Journal of Manual & Manipulative Therapy

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

--Manuscript Draft--

Manuscript Number:	
Full Title:	The intra-rater reliability of a revised 3 point grading system for accessory joint mobilizations
Article Type:	Original Research Paper (4,000 words limit)
Keywords:	joint mobilization movement diagram grades of mobilization intra-rater reliability
Corresponding Author:	Jennifer Rosemary Ward, MSc Neuromusculoskeletal Physiotherapy University of Brighton Eastbourne, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Brighton
Corresponding Author's Secondary Institution:	
First Author:	Jennifer Rosemary Ward, MSc Neuromusculoskeletal Physiotherapy
First Author Secondary Information:	
Order of Authors:	Jennifer Rosemary Ward, MSc Neuromusculoskeletal Physiotherapy
	Dr Nicola J Petty, PhD
	Dr Clair Hebron, PhD
Order of Authors Secondary Information:	
Abstract:	Objectives Joint mobilizations are often quantified using a 4 point grading system based on the therapist's detection of resistance. It is suggested that the initial resistance to joint mobilizations is imperceptible to therapists, 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 individual therapists. 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 therapist maximum force application, is divided into 3, creating 3 grades of mobilization. Method Thirty two post-registration physiotherapists and 19 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.
Funding Information:	

Title page

The intra-rater reliability of a revised 3 point grading system for accessory joint mobilizations, tested on the ankle

Authors

Jennifer Ward affiliated with;
School of Health Sciences
University of Brighton
Aldro Building
49 Darley Road
Eastbourne BN20 7UR
United Kingdom

Dr. Clair Hebron
Senior Lecturer Physiotherapy
Course Leader MSc Neuromusculoskeletal Physiotherapy and MSc Professional Health and Social Care Practice
School of Health Sciences
University of Brighton
Aldro Building
49 Darley Road
Eastbourne BN20 7UR
United Kingdom

Dr. Nicola J Petty Principal Lecturer
Centre for Health Research
School of Health Sciences
University of Brighton
Aldro Building
49 Darley Road
Eastbourne BN20 7UR
United Kingdom
Surrey and Sussex Regional Hub Leader, Council for Allied Health Professions Research

No external funding
No conflicts of interest among the authors

Maney Publishing
Research • Knowledge • Innovation

www.maneypublishing.com

Journal of Manual and Manipulative Therapy ('the Journal')

ASSIGNMENT OF COPYRIGHT

In order for W. S Maney & Son Ltd, trading as Maney Publishing ('the Publisher'), to ensure the widest possible dissemination and protection of articles published in the Journal, we request authors to assign worldwide copyright in print, digital and other media in their article, including abstracts, to W. S. Maney & Son Ltd. This enables the Publisher to ensure maximum international copyright protection against infringement, and to disseminate your article, and the Journal, to the widest possible worldwide readership. Should the article be rejected from the Journal for any reason, all rights assigned within this document revert back to the author(s) of the article.

Please read the explanatory notes of this agreement on the following page.

1. In consideration of the undertaking set out in paragraph 2, the author(s) as beneficial owner(s) hereby assign(s) to W. S. Maney & Son Ltd all rights, title and interest in the copyright throughout the world in the article named below (**which shall be taken to include any or all supplementary material associated with the article***) for the full legal term of copyright.

Article title: The intra-rater reliability of a revised 3 point grading system for accessory joint mobilizations.
Name(s) of Author(s): Jennifer Ward, Clair Hebron, Nicola Petty
Manuscript no./ref. (if known)

If this paper is sponsored for Open Access by your grant funders, please follow the MORE OpenChoice route by signing and returning this copyright assignment to the address below, and find details of the policy and how to make payment at www.maneyonline.com/openaccess.

So that there should be no doubt, it is understood and agreed that this assignment includes the right to publish or adapt (subject to paragraphs 3 and 4) the material in the article for use in conjunction with computer systems, including networks. The article may be published in printed, CD-ROM, microfiche, on-line, or other machine-readable form, or in any other format or medium (electronic or otherwise) approved by W. S. Maney & Son Ltd; and subject to data mining.

2. Maney Publishing, as the Journal's publisher, hereby undertakes to prepare for publication and publish in the Journal, the Article named in paragraph 1.

3. The editor of the Journal, and the Publisher, are empowered to make such editorial changes as may be necessary to make the Article suitable for publication. Every effort will be made to consult the Author(s) if substantive changes are required.

