidence level. Groups were analyzed together where no group differences were found, and separately where group differences were evident. Group differences were analyzed using t-tests on transformed data, which was transformed using log10 algorithm.[23]

Figure 4. Example of force data collection from one participant demonstrating the oscillations performed during the first, middle, and last third of resistance and R2.

Results

There was no statistically significant difference between pre- and post-registration groups ($p > 0.05$), except on intra-day tests in the first third of resistance, where the post-registration group applied lower mean forces on initial testing ($p = 0.03$) and on repeat testing ($p = 0.04$). Therefore, group results were combined for intra-rater reliability, in all tests except the first third of resistance on intra-day tests. The mean, standard deviation, and range of data for all participants during repeated tests was calculated (see Table 2).

The ICC_{3,1} values in the study of between 0.96 and 0.97 for intra-day testing and 0.85 and 0.93 for inter-day testing represented excellent reliability (Table 3).

Only small variation in force was found on SEM calculations. One SEM represents a confidence level of 68%, so for example, when applying R2 forces, physiotherapist forces would vary by 12.4 N on 68% of occasions on intra-day tests and by 25.6 N on inter-day tests (Table 3).

The SEM was used to calculate the Minimal Detectable change and this value was calculated as a percentage of the original first force application (Table 4). These results also showed small percent differences in physiotherapists force applications on intra-day testing of between 25 and 34%. The percent differences on inter-day testing were larger between 38 and 61%.

Table 2. Descriptive statistics between repeat tests.

Repeat test	Point in range	Day 1/test 1 (N)	Day 1/test 2 (N)	Day 2 (N)
Mean (SD)	1st third	Pre reg 100 (39)	91 (40)	81 (36)
		Post reg 79.2 (36)		
	2nd third	116 (43)	113 (42)	109 (42)
	3rd third	160 (52)	158 (54)	151 (54)
	R2	172 (55)	170 (55)	163 (54)
Range	1st third	Pre reg 41–200	24–195	19–200
		Post reg 19–187		
	2nd third	44–239	47–241	27–259
	3rd third	77–311	69–309	60–316
	R2	78–322	79–314	67–313

Note: (N) = Newton's.

Table 3. ICC with 95% confidence interval and SEM(N).

	ICC intra-day	SEM	ICC inter-day	SEM
1st third	Pre reg 0.97 (0.93–0.98)	9.1	0.93 (0.82–0.97)	13.7
	Post reg 0.97 (0.93–0.99)	9.7	0.87 (0.73–0.94)	17.2
2nd third	0.96 (0.93–0.98)	12.3	0.85 (0.74–0.91)	21.9
3rd third	0.97 (0.96–0.99)	12	0.88 (0.80–0.93)	24.5
R2	0.97 (0.96–0.99)	12.4	0.88 (0.79–0.93)	25.6

Notes: ICC – Intra-class correlation coefficient (95% confidence interval).
SEM = standard error of measurement.

Discussion

Excellent results were found for intra-rater reliability on intra-day testing with ICC_{3,1} values of 0.96–0.97 which is better than those found by Snodgrass et al. [11] of 0.84–0.93. The MDC also demonstrated excellent results, particularly on intra-day testing with forces varying by 25–34%. These results are favorable compared to Harms and Bader [8] who calculated the 95% limit of agreement, which is similar to the MDC and expressed this as a percentage of the original force application. Harms and Bader [8] found forces varied by 63 and 44% for grade I and II mobilizations, and by 24–40% for grade III, IV and end feel mobilizations when participants were asked to apply grade I and II mobilizations before the onset of perceptible resistance (R1). However, R1 has been found difficult to accurately identify during examination of force displacement data.[7] The results in this study suggest that redefining grades of mobilization improves intra-rater reliability on intra-day testing. Intra-day reliability is relevant in clinical practice, as physiotherapist will often apply several different joint mobilization doses within a physiotherapy session, commonly lasting 30–60 s, with a short break between doses.[3] It seems important that physiotherapists can reliably replicate force between repetitions within a treatment session, to ensure consistency and to have the ability to deliberately reduce or increase force to regress or progress the treatment dose. Lack of control of force application could potentially be detrimental to patient care.

