Comparison of Intervertebral ROM in Multi-Level Cadaveric Lumbar Spines Using Distinct Pure Moment Loading Approaches

Brandon Santoni, Andres F Cabezas, Daniel J Cook, Matthew S Yeager, James B Billys, Benjamin Whiting and Boyle C Cheng

*Int J Spine Surg* 2015, 9 ()
doi: https://doi.org/10.14444/2032
http://ijssurgery.com/content/9/32

This information is current as of January 25, 2020.

**Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at:
http://ijssurgery.com/alerts
Comparison of Intervertebral ROM in Multi-Level Cadaveric Lumbar Spines Using Distinct Pure Moment Loading Approaches

Brandon Santoni PhD,1 Andres F Cabezas ME,1 Daniel J Cook, MS,2 Matthew S Yeager, BS,2 James B Billys, BA,2 Benjamin Whiting, BA,2 Boye C Cheng PhD2
1Foundation for Orthopaedic Research & Education, Tampa, FL 2Allegheny Health Network, Department of Neurosurgery, Pittsburgh, PA

Abstract

Background
Pure-moment loading is the test method of choice for spinal implant evaluation. However, the apparatuses and boundary conditions employed by laboratories in performing spine flexibility testing vary. The purpose of this study was to quantify the differences, if they exist, in intervertebral range of motion (ROM) resulting from different pure-moment loading apparatuses used in two laboratories.

Methods
Twenty-four (laboratory A) and forty-two (laboratory B) intact L1-S1 specimens were loaded using pure moments (±7.5 Nm) in flexion-extension (FE), lateral bending (LB) and axial torsion (AT). At laboratory A, pure moments were applied using a system of cables, pulleys and suspended weights in 1.5 Nm increments. At laboratory B, specimens were loaded in a pneumatic biaxial test frame mounted with counteracting stepper-motor-driven biaxial gimbals. ROM was obtained in both labs using identical optoelectronic systems and compared.

Results
In FE, total L1-L5 ROM was similar, on average, between the two laboratories (lab A: 37.4˚ ± 9.1˚; lab B: 35.0˚ ± 8.9˚, p=0.289). Larger apparent differences, on average, were noted between labs in AT (lab A: 19.4˚ ± 7.3˚; lab B: 15.7˚ ± 7.1˚, p=0.074), and this finding was significant for combined right and left LB (lab A: 45.5˚ ± 11.4˚; lab B: 35.3˚ ± 8.5˚, p<0.001).

Conclusions
To our knowledge, this is the first study comparing ROM of multi-segment lumbar spines between laboratories utilizing different apparatuses. The results of this study show that intervertebral ROM in multi-segment lumbar spine constructs are markedly similar in FE loading. Differences in boundary conditions are likely the source of small and sometimes statistically significant differences between the two techniques in LB and AT ROM. The relative merits of each testing strategy with regard to the physiologic conditions that are to be simulated should be considered in the design of a study including LB and AT modes of loading. An understanding of these differences also serves as important information when comparing study results across different laboratories.

Keywords: lumbar spine, biomechanics, pure moment loading, range of motion, testing apparatus

Introduction
Since introduction almost three decades ago,1 pure moment load application has been widely adopted as the biomechanical test method of choice in quantifying the safety and efficacy of spinal implants and surgical techniques in cadaveric human2-5 and animal5-7 spine specimens. The general acceptance of this approach to spine evaluation theoretically enables existing and emerging devices to be compared across previous literature and between laboratories. Indeed, its acceptance as a standard testing protocol8,9 renders the general field of spine biomechanics unique, as no other sub-specialty field in orthopaedics has adopted a uniform laboratory test method with which to evaluate implants prior to clinical introduction.

