ELSEVIER

Clinical Biomechanics 14 (1999) 505–514

CLINICAL
BIOMECHANICS

www.elsevier.com/locate/clinbiomech

Variability in spine loading model performance ☆

W.S. Marras [a,*], K.P. Granta [b], K.G. Davis [a]

[a] Biodynamics Laboratory, The Ohio State University, 1971 Neil Ave, Columbus, OH 43210, USA

[b] Motion Analysis and Motor Performance Laboratory, University of Virginia, 2270 Ivy Road, Charlottesville, VA 22903, USA

Received 15 April 1998; accepted 30 September 1998

Abstract

Objective. To assess the sources of variability associated with an EMG-assisted model of spine loading.

Design. In vivo measurements of trunk dynamics, lifting moments and muscle activities were used as inputs into an EMG-assisted spine loading model.

Background. Several types of variability are inherent in biomechanical assessments of risk associated with trunk bending motions during lifting. Variability may occur as a function of variations in spine loading due to either subject variations in motion profiles (kinematics) or biomechanical model performance.

Methods. Twelve experienced and inexperienced materials handlers performed 10 repeated lifts where load weight, asymmetry, and velocity were varied. The experiment was replicated on a second day to assess day to day variability.

Results. These model performance variables indicated that variability was mainly a function of subject characteristics and experience. Minor variations in variability were associated with the task asymmetry and weight lifted. Advanced analyses suggested that experienced workers had a greater range of back motion compared to inexperienced workers which would affect the length–strength component of the model calibration.

Conclusions. This study indicates that for the results of an EMG-assisted model to be accurate, it is important to ensure that the model reflects a realistic relationship between the trunk muscle length and the muscle force production capacity. Underestimation if this relationship can degrade model fidelity and robustness.

Relevance

These results imply that by properly calibrating the model it is then reasonable to assume that the vast majority of variations observed in repeated exertions of a particular trial are due to kinematic and kinetic differences inherent in the muscle control system and not a function of model randomness.

Keywords: Spine loads; Biomechanical modeling; Low back disorders; Variability

1. Introduction

Variability is a natural and expected component of human performance. All statistical assessment techniques recognize and attempt to distinguish the variability associated with the factor or factors of interest from that expected due to natural variability. It has been well established that humans behave according to a Gaussian or normal distribution. Thus, repeated performance of a particular task would be expected to result in a given range of variability.

From a biomechanical standpoint, one can compare the level of loading imposed upon a structure vs the inherent level of tolerance associated with a structure in order to get an idea of the degree of risk associated with a particular task. It is believed that these loading and tolerance levels also behave according to a probabilistic function. For example, Mirka and Marras [1] demonstrated that one could quantitatively describe the range of spinal loads expected, given task parameters, based upon the variability in muscle activities observed during bending motions. Jager et al. [2] have also described the variability in load tolerance to compressive load expected in the spine. These researchers have shown that

☆ This paper is a companion to Granata KP, Marras WS, Davis KG. Variation in spinal load and trunk dynamics during repeated lifting exertions. Clin Biomech 1999; 14: 367–375, where it appears as citation #3 with an incomplete reference.

* Corresponding author. E-mail: wsm@osuergo.eng.ohio-state.edu

0268-0033/99/$ - see front matter
PII: S 0 2 6 8 - 0 0 3 3 (9 8) 0 0 0 9 0 - 4

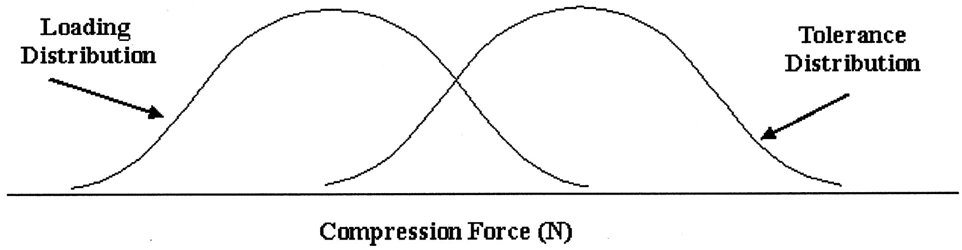

Fig. 1. Hypothetical distributions of spinal load and tolerance for compression force.

load tolerance can vary from about 2000 N to over 8000 N. Others have found the compressive tolerance to be even higher (ranging from 3700 to 13000-13000 N) [3]. Collectively, these loads and tolerance distributions can be viewed relative to one another (Fig. 1) in terms of a signal detection-type model to define the overall risk of suffering a micro fracture of the vertebral end plate. This is believed to lead to cumulative trauma of the lower back [4].

This model of low back disorder risk description is obviously dependent upon our ability to accurately characterize the load and tolerance distributions. Since there are currently no feasible means to characterize load or tolerance in vivo, we must rely on spine models to predict load and in vitro experimentation to assess tolerance. The load distribution model, human performance characteristics, and the load tolerance estimates all have inherent variability associated with them. Yet, there is a paucity of literature addressing the issue of the extent of the variability expected from the human vs the variability associated with the assessment technique. Thus, the goal of this study was to assess variability associated with a model currently used to assess spinal loads. A companion paper [5] addresses the issue associated with subject kinematic variability. Thus, collectively these two studies are offered in an effort to facilitate our understanding of how low back disorder risk might be associated with a particular task.

