Ergonomics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/terg20

Comparison of the observer, single-frame video and computer vision hand activity levels

Robert G. Radwin, Yu Hen Hu, Oguz Akkas, Stephen Bao, Carisa Harris-Adamson, Jia-Hua Lin, Alysha R. Meyers & David Rempel

To cite this article: Robert G. Radwin, Yu Hen Hu, Oguz Akkas, Stephen Bao, Carisa Harris-Adamson, Jia-Hua Lin, Alysha R. Meyers & David Rempel (2023) Comparison of the observer, single-frame video and computer vision hand activity levels, Ergonomics, 66:8, 1132-1141, DOI: 10.1080/00140139.2022.2136407

To link to this article: https://doi.org/10.1080/00140139.2022.2136407

Published online: 26 Oct 2022.

ERGONOMICS
2023, VOL. 66, NO. 8, 1132–1141
https://doi.org/10.1080/00140139.2022.2136407

Taylor & Francis
Taylor & Francis Group

Comparison of the observer, single-frame video and computer vision hand activity levels

Robert G. Radwin[a], Yu Hen Hu[a], Oguz Akkas[a], Stephen Bao[b], Carisa Harris-Adamson[c], Jia-Hua Lin[b], Alysha R. Meyers[d] and David Rempel[c]

[a]Industrial and Systems Engineering, University of Wisconsin, Madison, WI, USA; [b]Washington State Department of Labor and Industries, SHARP Program, Olympia, WA, USA; [c]Medicine, University of California-San Francisco, San Francisco, CA, USA; [d]Division of Field Studies and Engineering, National Institute for Occupational Safety and Health, Cincinnati, OH, USA

ABSTRACT
Observer, manual single-frame video, and automated computer vision measures of the Hand Activity Level (HAL) were compared. HAL can be measured three ways: (1) observer rating (HAL_O), (2) calculated from single-frame multimedia video task analysis for measuring frequency (F) and duty cycle (D) (HAL_F), or (3) from automated computer vision (HAL_C). This study analysed videos collected from three prospective cohort studies to ascertain HAL_O, HAL_F, and HAL_C for 419 industrial videos. Although the differences for the three methods were relatively small on average (<1), they were statistically significant ($p < .001$). A difference between the HAL_C and HAL_F ratings within ±1 point on the HAL scale was the most consistent, where more than two thirds (68%) of all the cases were within that range and had a linear regression through the mean coefficient of 1.03 ($R^2 = 0.89$). The results suggest that the computer vision methodology yields comparable results as single-frame video analysis.

Practitioner summary: The ACGIH Hand Activity Level (HAL) was obtained for 419 industrial tasks using three methods: observation, calculated using single-frame video analysis and computer vision. The computer vision methodology produced results that were comparable to single-frame video analysis.

ARTICLE HISTORY
Received 14 April 2022
Accepted 10 October 2022

KEYWORDS
Hand activity level; repetitive motion; computer vision

1. Introduction

Numerous methods have been developed for quantifying the risk of upper limb musculoskeletal injuries, including self-reports, observation, video-based single-frame analysis, and direct measurements (Joshi and Deshpande 2019). Each technique has advantages and disadvantages (David 2005). The hand activity level (HAL) rating is a 10-point visual analog scale based on hand speed and rest pauses, where HAL = 0 when the hands are idle most of the time and there are no regular exertions, and HAL = 10 when there is rapid steady motion, continuous exertion and difficulty keeping up. The observed HAL rating is recorded as an integer value. Originally introduced by Latko et al. (1997), the HAL rating has been adopted by the American Conference of Governmental Industrial Hygienists (ACGIH) in 2001 Hand Activity Threshold Limit Value® (TLV®) (ACGIH, 2021).

Epidemiological investigations demonstrated significant relationships between the TLV® action limit and elbow/forearm tendonitis, Bao et al. (2016) carpal tunnel syndrome (CTS) (Burt et al. 2011, Franzblau et al. 2005), and musculoskeletal pain (Werner et al. 2005). Longitudinal studies of manufacturing workers found that the TLV® for HAL was predictive of increased risk for carpal tunnel syndrome (Garg et al. 2012, Yung et al. 2019, Kapellusch, Gerr, et al. 2014), predicted increased risk for CTS while controlling for obesity and job strain (Burt et al. 2013), and was associated with increased risk for flexor tendon entrapment of the digits (Kapellusch, Garg, et al. 2014).