In practice, pure moment load application to cadav-
eric spinal constructs has been accomplished with varying laboratory set-ups.\textsuperscript{4,10-15} Mechanistically, the pure moment is applied to the end vertebrae ensuring equal loading to all segments of the spinal column,\textsuperscript{1} even as the spine deforms during testing. This has been accomplished in multiple forms with varying complexity and ease of use, including cable and pulley systems,\textsuperscript{2,11,13} in which load application is assisted with static weights or load frame actuation, motor-actuated robotic arms,\textsuperscript{15-17} and hydraulically- or pneumatically-actuated gimbal systems\textsuperscript{10,12,18} that can be configured to apply pure moments at one or both ends of the spinal segment. Stepwise moment application using cable and pulley systems has been the predominant method of flexibility testing\textsuperscript{19,20} and is still prevalent in the current literature.\textsuperscript{14,21-23} More recently, continuous loading of cadaveric spine specimens has been introduced,\textsuperscript{10,12,15,18} representing a more physiologic laboratory technique with which to evaluate spine flexibility, with the major advantage that it permits the analysis of the spine’s behavior in the neutral zone.\textsuperscript{24}

Despite the perceived notion that the kinematic response of the test constructs under pure moments applied with the varying approaches are similar, few studies have quantified the differences, if they exist, in intervertebral range of motion (ROM) resulting from laboratory-specific loading apparatuses.\textsuperscript{17,24} Thus, the goal of our study was to compare the kinematic response of multi-segment lumbar spine constructs tested in two independent laboratories that differed based on the moment loading approach used to test the specimens (stepwise versus continuous) and the boundary conditions (i.e. the physical constraints at each end of the multi-segment constructs) imposed upon the specimens during kinematic evaluation. We hypothesized that, regardless of the approach used, multi-directional ROM findings would be comparable between the laboratories.

### Materials & Methods

**Specimen Preparation**

Twenty-four (n=24: m/15, f/9; avg. age: 60 yrs.; range: 30-70 yrs.; avg. height: 68 in.; range: 61-74 in.; avg. weight: 198 lbs.; range: 120-330 lbs.) and forty-two (n=42: m/23, f/19; avg. age: 58 yrs.; range: 24-74 yrs.; avg. height: 66 in.; range: 41-75 in.; avg. weight: 212 lbs.; range: 90-400 lbs.) intact fresh-frozen cadaveric T12/L1-S1 specimens tested at laboratories A (Foundation for Orthopaedic Research and Education, Tampa, FL) and B (Allegheny Health Network, Pittsburgh, PA), respectively, were included in this study. The aggregate total specimen numbers from each lab, and included for comparison here, were originally tested as part of multiple prior research studies involving spinal implants or surgical techniques. Thus, this study was a retrospective, post hoc analysis of existing data collected for prior investigations. In general, specimen preparation techniques were similar between groups. In both laboratories, the specimens were denuded of musculature and adipose tissue keeping the ligamentous and bony structures intact. They were wrapped in gauze, hydrated with 0.9% saline and frozen to -20°C in sealed plastic bags. For all specimens, the L1 and S1 vertebrae were potted in high strength resin (Bondo body filler, 3M Collision Repair Solutions, St. Paul, MN USA) using embedded fixation screws. At both laboratories, the S1 body was potted such that the L3-L4 disc was aligned parallel to the horizontal. Specimens were thawed at room temperature for 10-24 hours prior to biomechanical evaluation.

**Pure Moment Testing Protocol**

The intact specimens were tested under pure moment loading conditions in accordance with literature-based recommendations.\textsuperscript{8,9} At each lab, unconstrained pure moments of ±7.5 Nm were applied in flexion-extension (FE), right and left lateral bending (LB) and axial torsion (AT). The spines from each lab included for comparison were tested without application of a compressive follower load.