4. The author(s) hereby assert(s) their moral rights under the terms of the Copyright Designs and Patents Act 1988 to be identified as the Author(s) of the Article.

5. The author(s) warrant(s) that the article is the author(s)' original work, has not been published before, either in part or whole, and is not currently under consideration for publication elsewhere. The author(s) also warrant(s) that (a) the article contains no libellous or unlawful statements, (b) it in no way infringes the rights of others (including the holders of copyright in material used within the article), (c) that permission has been obtained to reproduce any material or illustrations for which they do not hold copyright, and that (d) the author(s), as the owners of the copyright, is (are) entitled to make the assignment set out in this agreement. The author(s) hereby indemnifies/indemnify the Publisher against any claims for breach of the warranties given above.

6. In cases where the article is a translation of a previously published article, the author(s) warrant(s) that full permission has been obtained to carry out and publish the translation of the article from its original language, and that it in no way infringes the rights of others (including the holders of copyright in material used within the original Article).

7. The author(s) warrant(s) that the article contains no plagiarised material and agree(s) that the article may be checked for plagiarism and other ethical questions (including checks using automated software packages). The Publisher reserves the right to withdraw or retract any article for which, in its reasonable opinion, valid concerns exist over the originality of the article, or its compliance with the Journal's ethical policy. Maney's plagiarism and ethical policy may be read in full at www.maneyonline.com/page/authors/publishingethics-general.

*If the author(s) cannot assign copyright for any or all the supplementary material, the clause in parentheses should be struck out and a licence to publish provided for this material. If copyright is held by a third party, permission to reproduce this material must be obtained from the copyright holder as laid out in paragraph 5 of this form and submitted to the publisher.

Copyright Assignment

I confirm that all the co-authors, named above in 1, know that the article has been submitted to the Journal.

I hereby sign this Assignment of Copyright with the full knowledge and agreement of all authors.

Signed: ..... **Date:** 22/12/15

Printed name: JENNIFER R WARD

Please note: it is essential for publication that Maney Publishing receive this signed Assignment of Copyright form. Without it, we will be unable to proceed to publication.

Please complete all specified fields and return this signed form to the ***Production Editor*** *of the Journal.*

Production Editor, JMT, for Maney Publishing
Email: jmt@maneypublishing.com
297 Euston Road, London, NW1 3AQ, UK

Introduction

Accessory joint mobilizations, widely used by physiotherapists, involve the application of passive rhythmical oscillatory forces to the soft tissues overlying the joint.[1] This technique is often quantified using a subjective grading system. Margarey[2] and Maitland[3] advocated a 4 point grading system based on the therapists' 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 to the therapist. 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 (small amplitude) and II (large amplitude) movements be applied in resistance free range, before R1. Grades III (large amplitude) and IV (small amplitude) 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. For example a grade IV- movement would be performed into a small amount of resistance, whereas a grade IV++ movement would be into a large amount of resistance and might be considered to reach the limit of normal joint range. The grading system was defined according to the onset of resistance (R1) and increase in resistance to end range (R2) as well as the amplitude of movement.

Maitland[1] and Magarey[2] developed a movement diagram to aid communication between therapists 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 therapist identified on the x-axis the point in range where resistance was first felt (R1). Where resistance limited movement and the therapist was not prepared to apply any more force (R2) was depicted as a thick black line at the end of available range. The quality of the resistance felt during the range was identified by a line drawn between R1 and R2. 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 transition point between toe and linear region of resistance.[4-6] Once the behavior of pain or spasm was added to the movement diagram, it could then be used to guide the application of a treatment grade of movement. [1,2]

Figure 1 about here

Petty, Maher, Latimer and Lee[7] sought to more accurately define R1 by examining 30 force displacement graphs from spinal and peripheral joints. While there was in most graphs a distinct toe region, they failed to find a clear point of inflexion where the linear region began. A lack of demarcation between toe and linear region challenged the use of R1 by therapists when examining joint movement. Petty et al[7] thus proposed R1 would be more accurately considered to occur at the start of movement, 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 therapist's application of grades of movement defined by R1 has shown poor reliability. The majority of research has focused on inter-rater reliability with ICC values as low as 0.03-0.05,[8,12,13] suggesting poor reliability. However, variable forces between therapists during grades of mobilizations is expected and warranted as both patient and therapist factors have been shown to influence forces used.[9,14,15] Patient factors such as age, disability, area and bothersomness of symptoms,[14] weight, spinal stiffness, range of movement,[15] and therapist factors such as experience, qualification, and frequency of use of mobilizations have all been shown to influence forces used.[14] Variable forces between therapists might therefore represent best clinical practice, as the therapist is adapting their handling to patients own individual requirements. Additionally we would argue that inter-rater reliability is less clinically relevant than intra-rater reliability as patients are often assessed and treated by only one therapist during their course of treatment.

