

# ACCURACY OF AN ALGORITHM USING MOTION DATA OF FIVE WEARABLE IMU SENSORS FOR ESTIMATING LIFTING DURATION AND LIFTING RISK FACTORS

Menekse S. Barim<sup>1</sup>, Ming-Lun Lu<sup>1</sup>, Shuo Feng<sup>2</sup>, Grant Hughes<sup>2</sup>, Marie Hayden<sup>1</sup> and Dwight Werren<sup>1</sup>

<sup>1</sup> National Institute for Occupational Safety and Health, Cincinnati, Ohio

<sup>2</sup> FocusMotion, Focus Ventures Inc., Santa Monica, California

The objective of this study was to assess the accuracy of an algorithm for processing data from five inertial measurement unit (IMU) sensors for measuring the vertical distance (V) and horizontal distance (H) of two handed lifting, trunk flexion angle (T) and lifting duration (LD). The sensors were placed on five body segments including the left wrist, right wrist, upper arm of the dominant hand, upper back, and thigh of the dominant leg. A laboratory-grade optical motion capture system was used as the ground truth for the assessment. Data were collected on ten subjects that performed 12 two-handed lifting tasks varying in height of the hands and horizontal distance between the body and the lifted object. Results showed that the algorithm performed well for determining the LD (~1 sec error) and T (~2° error). The average errors for V and H were about 33 and 6.5 cm, respectively.

## INTRODUCTION

Musculoskeletal disorders (MSDs) are a major burden on individuals, employers and the health care system (Woolf & Pfleger, 2003). Work-related MSDs (WMSDs) occur among workers in many industry sectors, in particular among manual materials handlers (Bureau of Labor Statistics or BLS, 2016). According to the most recent data from BLS, MSDs represented 31% of all workplace injuries and illnesses requiring days away from work in 2015 with a median of 12 lost work days per case. The Liberty Mutual Institute for Safety Research reported that total work compensation costs for MSDs was estimated to be \$15 billion in 2015 (Liberty Mutual Workplace Safety Index, 2016). Prevention of MSDs primarily relies on proper job designs or interventions that are effective in reducing the physical job risk factors, such as forceful movements, repetitive motions, and awkward postures (Niu, 2010). Substantial evidence has shown that work-related physical risk factors are the main source of the WMSDs, including low back disorders (LBDs) (NIOSH 1997; NRC 2001; da Costa and Vieira E., 2009). Combination of job physical risk factors (awkward posture, heavy/repetitive manual lifting) may lead to an increasing risk of developing LBP (Lu et al., 2014).

Several risk assessment methods such as self-reports and observational methods have been developed to evaluate the risk of LBDs resulting from manual lifting tasks (Barim et al., 2019; Li and Buckle, 1999; David, 2005). However, these methods have many limitations, mainly due to the subjective ratings of risk factors. One of the limitations is their inability to determine reliable risk factors for variable job tasks because of the subjective identifications of risk factors (Plantard et al., 2017). Recent developments in wearable IMU sensor technology offer a great potential for improving the accuracy of the traditional risk assessment methods. Because inertial

measurement unit (IMU) sensors are small, inexpensive, wearable, and require little power for data collection, they are suitable for monitoring the kinematics of a worker's postures in a long period and potentially in real-time (Breen et al., 2009).

Some recent studies have developed useful algorithms for measuring body postural angles, such as the trunk flexion angle (Seel et al., 2014, Aoki et al., 2016, Fang et al., 218). However, algorithms for processing data from IMU sensors have not been developed for identifying the lifting duration as well as other lifting risk factors, such as the vertical height (V) of the lifted object to the ground and the distance (H) between the lifted object and the body during a lifting task. In addition to trunk flexion, these two lifting risk factors (i.e., V and H) have been shown to be good risk predictors in many studies (Arjmand et al., 2015).

The objective of this study was to evaluate the accuracy of the algorithms that were developed using machine learning and equation based hybrid approach to estimating V, H, T and LD during two-handed lifting tasks.