At laboratory A, non-destructive pure moments were applied to each specimen through a system of cables, pulleys, and suspended weights (Figure 1A-D). The inferior base (S1) of the specimen was mounted on a 6-component load cell (Model MC3A-1000, AMTI Transducers, Watertown, MA) and the specimen was allowed to move freely at the superior (L1) end. Using a 10in moment arm and slotted weights, moments were applied at 1.5 Nm increments up to a maximum of 7.5 Nm (based on visual feedback from the load cell’s readout) about the appropriate...
anatomical axes to induce six different motions: flexion, extension, left and right lateral bending, and left and right axial torsion. To account for the spine’s viscoelastic properties, for each loading scenario, three preconditioning cycles up to 7.5 Nm were applied to the specimen, and incrementally applied moments were maintained for approximately 30 seconds before recording ROM. Intervertebral ROM was obtained using an optoelectronic motion analysis system (Optotrak® Certus™, Northern Digital Inc., Waterloo, Ontario, Canada) with infrared light-emitting diode marker arrays rigidly coupled to each vertebral level. With the use of a digitizing probe, a local coordinate system for each vertebra was defined using three anatomical landmarks per body using the convention recommended by White and Panjabi with the (+)x direction to the left, (+)y directed superiorly and the (+)z directed anteriorly. The intervertebral ROM (degrees) was calculated as the range of the Euler angle corresponding to flexion-extension (X), axial rotation (Y) and lateral bending (Z). The Euler angle decomposition sequence was defined by the Optotrak system as $R_z(\phi) - R_y(\theta) - R_x(\psi)$. ROM values are reported as the dominant Euler angle during a given loading mode.

At laboratory B, each specimen was tested on a 6 degree-of-freedom custom-built test machine consisting of counteracting superior (T12/L1) and inferior (S1) flexion-extension and lateral bending stepper-motor-driven gimbals mounted on a biaxial and pneumatically-actuated Bose test frame (Smart Test Series, Eden Prairie, MN) (Figure 2). Torque cells, mounted in-line with each actuator within the gimbal set, were used for feedback and control for the FE and LB modes of loading. A six axis load cell integrated into the test machine was used for feedback and control during AT and for verification measurement of shear loads and bending moments. Specimens were subjected to a load control protocol during which 3 cycles of FE, LB, and AT moments were applied to a maximum of 7.5 Nm at a frequency of 0.005 Hz for three cycles with the third cycle used for ROM analysis. Similar to laboratory A, the ROM response of each intervertebral level was quantified with the Optotrak system with ROM calculated as the range of the Euler angle corresponding to FE, LB and AT loading modes.

**Statistical Analysis**

Data are reported as mean ± standard deviation. Differences in donor age, height, and weight between the two laboratories were compared with a student’s t-test. The ratio of male-to-female specimens at each laboratory was compared with a Fischer’s exact test. Level-wise ROM and total L1-L5 ROM measured at laboratory A and laboratory B were compared using the generalized linear modeling (GLM) function in
SPSS v22 (IBM, Corp, Armonk, NY). Specifically, a two-way ANCOVA was used to determine the effect of individual level and laboratory on intervertebral ROM. Age was included in the analysis as a covariate since ROM has been reported to vary based on age.\textsuperscript{7} Post-hoc multiple comparisons were performed with Bonferroni correction. Significance was set at the 0.05 level.

**Results**

The donor demographics of the specimens tested in both laboratories were similar. Specifically, no significant differences in donor age (p=0.105) or male-to-female ratio (p=0.221) of the specimens tested at the two facilities were identified. Further, there was no difference in donor height (p=0.114) or weight (p=0.404).

In flexion-extension and axial torsion loading, there was no significant difference in L1-L5 ROM between laboratories (p=0.289 and p=0.074, respectively) (Table 1). Significant differences in total right and left L1-L5 LB (laboratory A: 45.5˚ ± 11.4˚; laboratory B: 35.3˚ ± 8.5˚) ROM were identified between laboratories (p<0.001) (Table 1).

When ROM was compared between laboratories by individual intervertebral level, no significant differences were identified for any level in FE loading (p>0.170) (Figure 2). Pooling the ROM data between labs in FE, we found that L4-L5 displayed the greatest ROM on average (11.1˚ ± 3.3˚) and this was significant relative to the L1-L2 (7.3˚ ± 2.5˚, p<0.001), L2-L3 (8.5˚ ± 2.5˚, p<0.001), and L3-L4 (9.3˚ ± 2.6˚, p=0.003) levels.