The research on intra-rater reliability using grades of movement defined by R1, has also however shown poor reliability. Previous work has focused on spinal rather than peripheral joint mobilizations, and therefore is not directly comparable with this research, but can be used to inform common themes identified in therapist 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 therapists were recruited

and applied intra-day and inter-day force applications. The results, when calculated to represent the 95% confidence interval of the population, showed highly variable forces used within therapists repeated measures. Most variability was found for grade I and II mobilizations, which on intra-day testing varied by 63% and 44% respectively. On inter-day testing grades I and II varied by 114% and 94%.

Snodgrass, Rivett, Robertson and Stojanovski[9] found more favorable results for intra-rater reliability. One hundred and sixteen therapists applied 4 grades of mobilization to the C2 and C7 spinal levels, on one of 35 asymptomatic models. The results showed good intra-rater reliability on intra-day testing, with ICC values for vertical forces of 0.93 (0.92-0.94). However there was no standardized definition of grades of mobilization given, rather the participants chose their own definitions, based either on the resistance felt, [3] or the range in movement.[10] When asked what grading system they had used, participants provided a wide range of descriptions, with up to 22 different variables,[9] often not related to the detection of R1, but rather related to the available range. The authors found that participants who used range to define their grades of mobilization, rather than resistance, tended to use higher average forces for grade II mobilizations and it is argued that eliminating the ambiguities of detecting the onset of resistance, improved the intra-rater reliability.

Snodgrass et al[9] also examined the impact of therapist characteristics on the forces used during mobilizations and found that therapists with higher academic qualifications tended to use lower forces, but the level of therapist

experience and frequency of use of PA mobilizations had no significant effect on forces used. They did not explore the impact of therapist experience and qualifications on the intra-rater reliability however, and there is currently no research available in this area.

In light of the challenges highlighted in the grading system defined by R1, Petty[16] proposed a revised system, where the onset of resistance, R1, was considered to start at the beginning of range. The therapist perception of end range, or the maximum force they were prepared to apply determined R2, and this was expected to vary between therapists. Three grades of movement were then defined, occurring within the first, middle or last third of resistance, relative to the therapist's individual assessment of R2. If the resistance were assumed to be linear, then the three grades would be in the first, middle and last third of range. The amplitude of oscillation was disentangled from the definition of grade, and noted separately as a small or large oscillation. So, for example the treatment dose may be described as a *small amplitude movement in the middle third of resistance*. This revised and cruder grading system, which removes the ambiguities of assessing R1, may enhance intra-therapist reliability, considered in this paper to more clinically relevant than inter-therapist reliability.

The aim of this study was to test the intra rater reliability of applying forces within the first, middle and last third of resistance relative to R2. Improving the accuracy and reliability of grades of movement may help therapists to determine treatment dose more accurately, and progress and regress

treatments more reliably in clinical practice. As very few studies have focused on the reliability of peripheral joint mobilizations, this study explored AP mobilization on the talus which have been shown to improve joint stiffness and dorsiflexion ROM following ankle sprain.[17,18]

The aim was also to compare therapists 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 28 chartered physiotherapists (with a minimum of 5 years experience), all of who were attending a postgraduate MSc module in neuromusculoskeletal physiotherapy. In addition, there were 4 lecturers with MSc qualification in Musculoskeletal Physiotherapy and members of the Musculoskeletal Association of Chartered Physiotherapists (MACP). Group 2 consisted of 19 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 therapists is illustrated in Table 1.

Table 1 about here

Three models were recruited for the study, with an average BMI of 23.98 (range 22.62-25.71) and over 18 years of age.

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 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.[19] A lean bar ensured that all the participants force was transferred to force plate.[19] The models ankles were strapped to a wooden platform to maintain plantar grade.

Procedure

For each assessment, the model lay on the same wooden plinth and the same researcher marked the talus with the foot resting in a neutral position. This process was repeated for each model during each data collection period.

Only 5 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, by applying repeated oscillatory mobilizations to fully explore the range and judge where point R2 occurred. A minimum of 3 oscillations to end range has been shown to adequately precondition the tissues,[20] 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 6s 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 finally asked to find R2 and the force applied at this point was recorded. 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 minutes 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.