## METHODOLOGY

### Study Participants

Ten subjects (five males and five females) in the National Institute for Occupational Safety and Health (NIOSH) in Cincinnati, Ohio, volunteered to participate in the study. The subjects' mean and SD for age, stature and weight were  $51.50 \pm 9.83$  years;  $170 \pm 7.4$  cm and  $85.7 \pm 20.2$  kg, respectively. Prior to data collection, written consents were obtained from the subjects in accordance with NIOSH's IRB approved study protocol.

### Data Collection

Subjects' motion data were collected in a laboratory environment with five wearable IMU sensors (Kinetic Inc.) and a motion capture system (OptiTrack and MotionMonitor® system, Innovative Sports Inc.). Figure 1 demonstrates the sensor locations for the two body motion measuring systems.

During data collection, the IMU data were streamed continuously from 5 sensors to a data logger at a rate of 25 per second through Bluetooth connection. Prior to data collection, the internal clock of the sensor data logger was synchronized with the Universal Time Clock (UTC), which was used to synchronize the motion capture data and the videos recorded by a Microsoft web camera at a resolution of 480p. The camera viewing angle was approximately perpendicular to subjects' sagittal plane.

The subjects were asked to perform symmetrical lifting tasks to produce commonly used body postures. The lifting tasks were defined by two task variables: H and V distances used in the American Conference of Governmental Industrial Hygienists Threshold Limit Values (TLVs) for lifting (ACGIH, 2007).

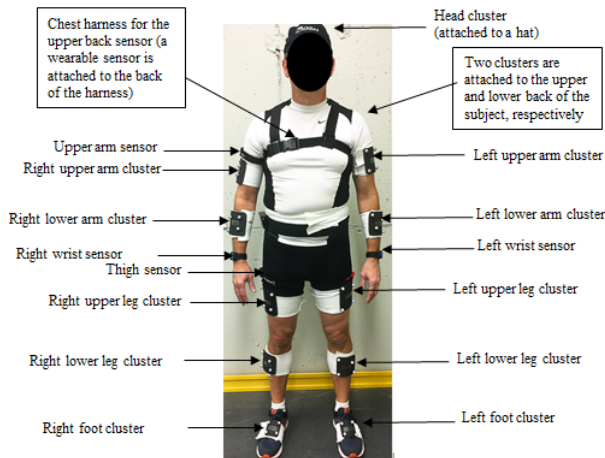


Figure 1. Body landmarks for Kinetic wearable sensors and the Optitrack motion capture system

ACGIH TLVs for lifting classifies 12 risk zones (Figure 2) using the combination of H and V for symmetrical lifting on the sagittal plane. The midpoints of the risk zones were used as the positions for starting the lifting tasks, except for zones 1-3, 4, 7, and 10. The alternative starting locations of the lift tasks for the exceptional zones were chosen for realistic lifting motion within subjects' reach envelope. Each task was repeated three times for a total 36 lifting trials for each subject. The initial lifting positions were adjusted according to each subject's anthropometric information. These trials were assigned to each subject in a random order to reduce learning effects.

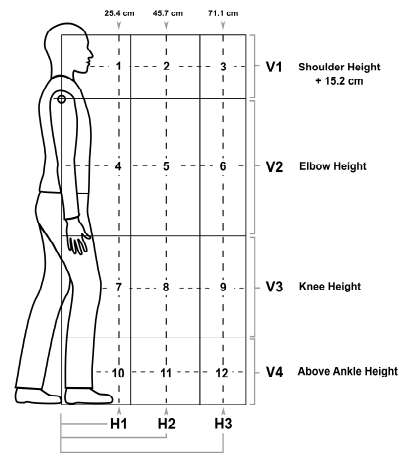


Figure 2. Initial lifting positions based on the ACGIH TLV for lifting (H1: Near Horizontal Distance from the basket, H2: Middle Distance, H3: Far Distance, V1: Vertical Height Shoulder Level, V2: Waist Level, V3: Knee Level and V4: Floor Level)