Right and left LB measured in laboratory A was significantly higher at each lumbar level (p<0.011) compared to the ROM measured in laboratory B (Figure 3B). The largest difference between testing approaches was observed at the L4-L5 level, with ROM measuring 12.3 ± 3.8˚ at laboratory A and 8.8 ± 2.9˚ at laboratory B (p<0.001). Right and left AT measured in laboratory A was significantly greater (p<0.008) at the L1-L2 (4.5 ± 2.4˚) and L4-L5 (5.7 ± 2.9˚) levels compared to the torsion ROM measured in laboratory B (2.7 ± 1.3˚ and 4.1 ± 2.4˚, respectively) (Figure 3C). No significant difference in axial torsion ROM at the L2-L3 or L3-L4 levels was identified between laboratories (p>0.411).

**Discussion**

Laboratory evaluation of existing and emerging spinal devices is an important step in assigning acute efficacy to the design of a new potential clinical technology. Historically, this has been accomplished by comparing the intact ROM of a single- or multi-segment cadaveric specimen to its instrumented condition in three physiologic loading directions.\textsuperscript{8,9} Specimen flexibility is usually assessed with the application of pure moments,\textsuperscript{1} in which a force couple is applied to the cranial segment and the resultant intervertebral ROM is measured. As a number of test set-ups exist in which pure moments can be applied, the current study aimed to determine if the measured multi-axial ROM over multiple lumbar levels of intact spine specimens were similar using two different

<table>
<thead>
<tr>
<th>Loading Mode</th>
<th>Laboratory A (˚)</th>
<th>Laboratory B (˚)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1-L5 FE ROM</td>
<td>37.4 (9.1)</td>
<td>35.0 (8.9)</td>
<td>0.289</td>
</tr>
<tr>
<td>L1-L5 LB ROM</td>
<td>45.5 (11.4)</td>
<td>35.3 (8.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>L1-L5 AT ROM</td>
<td>19.4 (7.3)</td>
<td>15.7 (7.1)</td>
<td>0.074</td>
</tr>
</tbody>
</table>

Table 1. Total L1-L5 intervertebral ROM (˚) comparisons between laboratories. Mean (SD).

Fig. 3. Intervertebral ROM measured at each level at each lab in FE (A), LB (B), and AT (C).

Downloaded from http://ijssurgery.com/ by guest on January 25, 2020
testing approaches in two independent laboratories. The primary differences between the laboratories include: (1) the boundary conditions of the multi-segment construct exposed to pure moment load application; and (2) the use of feedback-controlled stepper-motor driven biaxial gimbals that applied continuous pure moments versus a cable-driven pure moment system with stepwise incremental unidirectional loading assisted with hanging weights. An understanding of any apparatus-specific differences, should they exist, would serve as important information when comparing kinematic results across different laboratories.

The major findings of this study demonstrate remarkably similar FE ROM between two very different testing methods. Statistically significant differences of relatively small magnitude (2-3°) were found at two intervertebral levels in AT and at all levels in LB. These differences may best be explained by the differing boundary conditions that exist between the two test configurations in LB and AT. In laboratory A during AT, pure moments were applied in the plane of the top surface of the superior potting fixture throughout the range of movement. In laboratory B, AT moments were applied by the torsion actuator of the test frame, which maintains a fixed relationship relative to the laboratory floor and is independent of spinal posture. In laboratory A, LB moments were applied about an axis fixed to the superior potting fixture of the specimen. In laboratory B, LB moments were applied by independently driven motors fixed respectively to the upper and lower potting fixtures. Non-collinearity of these applied moments may result in off axis loading which cannot be mitigated by the test apparatus. In summary, the design of the test apparatus in laboratory B applies LB and AT moments with the spine situated in a predominantly upright posture, while that in laboratory A proscribes no fixed relationship between posture and applied loads. While the state of knowledge with regard to spine biomechanics is insufficient to state with certainty the relative merits of these two testing strategies, data has now been presented that demonstrates the magnitude of differences that may be encountered between them.