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 2). R2 force was identified as the final maximum peak force.

Figure 2 about here

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[4] 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.

Data representing the mean peak forces for each third and R2 was tested for normality using Shapiro-Wilks analysis and in the event of abnormally distributed data, it was transformed using Log10. Transformed data was then reanalyzed to ensure normality. An Independent two-tailed T-Test was used to determine whether differences between the pre-registration and post-registration groups existed.

Reliability was established by calculating ICC and SEM values each third and R2. These were calculated on untransformed data, as the tests are based on ANOVAs which have been shown to be robust to deviations in data.[21] Groups were analyzed together where no group differences were found, and separately where group differences were evident. The minimal detectable change was calculated using the SEM to represent the 95% confidence level.

Results

There was no statistically significant difference between pre and post registration groups ($p<.05$), 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.

Table 2 about here

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

Table 3 about here

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, therapist forces would vary by 12.4N on 68% of occasions on intra-day tests and by 25.6N 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 therapists force applications on intra-day testing of between 25-34%. The percent differences on inter-day testing were larger between 38-61%.

Table 4 about here

Discussion

Excellent results were found for intra-rater reliability on intra-day testing with ICC of 0.96-0.97 which is better than those found by Snodgrass[9] 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 et al[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 et al[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.

The study by Harms[8] required participants 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 therapist will often apply several different joint mobilization doses within a physiotherapy session, commonly lasting 30-60 seconds, with a short break between doses.[3] It seems important that therapists 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 0.85-0.93. 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.[9] 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

spine, was not standardized. Snodgrass[9] 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 therapist factors might vary. Inter-day reliability therefore remains an important aspect of the therapist's treatment, and the good ICC results found in this study suggest that therapists 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 et al[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 this 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. No previous research has explored the effect of therapist qualifications and experience on intra-rater reliability when applying grades of mobilization, but this study suggests that therapists of all levels of experience are able to reliably replicate the 3 point grading system.

Limitations

In order to accurately record vertical forces, participants had to apply the mobilization in 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 to the last third and then to 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 therapist 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 at the end of the three grades of movement, 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 therapist 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 seconds, which might have biased towards better reliability than measuring over a longer time period.

In order to limit the cumulative effect of repeated joint mobilizations, 3 models were recruited rather than 1. By using 3 models, the varying weight,[15] ankle range of motion[15] and age[14] may have also resulted in variable forces between participants. For example, a model with a stiffer ankle might have resulted in larger forces being applied to this model compared with others. On inter-day testing, efforts were made to ensure the experimental conditions remained similar, as previously described, by using the same model and asking them to do a similar level of physical activity on both days, however it was not possible to control for small variations.

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 PA 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, was found to have excellent intra-rater reliability on inter-day and intra-day testing. Therapists of all levels of qualification and

experience were equally able to apply this 3 point grading system with excellent intra-rater reliability.

References

1. Maitland G. *Vertebral Manipulation*. 5th edition. Butterworths Heinemann. Oxford; 1986.

2. Margarey ME. Examination and assessment in spinal joint dysfunction. In: Grieve G (ed.) *Modern Manual Therapy of the Vertebral Column*. Churchill Livingstone, Edinburgh, ch 44, 481-497; 1984.

3. Maitland G. *Vertebral Manipulation*. 6th edition. Butterworths Heinemann. Oxford; 2001.

4. Lee R, Evans J. Towards a better understanding of spinal posteroanterior mobilizations. *Physiotherapy*. 1994;80: 68-73.

5. Latimer J, Goodsell MM, Lee M, Maher C, Wilkinson B, Moran C. Evaluation of a new device for measuring response to posteroanterior force in a patient population, Part 1: reliability testing. *Phys Ther*. 1996;76(2):158-165.

6. Shirley D, Ellis E, Lee M. The response of posteroanterior lumbar stiffness to repeated loading. *Manual Ther*. 2002;7(1):19-25.

7. Petty NJ, Maher C, Latimer J, Lee M. Manual examination of accessory movements-seeking R1. *Manual Ther*. 2002;7(1):39-43.

8. Harms MC, Bader DL. Variability of forces applied by experienced therapists during spinal mobilization. *Clinical biomechanics*. 1997;12(6):393-399.

9. Snodgrass SJ, Rivett DA, Robertson VJ, Stojanovski E. Forces applied to the cervical spine during posteroanterior mobilizations. *J Manipulative and Physiol Ther*. 2009;32(1):72-83.