2. Methods

2.1. Subjects

Twelve healthy males with no prior history of low back disorder (LBD) volunteered to participate in this study. The subject population included seven college students and five experienced warehouse selectors from a local distribution center. The subjects ages ranged from 22 to 34 yr with and average age of 26.1 yr. The average (SD) stature of the subject population was 179.29(4.5) cm, and the average weight was 74.7(7.0) kg.

2.2. Task

Subjects were required to lift weighted boxes under various trunk velocity and asymmetry conditions from knee height to an upright posture. Lifting exertions were performed one at a time, with one minute of rest between exertions to minimize the possibility of fatigue.

2.3. Experimental design

Independent variables consisted of two box weights (13.6, 27.3 kg), two levels of task asymmetry (sagitally symmetric, 60° right), and two subjective lifting velocities (preferred lifting velocity, faster than preferred). Asymmetric tasks were achieved by requiring the subject to lift the weighted box from a knee height platform located 60° to the subject's right. The weight and asymmetry conditions were chosen to represent the typical range of conditions observed among low and high risk (of LBD) industrial jobs [6,7]. The subjective velocity levels of 'preferred' and 'faster than preferred' lifting styles were chosen to permit examination of two lifting velocities without artificially influencing the natural motion variability. Each condition was repeated 10 times. The trials were randomized with respect to weight, asymmetry, lifting velocity and repetition number.

Electromyographic (EMG) data were collected from bipolar surface electrodes over the right and left erector spinae, rectus abdominus, latissimus dorsi, external abdominal obliques and internal abdominal obliques as described by Marras and Mirka [8]. Myoelectric data were low pass filtered at 1 kHz, high pass filtered at 30 Hz, notch filtered at about 60 Hz, rectified, averaged using a 20 ms sliding window filter then normalized relative to values collected during maximum voluntary contraction (MVC) exertions. Maximum EMG values were collected during static flexion, extension, right twist, left twist, right lateral, and left lateral MVC exertions performed against a reference frame in an upright posture.

Trunk motion data were recorded from an electrogoniometer designed to measure sagittal, lateral and twisting motions of the lumbar region of the trunk.

Dynamic external loads were determined from a force plate (Bertec 4060A, Worthington, OH USA) at the subject's feet. An electro-mechanical vector monitor was employed to record the location of the lumbo-sacral junction relative to the center of the force plate [9]. External forces and moments applied to the lumbo-sacral junction of the spine were computed from the force plate data and the hip location and orientation kinematic data using the method of Fathallah et al. [9].

An EMG-assisted biomechanical model employed the EMG, kinetic and kinematic data as input to compute the dynamic loads on the spine [10–15]. The model incorporated the normalized muscle activities, dynamic trunk motion, and external loads to determine the contractile forces of the 10 co-contracting muscles. Spinal compression, lateral shear, and anterior–posterior (AP) shear forces were computed from the vector sum of the muscle forces. Thus, three-dimensional dynamic spinal loads were determined for each lifting exertion. The data collection methods, biomechanical model structure, and validation have been published previously [8,10–17].

Three measures of model performance were used to assess model sensitivity to factors that might contribute to model variability. These consist of gain, R^2, and average absolute error (AAE). Gain consists of the maximum force production capacity of the muscles represented in the model. This is a subject specific factor that is determined through model calibration. It represents the maximum force per unit area of muscle that is estimated for the model to perform well. Gain is computed by comparing muscle-generated, trunk moments with measured, applied moments about the lumbosacral junction. To satisfy the equations of dynamic equilibrium, muscle generated extension moment must equal the measured moment. Gain is appropriately and automatically adjusted to satisfy this condition. To be physiologically valid, the predicted gain level must fall within the range 30–100 N/cm^2 [18–20]. Muscle force per unit area is highly variable between subjects, based on subject conditioning and natural ability. On the other hand, gain predicted for a given subject must be constant throughout each of the experimental trials. Examination of whether the gain value is within the physiologic range and within subject variability provides one test of model validity.

The R^2 measure monitors the ability of the model to predict trunk moments relative to the moments that are measured by the experimental apparatus. Trunk moments are computed from dynamic, muscle, force vectors, and anthropometric moment arms. Three-dimensional moments predicted from the EMG data are displayed as a function of time and superimposed upon the trunk moments measured by the force plate. This provides a robust measure of dynamic model accuracy by comparing the predicted and measured moment profiles and quantitatively determined by means of statistical correlation value (R^2). A high R^2 value indicates the free-dynamic model which accounts for the variability in the lifting moment which implies the model generates and accurate simulation of spinal load during the lifting exertions. Since the mechanics used to estimate the external moment are directly related to the mechanics used to calculate spine loading, it is assumed that the R^2 statistic is also an indirect indicator of how well the model predicts spine loading. Thus, the R^2 measure indicates how well the measured dynamic trunk moment 'trend' matches the predicted trunk moment 'trend'. However, this measure provides no information about how well the magnitude of the predicted and measured moments match.

In order to provide a measure of magnitude error between the measured and predicted trunk moment the AAE is examined. This measure predicts how well the average magnitude of the measured trunk moment compared to the average magnitude of the predicted trunk moment through the exertion of interest. As with the R^2 statistic, the AAE is believed to indirectly indicate the magnitude of error in predicting spinal load.