A wired grid weighing about 0.45 kg and measuring 36 X 12 cm in size with two cutout handles was used to simulate lifting a tote box during the trials. The grid was designed to help subjects create a realistic lifting motion while minimizing obstructions for body motion measurements by the optical motion capture system. A small platform (12 X 12 cm) was used to hold the grid for setting the initial lifting height. A shelf (77.5 cm in height) as the ending position of the lifting trials was set up on the direct opposite side of the initial lifting location. The distance between the initial and ending positions of the lifting trials was 3.4 m.

Before each trial, three distances of H were marked on the floor to guide the subjects to lift from one of the designated zones classified by H (25.4, 45.7 and 71 cm from the center of the two ankles to the center of the grid). With the pre-determined locations of the feet and hands in the zones, the lifting postures from the 12 zones were assured.

**Identification of lifting task and risk variables**

Video recordings from the motion capture system were used to manually identify the beginning and ending of each lifting trial. Two NIOSH researchers reviewed the video recordings to identify the video frame number for the beginning of the lift (BOL) and that for the end of the lift (EOL) for each trial. The criteria for determining the frame numbers were based on the moment when the grid started to move for the BOL and the moment when the grid was set down completely by two hands for the EOL. The two analysts discussed questionable frame numbers if in doubt, then reached agreement on the final frame numbers.

V, H, and trunk flexion (T) were determined by the motion capture system and an algorithm that processed IMU data from the wearable sensors. A detailed description for calculating the variables by the motion

capture system is reported elsewhere in a previous study (Lu et al. 2014; 2015). The description for the algorithm follows.

**Algorithm for Processing IMU Data**

Sensor data were fed into two major modules including the lifting detection module and the sensor fusion module that runs in parallel. The lifting detection module detected the occurrence of a lifting event with the timestamps of BOL and EOL of each lifting trial. The sensor fusion module tracked device orientations in real-time at 25 Hz and provided the angle of the sensor in 3D relative to the earth gravity. The lifting detection module was used to detect whether the two wrists were in synchronization at a 0.5 second step size over a 2.5 second long window. Namely, a decision was made based on previous 2.5 second data and updated every 0.5 second to extract the motion synchronization feature. This lifting detection module achieved a training accuracy of 84% detection rate (Lu et al., 2019).

Trunk flexion angle (T) was calculated with the sensor fusion reported device orientation from the sensor attached to the upper back (i.e., T12). The trunk flexion angle was calibrated for the natural lordosis of the spine while standing upright.

The subject’s forearm length was used as the base for establishing a body segment ratio model (Figure 3). Using an anthropometric database (Chaffin, 1999), the length of the upper arm ( $L_{UA}$ ) was estimated to be equal to the forearm ( $L_{FA}$ ), whereas the length of the upper leg ( $L_{Thigh}$ ) or the lower leg ( $L_{Calf}$ ) was 1.2 times  $L_{FA}$ . The length of the spine ( $L_{back}$ ) was estimated to be 1.4 times  $L_{FA}$ . The angular data of four sensors on the body segments relative to the gravity direction and the body segment ratio model were used for estimating V and H (Figure 3). The angle of the calf was ignored for the simplicity of the model. The purpose of using the simple body segment ratio model was to expedite anthropometric data collection in the field.

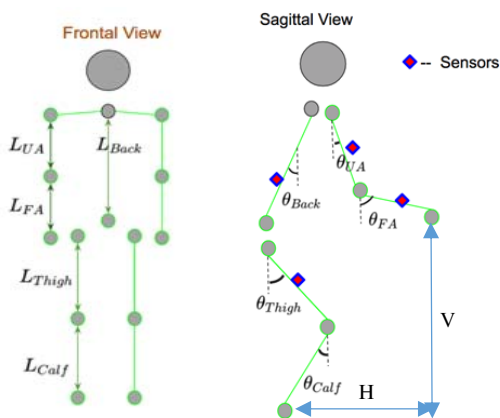


Figure 3 Body length ratio model and angular data of four sensors used for estimating V and H.