The observational method for ascertaining the HAL rating for a job requires a trained observer viewing workers performing the job on-site or observing a video of the job off-site (Lowe and Krieg 2009, Spielholz et al. 2001). Takala et al. (2010) recommended the observational HAL rating for general

CONTACT Robert G. Radwin rradwin@wisc.edu Department of Industrial and Systems Engineering, University of Wisconsin-Madison, 1550 Engineering Drive, Madison, WI 53706 USA

screening purposes. Without training, interobserver reliability is moderate (Kappa = 0.52) but with training, reliability increases (Kappa = 0.70) (Ebersole and Armstrong 2002).

In addition to observer ratings, the HAL may also be determined from a lookup table (HAL_T), which is part of the TLV® by measuring exertion frequency (F) and percent duty cycle (D) where:

$$F = \left(\frac{\text{no. exertions}}{\text{exertion time} + \text{rest time}}\right) \tag{1}$$

and

$$D = 100\left(\frac{\text{exertion time}}{\text{exertion time} + \text{rest time}}\right), \tag{2}$$

where exertion times and rest times are totalled within a cycle. Radwin et al. (2015) developed an equation for estimating HAL based on measurements of F and D using data from Latko et al. (1997) to continuously predict HAL values consistent with the TLV® look-up table. The equation:

$$\text{HAL} = 6.56 \ \ln D \left[\frac{F^{1.31}}{1 + 3.18 \ F^{1.31}}\right] \tag{3}$$

can be utilised to directly measure HAL and is also part of the TLV®. The direct measurement method is typically applied by using video single-frame analysis.

Video single-frame analysis has been used for objectively measuring F and D utilising software such as multimedia video task analysis (MVTA) by marking the video frames when exertions start and end in a cycle (Yen and Radwin 1995). This method has been frequently used for calculating HAL directly from videos (Bao, Spielholz, et al. 2006; Burt et al. 2013; Harris et al. 2011). Although the method is objective, it is time intensive, requiring the video to be viewed by an analyst in slow-motion or by single-frame.

Bao, Howard, et al. (2006) observed poor correlation between HAL observations, self-reports, and direct measurements and noted several reasons for different ratings, including using different definitions of an exertion. Wurzelbacher et al. (2010) compared more than 700 tasks for 484 workers and found that while correlated (Spearman rank = 0.49), the agreement between HAL ratings (within ±1point) from observations and calculated from videos was 61%. An objective automated method for measuring HAL is therefore desirable, especially for routine practice in industry.

Chen et al. (2013) developed an automated method for measuring HAL using a cross correlation-based template matching algorithm to track the motion trajectory of a selected region of interest (ROI) over successive video frames. The authors demonstrated that HAL can be calculated by automatically measuring F and D from hand displacement, speed and acceleration signals.

To further pursue the above goal, Akkas et al. (2015) developed an equation to measure HAL directly from RMS hand speed (S) and D since the HAL scale is anchored against the speed of hand motion/exertions and rest pauses. Furthermore, measuring F of repetitive motion is more challenging when the motion becomes more complex, such as when exertions and motions coincide. The equation was:

$$\text{HAL} = 10 \left[\frac{e^{-15.87+0.02D+2.25 \ \ln S}}{1 + e^{-15.87+0.02D+2.25 \ \ln S}}\right] \tag{4}$$

Finally, to measure D from single video cameras, Akkas et al. (2016) developed machine learning algorithms for laboratory simulations and tested them for real factory repetitive motion tasks to determine how accurately the algorithms work for different tasks (Akkas et al. 2017). A Feature Vector Training (FVT) algorithm was trained using corresponding phases (i.e. exertion and rest) based on the first cycle of each video clip. The first cycle of each video that were manually annotated using MVTA were inputted for the first cycle exertion and rest elemental times were used to train a k-nearest neighbour (kNN) classifier as the state estimator. The algorithm used $k=1$ since only one cycle was used for training. Subsequent exertion and rest states for the remaining nine cycles were automatically classified by the algorithm.

In the current study we compare the HAL values calculated using the three different methods; HAL based on an observer ratings (HAL_O), HAL calculated from single-frame-video analysis for measuring F and D (HAL_F) (Radwin et al. 2015), and HAL calculated using computer vision marker-less tracking for measuring hand speed and the kinematic properties from the hand tracking signal (i.e. location, speed and acceleration) to identify patterns when an exertion was made to measure D (HAL_C) (Akkas et al. 2015; 2016). Moreover, we provide recommendations based on these analyses for obtaining more consistent measurements of HAL.