10. Grieve GP. *Mobilisation of the spine*. 5th ed. Churchill Livingstone. Edinburgh; 1991.

11. Gorgos KS, Wasylyk NT, Van Lunen BL, Hoch MC. Inter-clinician and intra-clinician reliability of force application during joint mobilization: A systematic review. *Manual Ther*. 2013;19(2):90-96.

12. Mitchell WN. Reliability in the performance of grade II and grade IV mobilizations. Unpublished postgraduate diploma dissertation. Lincoln Institute of Health Sciences. School of Physiotherapy. Melbourne. 1983. Cited in Matayas B, Bach T. The reliability of selected techniques in clinical arthrometrics. *The Aust J Physio*. 1985;31(5):175-199.

13. Cook C, Turney L, Ramirez L, Miles A, Haas S, Karakostas T. Predictive factors in poor inter-rater reliability among physical therapists. *J Manual Manipulative Ther*. 2002;10(4):200-205.

14. Chiradejnant A, Latimer J, Maher C. Forces applied during manual therapy to patients with low back pain. *J Manipulative and Physiol Ther*. 2002;25(6):362-369.

15. Harms MC, Innes SM, Bader DL. Forces measured during spinal manipulation procedures in two age groups. *Rheumatology*. 1999;38:267-274.

16. Petty NJ, *Neuromusculoskeletal Examination and Assessment*: A Handbook for Therapists. 4th edition. Churchill Livingstone 2011.

17. Green T, Refshauge K, Crosbie J, Adams R. A randomized controlled trial of a passive accessory joint mobilization on acute ankle inversion sprains. *Phys Ther*. 2001;81(4):984-994.

18. Vincenzino B, Branjerdporn M, Teys P, Jordan K. Initial changes in posterior talar glide and dorsiflexion of the ankle after mobilization with movement in individuals with recurrent ankle sprain. *J Orthopaedic and Sports Phys Ther.* 2006;36(7):464-471.

19. Petty NJ, Messenger N. Can the force platform be used to measure the forces applied during a PA mobilization of the lumbar spine. *J Manual Manipulative Ther*. 1996;4(2):70-76.

20. Lee R, Evans J. Load-displacement-time characteristics of the spine under posteroanterior mobilization. *The Aust J Physio.* 1992;38:115-123.

21. Agresti A, Finlay B. *Statistical methods for the social sciences: With SPSS from A to Z: A brief step-by step manual.* 2nd edition. Pearson; 2009.

Tables

Table 1. Therapist participant demographic data

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

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) Post reg 79.2 (36)	91 (40)	81 (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 Post reg 19-187	24-195	19-200
	2nd third	44-239	47-241	27-259
	3rd third	77-311	69-309	60-316
	R2	78-322	79-314	67-313

(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 (.93-.98) Post reg 0.97 (.93-.99)	9.1 9.7	0.93 (.82-.97) 0.87 (.73-.94)	13.7 17.2
2nd third	0.96 (.93-.98)	12.3	0.85 (.74-.91)	21.9
3rd third	0.97 (.96-.99)	12	0.88 (.80-.93)	24.5
R2	0.97 (.96-.99)	12.4	0.88 (.79-.93)	25.6

ICC – Intraclass correlation coefficient (95% confidence interval)

SEM = standard error of measurement

Table 4,

Minimal detectable change (MDC) representing 95% confidence level. MDC expressed as a % of the initial force application (day1, test1)