The objective of the current study was to determine how these three measures of model performance varied as a function of experimental parameters. Thus, instead of examining how the spinal loads changed as a function of the experimental conditions, we were interested in how the measures of model performance would be affected by the experimental conditions. Ideally, the model performance measures should not respond in a statistically significant manner to any of the experimental conditions, indicating that all variability in performance would be a function of the experimental conditions.

2.4. Statistical analyses

Intra-class correlations (ICCs) were performed to identify the independent parameters that influenced the variability of the modeled data [21]. Repeated measures analysis of variance (ANOVA) statistical analyses were performed to augment the ICC results and for comparison with previous research. ANOVA were also performed on within subject variability measures to identify factors that influence distribution widths. For all significant independent variables, post hoc anlayses in the form of Tukey multiple pairwise comparisons were performed to determine the source of the significant effect(s). An alpha level of 0.05 for all statistical tests was selected.

3. Results

A summary of the factors that resulted in statistically significant differences in model gain, R^2, and AAE as a function of the data collected on the same day and on different days are shown in Table 1. For repeated trials performed on the same day, task asymmetry and weight by experience interaction affected gain. Table 2 reports the descriptive statistics associated with this portion of the experiment. Table 2 indicates that the asymmetric lifts resulted in muscle gain estimates that were about 7–5 N/cm^2 lower than those predicted for the sagitally symmetric exertions. However, both of these gain estimates were well within the range of physiologic feasibility. The weight by experience interaction was tracked to the experienced subjects only. Inexperienced subjects exhibited the same gain values regardless of the weight of lift. Experienced subjects exhibited gain values that were 2–6 N/cm^2 greater than the inexperienced group with the heavier load relating the greater gain value prediction. Other than these two effects, the model was very stable and no other factors significantly influenced gain estimation.

The R^2 statistic for repeated exertions of trials performed on the same day was affected by the weight of the object lifted, subject experience, the trial by asymmetry interaction, the lift velocity by experience interaction and the trial by experience interaction. Table 2 indicates that although significant, the effect of changing the magnitude of the weight lifted upon the R^2 was fairly minimal. The R^2 increased by 0.04 when the weight increased by 13.6 kg. However, it should be pointed out that both values of R^2 associated with the weight magnitude were excellent. Table 2 indicates that worker experience had a significant and fairly substantial effect on average R^2. The average R^2 value observed for experienced subjects was 0.1 lower than that for inexperienced subjects. The interaction of lift velocity and weight with experience was statistically significant, yet biomechanically trivial, as was the interaction of trial and asymmetry.

Table 1
Statistical significance results from ANOVA for the model performance variables for the same day exertions (within day). Figures in bold indicate significance at $\alpha = 0.05$)

Variable	Gain	R^2	Average absolute error (AAE)
Weight (W)	0.06	**0.006**	**0.05**
Lift velocity (L)	0.76	0.44	0.28
Asymmetry (A)	**0.01**	0.49	0.44
Experience (E)	0.56	**0.0004**	**0.02**
Trial (T)	0.23	0.14	0.19
Weight*lift velocity	0.14	0.55	0.63
Lift velocity*asymmetry	0.35	0.66	**0.03**
Weight*asymmetry	0.12	0.82	0.18
Experience*asymmetry	0.42	0.36	0.38
Experience*lift velocity	0.83	**0.05**	0.29
Experience*weight	**0.005**	0.84	0.09
Weight*trial	0.30	0.82	0.51
Lift velocity*trial	0.46	0.37	0.59
Asymmetry*trial	0.12	**0.02**	0.11
Experience*trial	0.42	**0.002**	**0.002**

Table 2
The mean (SD) for the model performance variables as a function of the independent variable for the same day exertions (within day)

	Gain		R^2		Absolute average error (AAE)	
	Mean	SD	Mean	SD	Mean	SD
Experience level						
Inexperience	39.39	15.67	0.895	0.074	17.24	8.32
Experience	44.20	13.03	0.794	0.142	25.69	14.29
Lifting velocity						
Preferred	41.5	15.89	0.867	0.119	19.65	13.38
Faster than preferred	40.37	14.08	0.857	0.103	20.27	8.71
Asymmetry						
0° (symmetric)	44.69	16.32	0.856	0.103	19.41	7.51
60° (right of origin)	37.16	12.54	0.869	0.119	20.51	14.14
Weight						
30 lb	40.32	14.48	0.841	0.117	18.68	11.67
60 lb	41.56	15.56	0.885	0.109	21.24	10.83
Trial						
1	41.77	13.97	0.841	0.120	22.42	10.01
2	41.16	14.73	0.855	0.123	20.39	10.39
3	40.53	13.71	0.860	0.097	19.39	8.41
4	39.88	14.28	0.857	0.108	20.28	10.41
5	42.43	22.20	0.854	0.134	19.96	13.63
6	41.11	14.31	0.876	0.086	19.74	10.37
7	39.96	13.72	0.878	0.083	18.38	7.22
8	41.09	15.24	0.868	0.092	19.24	8.11
9	41.57	12.78	0.871	0.127	21.46	20.34
10	39.92	13.21	0.865	0.133	18.31	7.64