The equations below were used for estimating the V and H variables. The development of the algorithm is described in detail in a separate proceeding paper (Lu et al., 2019).

$$V = L_{back} \times \cos(\theta_{Back}) + L_{Thigh} \times \cos(\theta_{Thigh}) + L_{calf} - L_{UA} \times \cos(\theta_{UA}) - L_{FA} \times \cos(\theta_{FA})$$

$$H = L_{UA} \times \sin(\theta_{UA}) + L_{FA} \times \sin(\theta_{FA}) + L_{Back} \times \sin(\theta_{Back}) - L_{thigh} \times \sin(\theta_{Thigh})$$

**Data Analysis**

The optical motion capture data was used as the ground truth for assessing the accuracy of the algorithms for the wearable system. Specifically, Bland-Altman plots were used to determine the accuracy (bias) of the lifting risk variables. In addition, the Pearson correlation coefficient (r) was calculated to determine the strength of the correlation between IMU and motion capture data. The 12 lifting risk zones were stratified into different grouped zones (H1-H3; V1-V4) for examining patterns of H and V measurements (Figure 2).

**RESULTS**

**Accuracy of Lifting Duration Measurements**

Table 1 shows the accuracy levels of LD (EOL–BOL) measured by the IMU sensors for the 12 lifting zones. For this accuracy measure, the lower the difference, the better. Generally, the average accuracy levels for the zones were about one second, with the lowest accuracy of 1.032 sec and the highest of 0.386 sec.

Table1. Lifting Duration Differences (zone number is presented in the parentheses)

Lifting Duration Differences (sec)		
Average ± SD		
Total difference within all subjects and zones is 0.939 ± 0.673		
(1) 1.056 ± 0.586	(2) 1.036 ± 0.596	(3) 0.949 ± 0.705
(4) 0.879 ± 0.820	(5) 0.622 ± 0.386	(6) 0.754 ± 0.528
(7) 1.052 ± 1.029	(8) 1.032 ± 0.976	(9) 0.913 ± 0.466
(10) 0.884 ± 0.548	(11) 1.014 ± 0.493	(12) 1.079 ± 0.565

The timestamps for the BOL and EOL were also measured by the algorithm with IMU data. There was a systematical delay in identifying BOL and EOL of each lifting trial by the algorithm. The average difference in BOL ranged from 1.144 to 1.837 sec. and the standard deviation ranged from 1.006 sec to 2.224 while the average difference in EOL varied from 1.521 to 2.203 sec

and the standard deviation ranged from 1.371 to 2.194 sec.

**Accuracy of V and H measurements**

The Bland-Altman plots and the scatter plot of the V variable measured by the algorithm and motion capture system are shown in Figure 4 and 5, respectively. In Figure 3, the Y axis represents the difference in V (Vw (IMU data) minus Vm (motion monitor data)). The X axis represents the mean value of the two variables. Bias represents the mean of the differences in all measurements. The lower and upper level of agreement (LOA) represent the 95% confidence interval ( $\pm 1.96$  SD) of the mean difference. The mean difference of all Vw-Vm measurements was 33 cm. The algorithm used for estimating V above the shoulder (i.e., in zone V<sub>1</sub>) performed the worst with the mean difference of 43cm. A high correlation ( $r=0.95$ ) between Vm and Vw was observed (Figure 5). Total of 4 Vertical zones were compared and it showed a positive trend as average differences increases the mean difference increase. For vertical distance, subject's height was used to determine the levels (above shoulder, waist, knee and floor). In each zone, this difference can be observed.