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

Figure 1

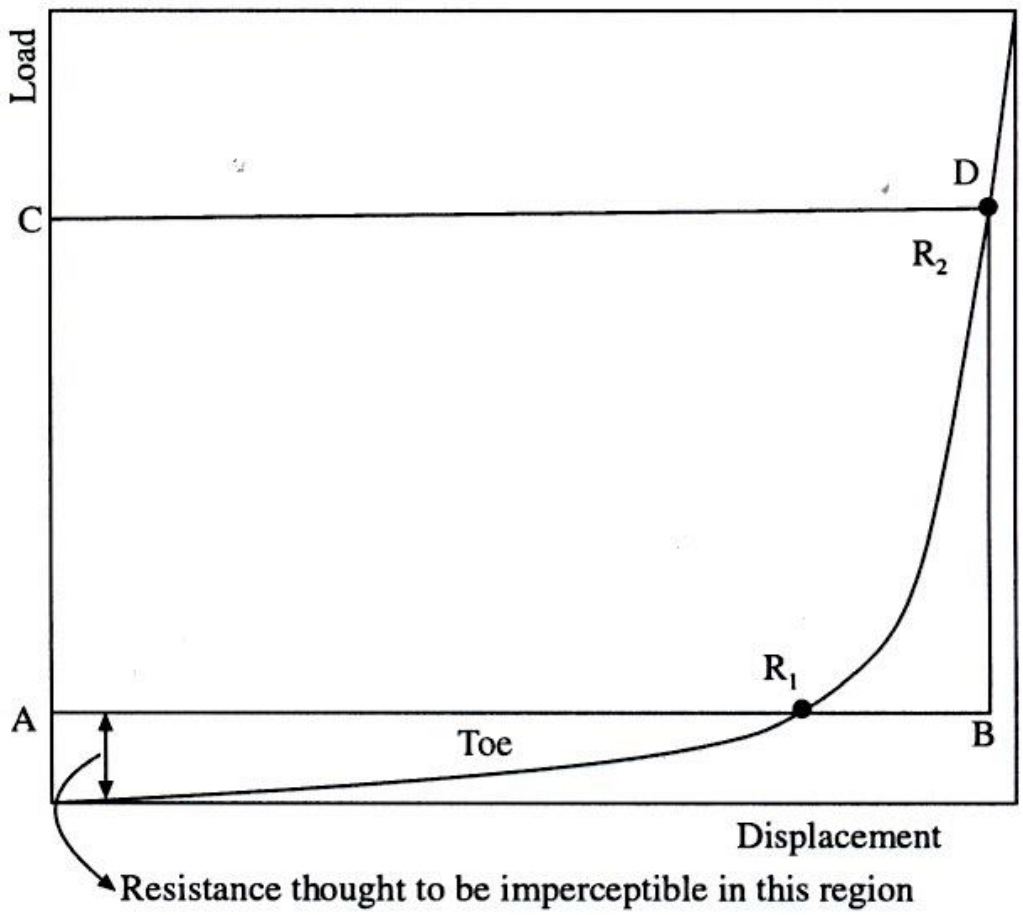

Figure 1

Relationship of movement diagram (ABCD) to load displacement curve (Reproduced with kind permission from Lee R, Evans J 1994 Towards a better understanding of spinal posteroanterior mobilizations Physiotherapy 80: 68-73)

Figure 2

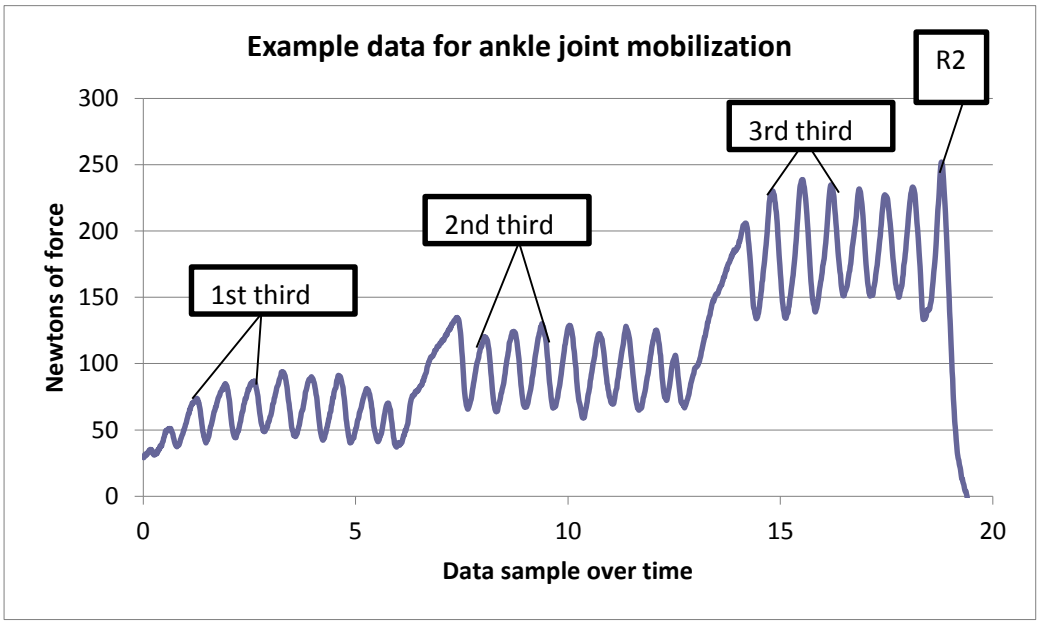

Figure 2

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