Table 3
Statistical significance results from ANOVA for the model performance variables for the different day exertions (between day). Figures in bold indicate significance at $\alpha = 0.05$

Variable	Gain	R^2	Average absolute error (AAE)
Weight (W)	0.75	0.16	**0.008**
Day (D)	0.49	0.48	0.27
Asymmetry (A)	**0.005**	0.21	0.23
Experience (E)	0.16	**0.002**	**0.01**
Trial (T)	0.20	0.96	0.84
Weight∗day	0.33	0.79	0.68
Day∗asymmetry	0.49	0.99	0.57
Weight∗asymmetry	0.34	0.45	0.22
Experience∗asymmetry	0.76	0.43	0.27
Experience∗day	0.22	0.54	0.21
Experience∗weight	0.34	**0.05**	**0.004**
Weight∗trial	0.73	0.19	0.40
Day∗trial	0.37	0.84	**0.05**
Asymmetry∗trial	0.54	0.61	0.43
Experience∗trial	0.57	0.22	0.76

Statistically significant differences in AAE were also observed for the magnitude of the weight lifted and subject experience, as well as the lift velocity by asymmetry interaction and the trial by experience interaction (Table 1). Table 2 indicated that AAE was lower for the lighter weight by about 2.5 N. This was not unexpected since the weight values are lower for the lighter weight. Therefore, relative error might not indicate a significant difference. This difference, although statistically significant, is biomechanically trivial. Differences as a function of experience were more pronounced. AAE was about 8–1/2 Nm lower for the inexperienced subjects compared to the experienced subjects. The standard deviation associated with the inexperienced subject performance was also about half of that for the experienced subjects. Thus, experience did play a substantial role in model fidelity relative to absolute error. The lift velocity by asymmetry interaction indicated that lifts performed at faster than preferred velocities yielded about the same AAE regardless of the asymmetry of the lift. However, under preferred lift velocity conditions the 60° asymmetric condition resulted in slightly greater AAE (about 5 Nm). The trial by experience interaction indicated that AAE was also rather stable for inexperienced subjects among trials, whereas the AAE was greater and much more variable for experienced subjects.

The statistically significant differences observed as a function of days are summarized in Table 3. In general, fewer variability trends (significant differences) were observed between days compared to variability within the same day (Table 1). Similar trends relative to asymmetry were observed for this analysis as were observed for the within day variability analysis. Here again, a statistically greater gain was observed for the sagitally symmetric conditions compared to the asymmetric conditions. Table 4 shows the descriptive statistics associated with model performance between trial

Table 4
The mean (SD) for the model performance variables as a function of the independent variable for the different day exertions (between day)

	Gain		R^2		Absolute average error (AAE)	
	Mean	SD	Mean	SD	Mean	SD
Experience level						
Inexperience	39.06	14.74	0.897	0.089	17.03	9.41
Experience	51.77	20.08	0.795	0.135	30.68	18.57
Day						
1	41.50	15.89	0.867	0.119	19.66	13.38
2	44.76	19.27	0.864	0.111	23.15	15.32
Asymmetry						
0° (symmetric)	47.01	18.73	0.859	0.106	19.57	10.02
60° (right of origin)	38.93	15.30	0.873	0.124	22.90	17.57
Weight						
30 lb	43.13	17.39	0.846	0.118	19.67	14.11
60 lb	42.81	17.75	0.886	0.109	22.81	14.51
Trial						
1	42.00	15.17	0.863	0.121	21.64	12.04
2	42.48	15.27	0.858	0.133	21.38	12.38
3	41.47	14.83	0.853	0.114	20.63	11.76
4	42.46	17.53	0.854	0.122	20.61	12.18
5	45.78	24.99	0.869	0.134	20.83	16.48
6	43.48	17.37	0.874	0.087	22.27	13.80
7	42.63	16.35	0.872	0.103	20.47	13.54
8	44.00	19.92	0.869	0.103	20.47	13.35
9	43.39	16.40	0.866	0.134	23.52	22.39
10	42.04	15.92	0.881	0.096	20.55	13.11

days. Table 4 shows that the gain was about 8 N/cm^2 greater for the symmetric conditions compared to the asymmetric conditions. This difference in magnitude was similar to that observed for the within day analysis which indicates a level that was well within the range of physiological feasibility. No other factors affected gain estimation.

The R^2 statistic was affected by experience and the weight by experience interaction. Table 4 indicates similar trends as observed for the within day analysis relative to the effects of experience. Here again, when the data for inexperienced subjects was modeled, the R^2 statistic was greater by an average of about 0.1 than for the experienced subjects. In addition, the inexperienced subjects yielded a lower standard deviation than the experienced subjects. It is also interesting to note that Table 4 indicates similar trends observed for the within day analysis relative to the effects of experience. When the data for the inexperienced subjects was modeled, the R^2 statistics for experience were just about the same for within day trial compared to between day trials. The interaction of experience and weight indicated that the R^2 values were about equivalent for experienced subjects regardless of the weight lifted, whereas, the inexperienced subjects exhibited slightly better R^2 statistic for the heavier weight.