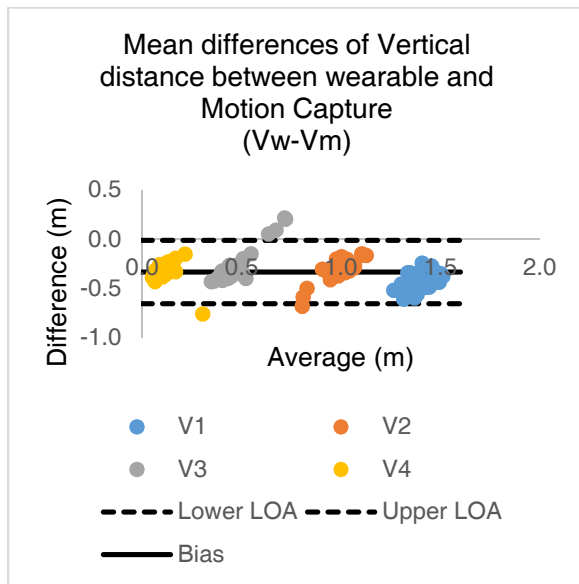


Figure 4. Comparison of V measurements (Vm: measured by the motion capture system sensors; Vw: measured by the algorithm using IMU data)

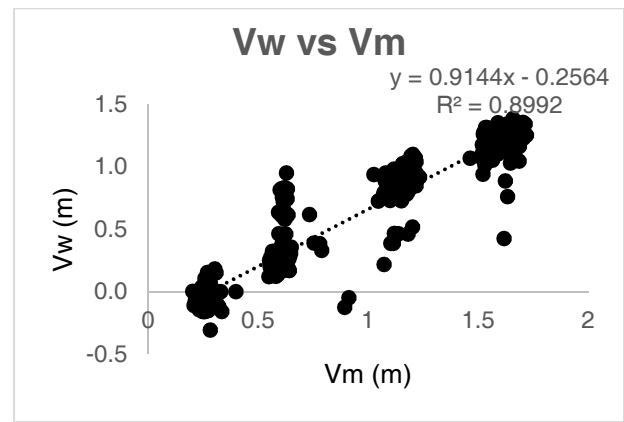


Figure 5. Correlation between Vw and Vm.

Similarly, the Bland-Altman plots and the scatter plot of the H variable are presented in Figures 6 and 7, respectively. Horizontal distance was measured with using subject's foot size. As it can be seen from Figure 6, 3 zones can be identified clearly. The overall mean and SD of H measured by the algorithm were  $0.37 \pm 0.131$  m. This mean value was 6.5 cm shorter than the estimated mean value by the motion capture system ( $0.435 \pm 0.161$  m). The mean and SD of H measured by the motion capture system for H<sub>1</sub>, H<sub>2</sub> and H<sub>3</sub> were  $0.250 \pm 0.04$ ,  $0.426 \pm 0.034$  and  $0.629 \pm 0.06$  m, respectively. The mean and SD of H measured by the algorithm for the three grouped zones were  $0.370 \pm 0.13$ ,  $0.355 \pm 0.14$  and  $0.394 \pm 0.13$ , respectively. Clearly, the H variable determined by the algorithm was significantly affected by the horizontal location of the lift in relation to the body, in particular for lifts in zone H<sub>3</sub>. A poor correlation ( $r=0.14$ ) between Hm and Hw was found (Figure 7).