Despite the multitude of laboratory means to apply pure moments, only a few studies have been designed to elucidate the effect of the subtle technical differences used by different groups on the displacement response of the spine under well-defined loading conditions. Goertzen et al. evaluated six porcine cervical spine specimens (C2-C4) on a custom-designed testing apparatus consisting of a motor-controlled articulating arm. In their study, pure moments were applied to each spine in random fashion to the C2 body under the following conditions: (1) stepwise loading to ± 2 Nm (FE, LB, AT) in four equal load steps with a 30 s pause at each target moment; and (2) continuous loading to ± 2 Nm in all loading modes at 0.25˚/s. They noted significant differences in C2-C4 ROM between the two loading conditions, with the continuous loading protocol resulting in significantly lower ROM than stepwise loading in FE (23.9˚ vs. 26.1˚), LB (35.2˚ vs. 37.9˚) and AT (8.5˚ vs. 9.7˚), a finding they attributed to the viscoelastic creep effect present during stepwise loading. The creep response was likely present in the current study; however, because FE ROM between the two labs was remarkably similar, we consider this difference between testing approaches to have had less of an impact on LB and AT ROM than the differences in boundary conditions imposed on the specimens that we describe above in detail. In a more recent study, Wheeler and colleagues compared the ROM at ± 7.5 Nm of seven intact lumbar spine cadaveric motion segments (n=6 L1-L2, n=1 L4-L5) tested at four independent laboratories, two of which used stepwise loading protocols while the others applied pure moments continuously with servohydraulically-actuated load fixtures. In contrast to our findings and those of Goertzen et al., they identified no significant differences in the measured ROM in any loading mode between the laboratories. Noteworthy differences between our study and that of Wheeler et al. include our use of multi-segment constructs, a dual gimbal system that applied continuous moments, and a stepwise loading approach that differs from the sliding-ring cable-pulley moment application approaches used in their study.

For the flexibility tests performed as part of this study, no compressive follower load was utilized, as a primary consideration was the utilization of similar...
test conditions between the two laboratories without introducing potentially confounding factors. The follower load, as initially introduced Patwardhan et al.,\textsuperscript{28} has been used by numerous authors in the evaluation of spinal devices since its incorporation into the laboratory setting is thought to more accurately replicate the compressive loads borne by the spine due to muscle activity in vivo.\textsuperscript{6,28,29} Despite its physiologic merit, it does suffer from a number of limitations. Due to technical difficulties, its use is currently restricted to flexion-extension loading only. Secondly, deviations of the compressive load’s line of action from an optimized path as it runs tangent to the lordotic curvature of the spine can introduce offset moments and confound biomechanical outcomes,\textsuperscript{30,31} further clouding the ability to make meaningful comparisons between studies in which spinal implants are evaluated under pure moment loading conditions.

Our study suffers from certain limitations. The lumbar segments tested at laboratory A and B were from different donors. Differences in male-to-female ratio, age and differing levels of degeneration may contribute to differences in ROM findings between laboratories, independent of the test apparatuses used. While we did not grade the degree of disc degeneration in the tested specimens, donor demographics between laboratories were not statistically different regarding age (p=0.105), male-to-female ratio (p=0.221), height (p=0.114) or weight (p=0.404). Future studies in which the same multi-segment lumbar specimens are tested with these or other pure moment loading approaches may help to substantiate the findings reported herein. Secondly, the singular biomechanical outcome that we report comparing the two testing apparatuses was intervertebral ROM. Additional measures pertinent to fully understanding the biomechanical behavior of the spine under moment loading conditions may include neutral zone, stiffness and intradiscal pressure were not reported here, most notably because this study was not designed prospectively, but instead undertaken using historical data already collected as part of previous studies done in both laboratories in which intervertebral ROM was the primary study metric.

**Conclusion**

The results of this study show that intervertebral ROM in multi-segment lumbar spine constructs is remarkably similar in FE between two fundamentally different testing methods. Differences in boundary conditions are likely the source of small, statistically significant differences at some intervertebral levels between the two techniques in LB and AT ROM. The characteristics of loading applied by exogenous forces and by the musculature to the lumbar spine during any specific activity, while an emerging field of study, are not empirically verified. Thus, the relative merits of each testing strategy with regard to the physiologic conditions that are to be simulated should be considered in the design of a study including LB and AT modes of loading. An understanding of these differences also serves as important information when comparing study results across different laboratories.

**References**


Disclosures
The authors declare no relevant disclosures.