Four factors significantly affected AAE when the data were evaluated for variability between days. AAE was slightly different as a function of weight lifted with values slightly higher for heavier weights (Table 4). The values were fairly similar to those for the within day data. The slight difference noted is also expected given the magnitude differences for the within day analyses. Since the magnitude of this difference is of the order of 3 N, this difference, although statistically significant, is not expected to be biomechanically meaningful. Differences in AAE as a function of experience between days were more substantial. Experienced subjects exhibited far greater AAE than inexperienced subjects between test days. The AAE for inexperienced subjects was similar to that for the within day analysis. However, the AAE was about 5 Nm greater, on average, for the experienced subjects when analyzed between days compared to within days. Variability as evidenced by the standard deviation also increased for experienced subjects between days (Table 4). The experience by weight interaction was significant for AAE between days indicating that the AAE for inexperienced subjects was fairly consistent and low (less than 20 Nm) for both weights, whereas, the AAE was greater (about 28 Nm) for the 13.6 kg load and became even greater (about 34 Nm) when the 27.3 kg load was lifted. The day by trial interaction was also significant for AAE when analyzed between days with differences in AAE being minimal as a function of early trials and greater AAE values occurring on the second testing day for latter trials.

ICCs were used to evaluate the source of variability in model performance for these results. ICCs indicate the source of variability over the experimental conditions. Figs. 2 and 3 show the sources of variability for gain, R^2, and AAE for the within day analyses and the between day analyses, respectively. Note that these results are not identical to statistical significance analyses (Tables 1 and 3). This is expected in that ICCs indicate the overall

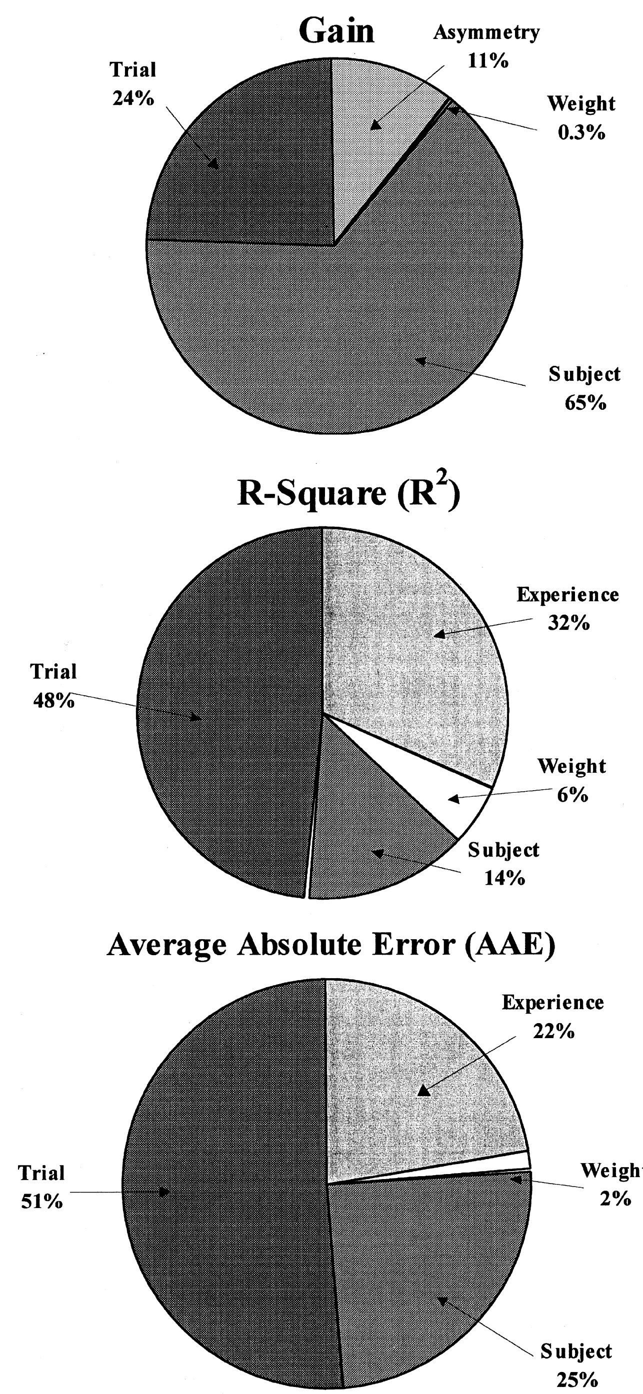

Fig. 2. The variability explained by each of the corresponding independent variables within days.