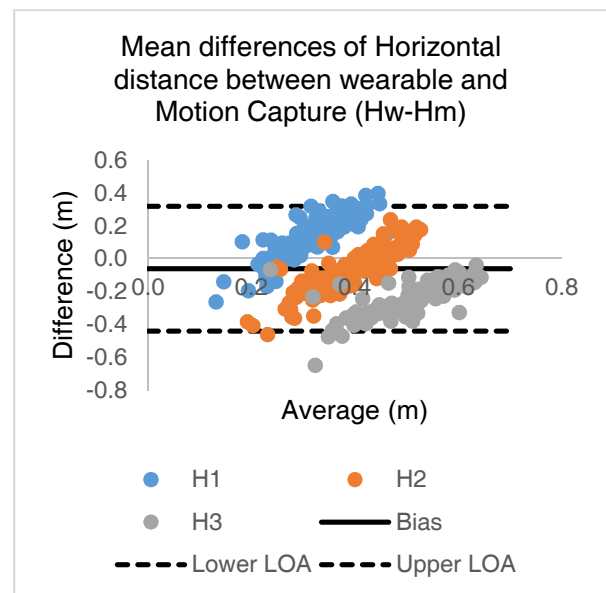


Figure 6. Comparison of the horizontal distance of the lifting trials (Hm: measured by the motion capture system sensors; Hw: measured by IMU sensors)

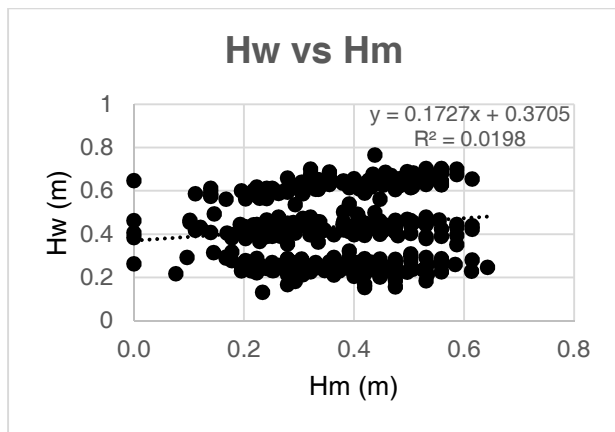


Figure 7. Correlation between Hw and Hm.

**Accuracy of trunk flexion angle**

For T measurements, the correlation ( $r=0.92$ ) between the two systems was high (Figure 8). The mean and SD of T measured by the motion capture system and the wearable system was  $37.5^\circ (\pm 32.8^\circ)$  and  $39.8^\circ (\pm 32.9)$ , respectively.

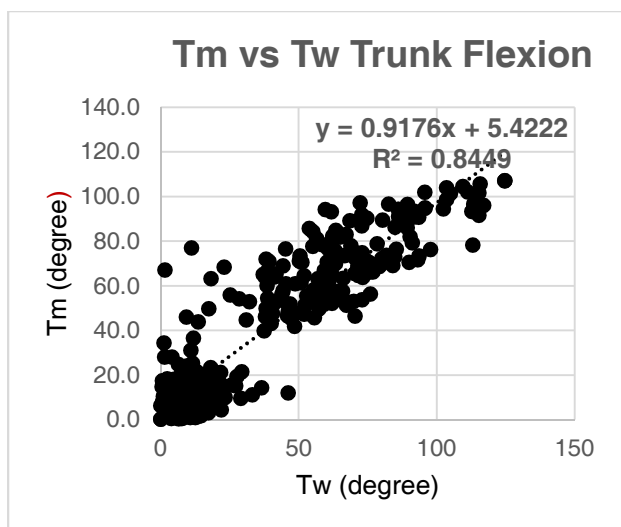


Figure 8. Correlation of the trunk flexion angle (Tm: measured by the motion capture system; Tw: measured by the algorithm using IMU data)

**DISCUSSION**

This paper presents estimated accuracy levels of an algorithm used to process IMU data for measuring H, V, T and LD, for a standardized experiment. The study results showed that the algorithm using two wrist sensors was capable of measuring LD within approximately one second of error. Although there was a systematic delay in timestamping both BOL and EOL, the large differences in LD between the two motion measuring systems were primarily caused by increased differences in the measurements for estimating the timestamps for EOL. The two analysts noticed that during EOL, many subjects may have turned their torso while releasing the grid. This inconsistent lifting behavior may have caused the

increased differences in the identification of EOL between the two motion measuring systems.