Good results were also found for intra-rater reliability on inter-day testing with ICC values of between 0.85 and 0.93. Although the reliability on inter-day testing is slightly lower than intra-day testing, this is expected as the viscoelasticity of tissues can vary with temperature,

time of day, and activity levels of the model.[11] While room temperature remained the same on each day of testing, the temperature of the model was not measured and may have varied on different days. Models were asked to avoid variation in activity levels between days but this could not be strictly controlled. Finally, while every effort was made to repeat measures at the same time of day this was not always the case and length of time since getting up from bed, which would affect the ankle, was not standardized. Snodgrass et al. [11] argued that inter-day testing should not be explored as these factors would make the results too variable, however in clinical practice it is standard to treat patients with multiple treatment sessions on different days [3] where both patient and physiotherapist factors might vary. Inter-day reliability therefore remains an important aspect of the physiotherapist's treatment, and the good ICC results found in this study suggest that physiotherapists were reliable in their force application, despite differences across days.

The MDC also demonstrated good results on inter-day testing with forces varying by 38–61%. These results are again favorable compared to Harms and Bader [8] who found much wider variations in force application on inter-day tests of up to 114 and 94% for grade I and II mobilizations respectively, and by 32–55% for grade III, IV, and end feel mobilizations. This suggests that the 3-point grading system resulted in improved inter-day reliability.

Overall there was no significant difference between pre- and post-registration groups in the reliability of applying the 3-point grading system. However, the pre-registration group, who had less qualification and experience than the post-registration group, applied higher mean peak forces in the first third of resistance. Chiradejnant et al. [12] also found that physiotherapists with less experience applied higher mean peak forces during grades of mobilization, and this suggests that experience and qualifications may improve physiotherapist's ability to apply mobilizations within the first third of resistance. No previous research has explored the effect of physiotherapist qualifications and experience on intra-rater reliability when applying grades of mobilization, but this study suggests that overall, physiotherapists of all levels of experience are able to reliably replicate the 3-point grading system.

Table 4. Minimal detectable change (MDC) representing 95% confidence level.

	MDC95 intra-day test	Intra-day MDC as a percentage of initial force (%)	MDC95 inter-day test	Inter-day MDC as a percentage of initial force (%)
1st 3rd				
Pre reg	25	25	38	38
Post reg	27	34	48	61
2nd 3rd	34	29	61	53
3rd 3rd	33	21	68	43
R2	34	20	71	41

Note: MDC expressed as a % of the initial force application (day1, test1).

Limitations

In order to accurately record vertical forces, participants had to apply the mobilization in a way that differed to their typical practice. They were required to use a lean bar, which kept them at a distance from the model and prevented them from applying the mobilization in their usual way. The participant's handhold was limited to an overhand grip over the anterior surface of the talus only. Several participants commented on how unnatural this felt. It seems reasonable to suggest that these alterations to the participant's normal procedure in applying an AP to the talus would affect their ability to accurately and reliably perform the mobilizations. Participants were required to move from the first third oscillations, to the middle third, and to the last third and R2 without taking their hands off the talus. Participants noted they would normally remove their hands and re-palpate when changing the grade of movement, and that the successive application of different grades felt quite unnatural. In addition, most participants were not familiar with the new grading system and therefore would have been learning how to apply a new technique during data collection. These requirements would be expected to lead to a reduction in reliability.

The physiotherapist used the initial exploration of R2 as the reference point from which to apply oscillations in the first third and subsequent grades, but this R2 force was not measured; R2 force was instead measured during the final third of resistance, in order to reduce the complexity of data collection. It is uncertain whether the measured R2 force accurately reflects the R2 force used by the physiotherapist to determine the grade of movement and could have led to inaccurate data. During data analysis, the reliability was captured from the initial oscillations and not from the full 6 s, which might have biased towards better reliability than measuring over a longer time period.

The results of this study are limited by the use of asymptomatic models, and cannot be generalized to a patient population, where pain and abnormal movement may be present during the AP movement. The results are also limited to the ankle, and future work focusing on other peripheral and spinal joints are needed.

Conclusion

The redefined grades of mobilizations that assumed R1 started at A in the movement diagram were found to have excellent intra-rater reliability on inter-day and intra-day testing. Physiotherapists of all levels of qualification and experience were equally able to apply this 3-point grading system with excellent intra-rater reliability.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Jennifer Ward was responsible for study design, obtaining ethics approval, logistic, administrative and technical responsibilities of data collection, data analysis, drafting the paper, critical revision of paper and final approval of article.

Clair Hebron was responsible for study design, drafting ethical approval, methodology, supervision of data collection, data analysis, critical appraisal of paper and revision of final approval of article.

Nicola J. Petty was responsible for study design, methodology, critical appraisal of paper and revision of final approval of article. Paul Conrad and Victoria Hill were responsible for assistance in data collection.