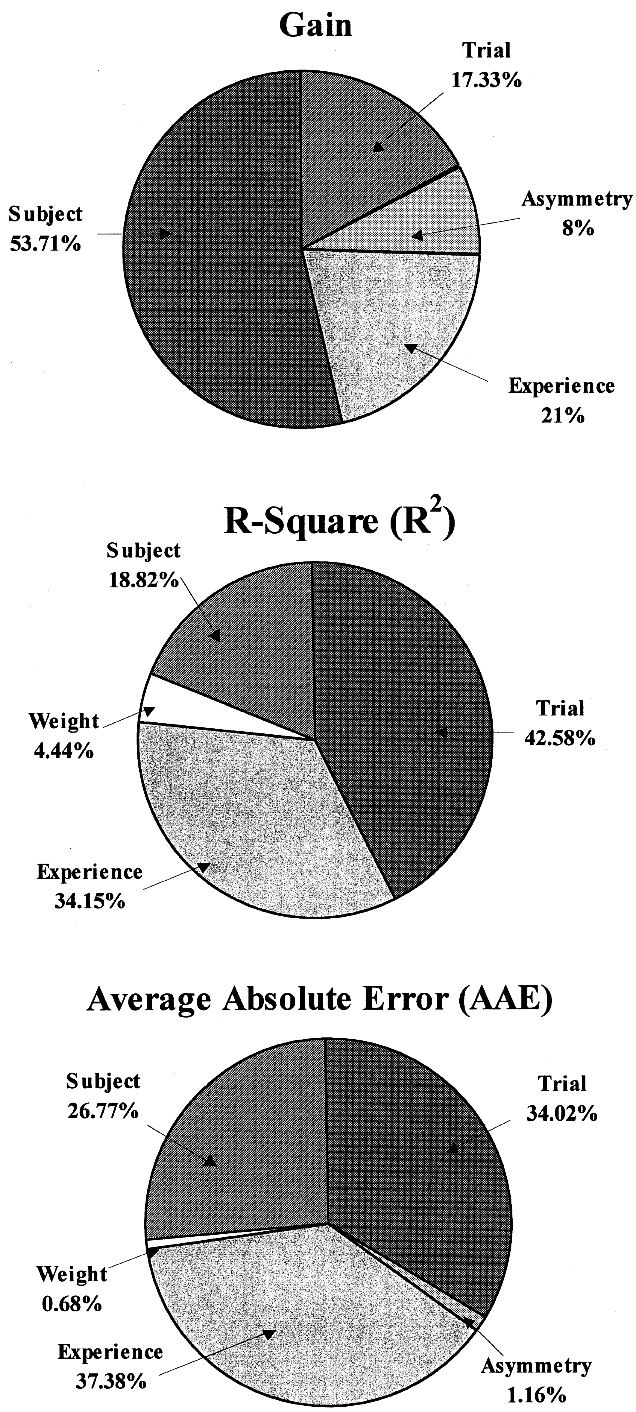

Fig. 3. The variability explained by each of the corresponding independent variables between days.

variability associated with a variable, whereas, the ANOVA would indicate how the variability distributions would change as a function of the experimental variables. Both figures indicate that for both within day and between day analyses the greatest portion of variability in gain is associated with specific subject properties. Negligible variability is associated with weight lifted and velocity of lift within days, and is associated with weight lifted, velocity and the day of the experiment. In both analyses 54–65% of the variability is associated with specific subject characteristics; 17–25% of the variability is associated with the specific trial over the two analyses, and 8–11% of the variability is attributable to changes in asymmetry. Between day analyses also indicated that 20.6% of the variability in gain was associated with experience.

Both within day and between day analyses indicate that the R^2 performance was influenced primarily by trial (42–49% of variability) and experience (31–34% of variability). R^2 was affected to a lesser extent by the subject characteristics (13–19% of variability) and by the weight lifted (4–6% of variability).

Finally, the AAE ICCs were fairly similar to those for the R^2 statistic. Trial and experience explained most of the variability and subject characteristics explained about 25% of the variability. However, weight magnitude explained less than 2% of the variability in both the within day and between day analyses.

4. Discussion

Observing how the independent variables of a biomechanical lifting study can affect model performance measures, in an experiment like this, can yield valuable information about model stability and help one separate biomechanical variability from inherent model variability. Collectively, these results have shown that the model was, for the most part, very stable and repeatable when applied to lifting exertions within the same testing period as well as on different testing days. Both the between day data as well as the within day data have shown that the largest source of variability (via the ICCs) when estimating subject gain is inherent to subject characteristics. It is not unexpected, given that it is a reasonable assumption that in an EMG-assisted model which calibrates the model parameters to each subject, that every subject should have a slightly different muscle gain given differences in fiber composition and training levels. Thus, the large amount of variability associated with each subject serves to reassure one that the model is behaving as expected, and does not adversely affect model stability or robustness. The R^2 statistic and AAE were affected to a lesser extent by individual subject characteristics. Variability associated with these model parameters was most likely due to individual variation in muscle cross-sectional area estimates and associated with muscle lines of action between subjects. This variability would be inherent in applying any type of population based relationship to a given individual.

Trial and asymmetric variability in gain estimation speak more to the issue of model stability. Trial variability in gain estimation for both the between day trials and within day trials are shown in Fig. 4. As seen in

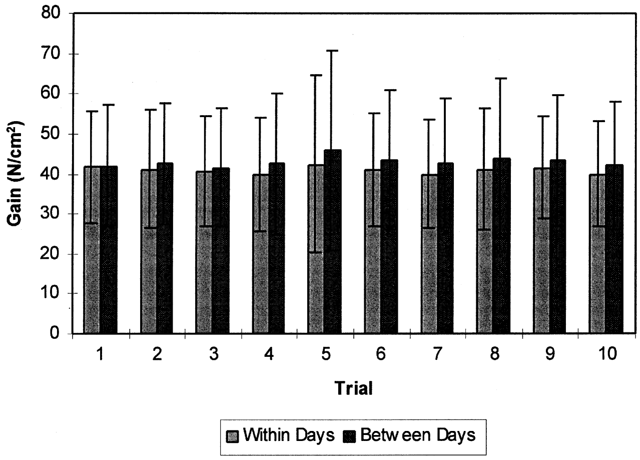

Fig. 4. Muscle gain as a function of trial for the within day and between day exertions.