The body ration model using the gyroscope data relative to the gravity direction did not perform well in measuring V and H. The poor performance (mean errors of 33 and 6.5 cm for V and H, respectively) of the equations was probably caused by two main factors. First, the unrealistic body segment ratio model. As seen in the equations used for calculating the two variables, a small change in the body segment ratio may result in a large measurement. A precise measurement of the body segments may improve the performance of the equations. Inclusion of the angular information of the flexed lower legs may further improve the model’s performance. The additional measurements, however, may increase burden to the user in the field. Second, the rotations of the sensors on the arm may have caused inaccurate angular data. The x axis of the gyroscope data relative to the gravity direction was used for calculating V. The pronation of the arms, in particular the lower arms observed during the trials for lifts in the  $V_1$  grouped zone, changed the projection of the x axis on the sagittal plane. Consequently, the angular data of the arm sensors were no longer valid for calculating V. This finding is substantiated by the increased errors in the  $V_1$  grouped zone.

It should be noticed that the high correlation of Vw and Vm ( $r=0.92$ ) does not warrant good measurements of V with the wearable system. The high correlation of Vw and Vm resulted from a systematic underestimation in measuring V with the wearable system.

Although the mean of the differences in H measured by the two systems was poor ( $-6.5$  cm), the differences varied greatly between grouped H zones. The mean difference was the smallest for grouped zone  $H_2$ . However, the mean differences for grouped zones  $H_1$  and  $H_3$  suggested an over- and under-estimation of the variable by 12 and 24 cm, respectively. The wide spread ( $SD=20$  cm) of the differences in Hw and Hm for all three grouped zones caused a poor correlation ( $r=0.14$ ), which suggested an unreliable measuring system.

Upon a closer review of the motion capture data of the lifting trials, we found that the lifting motion of a small percentage (7%) of the lifting trials was not even (i.e., difference in Vm of both wrists  $> 3$  cm) during BOL. This uneven motion may have produced two effects on the study results. First, it may have caused inconsistent identifications of the frame numbers for BOL and EOL for even and uneven lifts. Second, it may have resulted in increased differences of V and H between the two motion measuring systems. The motion data of both wrists/arms were correctly captured by the motion capture system because motion data from the entire body including both arms were used. The wearable system, however, could only record the movements of the right arm. Consequently, inaccurate measurements of V and H by the wearable system were anticipated due to the uneven lifting motion of both arms.

There was a small percentage (8%) of missing motion data captured by the optical motion capture system. The proxy data were used based on the data available within 15 frames of the frame numbers identified for BOL and EOL. We decided to use the proxy data after a careful review of the hand and foot positions that matched the position for the BOL and EOL. This manual identification of the matched body movements may have caused slight biased results in the comparisons of the V and H measured by the two systems.

Since IMU's estimate three-axis angles using accelerometers and gyroscopes, the values will include errors. Marras et al (1992) developed a lumbar motion monitoring (LMM) device that uses a three-axis electrogoniometer. In 10 repetitions of LMM motions were performed for 20 different ranges of motion and the average deviations was 1.71 degrees. In this study, the wearable system performed well for measuring the trunk flexion angle. The error in measuring the trunk flexion angle by the sensor attached on the upper back was fairly small (~2°). This mean error is in agreement with previous studies in the literature (Faber et al., 2009, Marras et al., 1992).

The H and V variables are very critical factors for using the ACGIH TLV for lifting and the revised NIOSH lifting equation (RNLE). To take measurements for the two variables in the field, the interruption of the worker's posture is inevitable. Using a limited number of wearable IMU sensors for measuring the variables provides a practical solution to the challenge of field data collection. Although results from this study showed large errors in the measurements of V and H for a variety of lifting postures, the identifications of BOL, EOL, LD and T by the developed algorithm may be useful.

#### FUNDING AND ACKNOWLEDGEMENT

The study was financially supported by intramural funding of National Institute for Occupational Safety and Health (NIOSH). No conflict of interest is declared. The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention.