ORCID

Nicola J. Petty <http://orcid.org/0000-0002-4327-7195>

References

- [1] Maitland G. Vertebral manipulation. 5th ed. Oxford: Butterworths Heinemann; 1986.
- [2] Margarey ME. Examination and assessment in spinal joint dysfunction. Chapter 44. In: Grieve G, editor. Modern manual therapy of the vertebral column. Edinburgh: Churchill Livingstone; 1984. p. 481–497.
- [3] Maitland G. Vertebral manipulation. 6th ed. Oxford: Butterworths Heinemann; 2001.
- [4] Lee R, Evans J. Towards a better understanding of spinal posteroanterior mobilisation. *Physiotherapy*. 1994;80:68–73.
- [5] Latimer J, Goodsell MM, Lee M, et al. Evaluation of a new device for measuring response to posteroanterior force in a patient population, Part 1: reliability testing. *Phys. Ther.* 1996;76:158–165.
- [6] Shirley D, Ellis E, Lee M. The response of posteroanterior lumbar stiffness to repeated loading. *Man. Ther.* 2002;7: 19–25.
- [7] Petty NJ, Maher C, Latimer J, et al. Manual examination of accessory movements-seeking R1. *Man. Ther.* 2002;7:39–43.
- [8] Harms MC, Bader DL. Variability of forces applied by experienced therapists during spinal mobilization. *Clin. biomech.* 1997;12:393–399.
- [9] Mitchell WN. Reliability in the performance of grade II and grade IV mobilizations [Unpublished postgraduate diploma dissertation]. Melbourne: Lincoln Institute of Health Sciences. School of Physiotherapy; 1983. Cited in Matayas B, Bach T. The reliability of selected techniques in clinical arthrometrics. *Aust. J. Physiother.* 1985;31:175–199.
- [10] Cook C, Turney L, Ramirez L, et al. Predictive factors in poor inter-rater reliability among physical therapists. *J. Man. Manipulative Ther.* 2002;10:200–205.
- [11] Snodgrass SJ, Rivett DA, Robertson VJ, et al. Forces applied to the cervical spine during posteroanterior mobilization. *J. Manipulative Physiol. Ther.* 2009;32:72–83.
- [12] Chiradejnant A, Latimer J, Maher C. Forces applied during manual therapy to patients with low back pain. *J. Manipulative Physiol. Ther.* 2002;25:362–369.
- [13] Harms MC, Innes SM, Bader DL. Forces measured during spinal manipulative procedures in two age groups. *Rheumatology (Oxford)*. 1999;38:267–274.

- [14] Grieve GP. Mobilisation of the spine. 5th ed. Edinburgh: Churchill Livingstone; 1991.
- [15] Silvernail JL, Gill NW, Teyhen DS, et al. Biomechanical measures of knee joint mobilization. *J. Man. Manipulative Ther.* 2011;19:162–171.
- [16] Tragord BS, Gill NW, Silvernail JL, et al. Joint mobilization forces and therapist reliability in subjects with knee osteoarthritis. *J. Man. Manipulative Ther.* 2016;21:196–206.
- [17] Petty NJ. Neuromusculoskeletal examination and assessment: a handbook for therapists. 4th ed. Edinburgh: Churchill Livingstone; 2011.
- [18] Green T, Refshauge K, Crosbie J, et al. A randomized controlled trial of a passive accessory joint mobilization on acute ankle inversion sprains. *Phys. Ther.* 2001;81:984–994.
- [19] Vincenzino B, Branjerdporn M, Teys P, et al. Initial changes in posterior talar glide and dorsiflexion of the ankle after mobilization with movement in individuals with recurrent ankle sprain. *J. Orthop. Sports Phys. Ther.* 2006;36:464–471.
- [20] Petty NJ, Messenger N. Can the force platform be used to measure the forces applied during a PA mobilization of the lumbar spine? *J. Man. Manipulative Ther.* 1996;4:70–76.
- [21] Lee R, Evans J. Load–displacement–time characteristics of the spine under posteroanterior mobilisation. *Aust. J. Physiother.* 1992;38:115–123.
- [22] Field A. Discovering statistics using IBM SPSS statistics. 3rd ed. London: Sage; 2009.
- [23] Agresti A, Finlay B. Statistical methods for the social sciences: with SPSS from A to Z: a brief step-by step manual. 4th ed. Boston (MA): Allyn & Bacon; 2009.