Fig. 4 even though 17–25% of the variability in gain estimation is associated with the trial the magnitude of this difference is small and at most 2 N/cm^2. Neither analysis indicated that this difference was statistically significant. Variability in gain estimation associated with asymmetry point to more inherent model limitations. Of the ICCs associated with gain, this source of variability was the smallest. Figs. 2 and 3 indicated that variability was slightly less for the within day trials than for the between day trials. This difference might be explained by the fact that it is virtually impossible to place the EMG surface electrodes in exactly the same location on 10 muscles on different days. This, slightly different EMG pick up volumes would be recorded that would calibrate differently and manifest themselves as differences in gain. Variability associated with the R^2 as a function of trial, although representing 31–34% of the variability in R^2, was biomechanically inconsequential. Average R^2 varied by at most 0.04 and could be considered excellent for all trials. Thus, R^2 performance was not negatively affected by the trial. Similar arguments could be made for the AAE. Although there was variability in AAE, the difference in variability from trial to trial, although statistically significant, had very little biomechanical meaning.

The symmetric or asymmetric nature of the exertion was a source of variation for only the gain performance measure. The overall difference in gain estimation between sagitally symmetric exertions and asymmetric exertions most likely point to a more fundamental issue in the model. Gain is lower for asymmetric exertions compared to sagitally symmetric exertions. Asymmetric exertions involve the cooperative effort of more muscles, which would involve more total muscle mass. Since gain is estimated as the same value for all muscles in this model, when more muscles are recruited to perform a task supporting the same external moment, a lower gain is predicted to all muscles as seen in this study. This trend indicates that in reality different muscles might be better represented by assigning different gain values. However, it is also comforting to acknowledge that miscalculation of gain due to asymmetry affects the total gain variability by at most 10% (Figs. 2 and 3) Also, since the gain was determined by comparing the predicted and measured sagittal moments, some error in predicting muscle gain may have resulted for the asymmetric exertions due to additional lateral and axial moments. Further analysis of the trunk moments indicated that the use of the sagittal gain was reasonable since the trunk moments were predominantly in the sagittal plane. The symmetric conditions have average sagittal, lateral and axial trunk moments of 209.7, 40.8 and 11.4 Nm, respectively. In comparison, the asymmetric lifts had, on average, a 232.2 Nm sagittal moment, 89.2 Nm lateral moment, and 32.3 Nm axial moment. Thus, the sagittal gain appear to be the most appropriate value to use since all the lifts (symmetric and asymmetric) contain a substantial amount of sagittal moment.

Experience level affected R^2 and AAE for the within day trials, between day trials, as well as several interaction terms. An apparent irony associated with these significant effects is the observation that the inexperienced subjects yielded better model performance than did the experienced subjects. Experience, in this study, was defined as subjects having experience with materials handling activities. Experienced subjects consisted of warehouse order selectors who perform materials handling activities on a daily basis. The inexperienced subjects consisted of students who were employed in our laboratory. Initially, we hypothesized that both the R^2 and AAE measures could be affected by improperly normalized EMG signals. If this were the case, then trends associated with muscle force predictions would not be well described and the R^2 and AAE measures would be indicators of a sub-optimal model. In order to test this hypothesis, each trial was examined to determine if the level of EMG observed for the muscles exceeded the maximum EMG observed during the calibration exertions. For the inexperienced group, in only 1% of the exertions did a muscle exceed the maximum EMG observed during calibration exertions. However, for the experienced group nearly 16% of the exertions exceeded the observed maximum EMGs observed during the calibration exertions. Thus, these subjects may not have been performing maximum exertions during the calibration exertions. It was thought that this may have affected the calibration of muscle force which, in turn, affected the relationship between the measured trunk moments and the EMG-predicted trunk moments. Inexperienced subjects, on the other hand, were well accustomed to performing laboratory maximal exertions, since they were experienced at participation in laboratory studies. In order to further evaluate this hypothesis, the inexperienced maximum exertions were reduced and increased by 20% intervals. However, the model performance measures did not degrade appreciably. This indicates that it is important for modeling purposes to establish a relative EMG anchor point. However, it matters little whether the anchor point is a maximum or a submaximal effort as long as it is consistent.

Next, it was hypothesized that the experienced subjects may have a very different length–strength relationship to that of the inexperienced lifters. Since the model length–strength relationship had been developed using inexperienced subjects, it might have been the case that experienced lifters had a different relationship between trunk angle and the trunk muscle length–strength calibration. In order to test this hypothesis the passive portion of the length–strength relationship applied to the experienced subjects was extended by a 10% increment to see if the model performance measures improved. When this was done, the model performance matched that of the inexperienced subjects. Adams and Hutton [22,23] have demonstrated how the ligamentous (passive force) is a function of the degree of trunk bend for an individual. Thus, this suggests that the model fidelity can be improved by individually 'tuning' the length–strength relationship to the individual. This observation points to the importance of proper length–strength calibration in EMG-assisted models. It is evident from this analysis that the better this estimate is, the better the model behaves.