#### REFERENCES

- Aoki, T., Feng-Shun Lin, J., Kulic, D., Venture, G. (2016). Segmentation of human upper body movement using multiple IMU sensors, 2016 38<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 3163-3166.
- Arjmand, N, Amini, M., Shirazi-Adl, A., Plamondon, A., Parnianpour, M. (2015). Revised NIOSH Lifting Equation may generate spine loads exceeding recommended limits, *International Journal of Industrial Ergonomics*, **47**, 1-8.
- Barim, M.S., Sesek, R.F., Capanoglu, M.F., Drinkaus, P., Schall, M.C., Gallagher, S., Davis, G.A. (2019). Improving the risk assessment capability of the revised NIOSH lifting equation by incorporating personal characteristics, *Applied Ergonomics*, **74**, 67-73.
- Breen, PP, Nisar, A., O'laighin, G. (2009). Evaluation of a single accelerometer based biofeedback system for real-time correction of neck posture in computer users, *IEEE Engineering in Medicine and Biology Society Annual International Conference*, Minneapolis, Minnesota, USA, 7269-7272.
- Chaffin, D.B., Andersson, G.B.J., & Martin, B.J. (1999). *Occupational Biomechanics*. New York: Wiley.
- da Costa, B. and Vieira, E. R. (2010). Risk factors for work-related musculoskeletal disorders: A systematic review of recent longitudinal studies. *American Journal of Industrial Medicine*, **53**, 285-323
- David, G.C. (2005) Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders, *Occupational Medicine*, **55** (3), 190-199.
- Faber, G.S., Kingma, I., Bruijin, S.M., van Dieen, J.H. (2009). Optimal inertial sensor location for ambulatory measurement of trunk inclination, *Journal of Biomechanics*, **42** (14), 2406-09.
- Fang, Z., Yang, Z., Wang, R-B., Chen, S-Y. (2018). Inertial sensor-based knee angle estimation for gait analysis using the ant colony algorithm to find the optimal parameters for kalman filter, 2018 3<sup>rd</sup> International Conference on Automation, Mechanical and Electrical Engineering (AMEE 2018), 249-254.
- Lu, M., Waters, T.R., Krieg, E., Werren, D. (2014). Efficacy of the revised NIOSH Lifting Equation to predict risk of low back pain associated with manual lifting: A one-year prospective study, *Human Factors*, **56** (1), 73-85.
- Li, G and Buckle, P. (1999). Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods, *Ergonomics*, 674-695.
- Lu, M., Waters, T. and Werren, D. (2015). Development of human posture simulation method for assessing posture angles and spinal loads. *Human Factors and Ergonomics in Manufacturing and Services Industries*, **25**, 123-136.
- Lu, M., Feng, S., Hughes, G., Barim, M.S., Hayden, A.M. and Werren, D. (2019). Development of an algorithm for automatically measuring lifting risk factors using inertial measuring units. *Proceeding Paper in the Human Factors and Ergonomic Society Annual Meeting in Seattle, WA*.
- Marras, W.S., Fathallah, F.A., Miller, R.J., Davis, S.W., Mirka, G.A. (1992). Accuracy of a three-dimensional lumbar motion monitor for recording dynamic trunk motion characteristics, *International Journal of Industrial Ergonomics*, **9**, 75-87.
- NRC (2001). *Musculoskeletal Disorders and the workplace: low back and upper extremities*. Washington, DC.

- Plantard, P., Shum, H.P.H., Le Pierres, A-S and Multon, F (2017). Validation of an ergonomic assessment method using Kinect data in real workplace conditions, *Applied Ergonomics*, **65**, 562-569.
- Seel, T., Raisch, J., Schauer, T. (2014). IMU-Based joint angle measurement for Gait Analysis, *Sensors*, **14** (4), 6891-6909.
- Woolf, A.D. and Pfleger, B. (2003). Burden of major musculoskeletal conditions, *Bulletin of the World Health Organization*, **81** (9), 646-56.