The final factor that affected model performance was the weight of the object lifted. Load weight affected both the R^2 statistic and AAE for the within day exertions and the AAE for the between day exertions. The R^2 improved with increasing weight and the AAE became slightly worse (by about 3 Nm on average) for the heavier weight. The change in R^2 was about 0.04, which given the high R^2 values, was well within the acceptable range for both weight conditions. This difference was most likely due to a larger dynamic range of the moment imposed upon the spine which would tend to increase R^2 performance. The increase in AAE was most likely due to the fact that with the heavier weight a small percentage error would result in a larger absolute error. However, as shown earlier, the differences in model performance observed as a function of weight, even though statistically significant, were not significant from a biomechanical standpoint.

Collectively, this study provides some insight into the strengths and limitations of the EMG-assisted model. Overall, the model performed extremely well as evidenced by all three model performance measures. This study has shown that the model attributes are very sensitive to individual differences among subjects. However, the model is designed to be sensitive to these differences. Future models might adjust the models further for specific subject attributes such as muscle cross-section or specific muscle lines of action based upon anthropometry. The other large source of potential variation could be tracked back to the inability to properly calibrate the model EMG signal mediation factors. This study had demonstrated that for the results of an EMG-assisted model to be acceptable one must match the length–strength relationship to the experience and flexibility of the individual.

These results imply, therefore, that by properly calibrating the model it is then reasonable to assume that the vast majority of variation observed in repeated exertions of a particular trial are due to kinematic and kinetic differences inherent in the muscle control system and not a function of model randomness. Future studies should be able to build further upon these findings and better articulate how biomechanical risk occurs during materials handling activities.

References

[1] Mirka GA, Marras WS. A stochastic model of trunk muscle activities during trunk bending. Spine, 1993;8(11):1396–1409.

[2] Jager M, Luttmann A, Lauring W. Lumbar load during one-handed bricklaying. Int J Indust Ergon, 1991;8(3):261–277.

[3] Hutton WC, Adams MA. Can the lumbar spine be crushed in heavy lifting? Spine, 1982;7(6):586–590.

[4] NIOSH. Work practices guide for manual lifting. Department of Health and Human Services (DHHS), National Institute for Occupational Safety and Health (NIOSH) publication no. 81-122, 1981.

[5] Granata KP, Marras WS, Davis KG. Variation in spinal load and trunk dynamics during repeated lifting exertions. Clin Biomech, 1999;14(6):367–375.

[6] Marras WS, Lavender SA, Leurgans S, Rajulu S, Allread WG, Fathallah F et al. The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders: the effects of workplace factors, trunk position and trunk motion characteristics on injury. Spine, 1993;18(5):617–28.

[7] Marras WS, Lavender SA, Leurgans S, Fathallah F, Allread WG, Ferguson SA et al. Biomechanical risk factors for occupationally-related low back disorder risk. Ergonomics, 1995;38(2):377–410.

[8] Marras WS, Mirka GA. A comprehensive evaluation of trunk response to asymmetric trunk motion. Spine, 1992;17(3):318–326.

[9] Fathallah FA, Marras WS, Parnianpour M, Granata KP. A method for measuring external spinal loads during unconstrained free-dynamic lifting. J Biomech, 1997;30(9)975–978.

[10] Marras WS, Sommerich CM. A three-dimensional motion model of loads on the lumbar spine: I. model structure. Hum Fact, 1991;33(2):123–137.

[11] Marras WS, Sommerich CM. A three-dimensional motion model of loads on the lumbar spine: II. model validation. Hum Fact, 1991;33(2):139–149.

[12] Granata KP, Marras WS. An EMG-assisted model of loads on the lumbar spine during asymmetric trunk extensions. J Biomech, 1993;26(12):1429–1438.

[13] Granata KP, Marras WS. An EMG-assisted model of trunk loading during free-dynamic lifting. J Biomech, 1995;28(11):1309–1317.

[14] Marras WS, Granata KP. A biomechanical assessment and model of axial twisting in the thoraco-lumbar spine. Spine, 1995;20(13):1440–1451.

[15] Marras WS, Granata KP. An EMG-assisted model of trunk lateral bending. J Biomech, 1997;30(7)697–703.
[16] Marras WS, Mirka GA. Trunk responses to asymmetric acceleration. J Orthop Res, 1990;8(6):824–832.
[17] Marras WS, Mirka GA. Electromyographic studies of the lumbar trunk musculature during the generation of low level trunk acceleration. J Orthop Res, 1993;11(6):811– 817.
[18] McGill SM, Norman RW. Effects of an anatomically detailed erector spinae model on L4-S1 disc compression and shear. J Biomech, 1987;20(6)591–600.
[19] Reid JG, Costigan PA. Trunk muscle balance and muscular force. Spine, 1987;12(6):783–786.
[20] Weis-Fogh T, Alexander RM. The sustained power output from striated muscle. In: Scale effects in animal locomotion. London: Academic Press, 1977:511–25.
[21] Montgomery DC. Design and analysis of experiments, 3rd ed. New York: Wiley, 1991.
[22] Adams MA, Hutton WC. The effect of posture on the lumbar spine. J Bone Joint Surg, 1985;67B:625–629.
[23] Adams MA, Hutton WC. Has the lumbar spine a margin of safety in forward bending? Clin Biomech, 1986;1:3–6.