2. Methods

2.1. Pooled data from the NIOSH, SHARP and UCSF upper extremity consortium studies

This study considers exposure data from three sources, including prospective studies performed by the National Institute for Occupational Safety and Health (NIOSH), the Washington State Department of Labour and Industries Safety and Health Assessment and Research for Prevention Program (SHARP) and the University of California-San Francisco (UCSF). These data were part of a collaborative research program, the Upper Limb Musculoskeletal Disorder Consortium (ULMSDC), that included prospective cohort studies conducted by NIOSH, six universities and one state agency (Bao et al. 2015; Harris-Adamson, Eisen, et al. 2013; Harris-Adamson, You, et al. 2013; Harris-Adamson et al. 2014; Kapellusch et al. 2013; Kapellusch, Garg, et al. 2014; Fan et al. 2015). The original pooled data for the three research studies had 1649 workers from 16 industries from a variety of industries such as Medical and Surgical Hospitals, Concrete Products, and Manufacturing. Each study collected exposure data independently using observation, direct instrumentation measurement, single-frame video analysis, or self-reports. NIOSH data were collected for 483 workers from three employers in three states. SHARP had 719 workers, from 13 different sites in Washington State. UCSF had 447 participants from four sites. All the above studies collected force, repetition, posture, and vibration exposure variables.

2.2. Ethical statement

Ethical approval was obtained from the appropriate local ethics committee or Institutional Review Board (University of Wisconsin-Madison Federal Wide Assurance no. #FWA00005399), protocol number 2016-0297-CR006, and that informed consent has been obtained.

2.3. Inclusion criteria for automated analysis

While we report the number of workers in the prospective study, each worker had a job that may involve one or more tasks. This study involved individual video clips and did not consider which operators were included or excluded. Videos were created for baseline or follow-up data. No distinction was made if a particular video associated with a unique task was recorded during a baseline or during a subsequent visit.

All three studies provided videos of every task for the current study, as defined by Bao et al. (2009). Videos satisfying the pre-determined inclusion criteria for automated analysis, described below, were used in the hand tracking algorithm and extract kinematic data including velocity and acceleration.

To make an objective video selection, we defined inclusion criteria for computer vision analysis suitability. All videos provided by the ULMSDC prospective studies were originally produced in or converted to an MP4 format and had an accompanying coded MVTA data file. Because the videos were created for a different purpose, not all videos were suitable for computer vision analysis. Examples of non-suitable videos included instances of occlusion of the hand for much of the video by other body parts, tools or other workers, or non-stable camera recording which introduced an artificial signal when the hand was tracked. In the current study, work is comprised of a series of *tasks*. A *job* J_i could consist of primarily one or more tasks $T_j \in J_i$ performed throughout the workday. A task may have occurred one or more times, as part of a job. In general, tasks are defined as typically having a specific goal, such as loading a machine, operating a tool, or lifting materials, and are specific to an industry or business, and their associated products or services. A cycle is a series of motions or exertions that are performed repeatedly. All the jobs studied had repetitive cycles where the task performed was repeated. The criteria for inclusion were videos containing at least five contiguous cycles with limited camera motions.

Videos were excluded if breaks were taken that were outside the primary task or the hand location was unable to be determined due to an obstructed view of the hand and arm. The inclusion test determined that 1098 videos were initially deemed unsuitable. An additional 136 videos were subsequently excluded because they could not be tracked using the computer vison software because of extraneous body movements, occlusions, missing activity in the video, or they depicted insignificant exertion levels for determining HAL (i.e. HAL $=$ 0). A total of 419 videos remained suitable for computer vision analysis.

HAL estimates were all made at the task level. Exposures were based on a trained analyst's observation of each participant performing his or her usual work tasks augmented with interviews of workers and their supervisors. Frequency and duty cycle was based on detailed time studies of task-level videos.

2.4. Observer HAL (HAL_O)

Each video was viewed by the same trained analyst who rated HAL_o for the dominant hand according to the visual-analog scale in the TLV (ACGIH 2021). The same analyst who conducted the below-mentioned review and re-annotation of the videos, performed the observed HAL ratings but was blinded to the HAL_F calculated values.

2.5. Single-frame video analysis HAL (HAL_F)

The videos were previously coded for exertion and rest time by multiple analysts from the individual institutions of the ULMSDC. Since the videos were previously coded using MVTA for a different purpose, the single-frame analysis was conducted for the current study using the following definition of exertions. An exertion was defined as a visible hand or forearm muscular effort while grasping an object or applying a force (e.g. hold, manipulate, trigger, push, pull, or handle an object) during task performance, regardless of the force required. Exertions that were less than 10% MVC were excluded and contributed to rest time. Every qualifying exertion was included, even if it was not followed by a pause (e.g. when a sequence of exertions occurred while retaining an object, or when grasp of an object was released and immediately followed by another grasp).

A review of the MVTA video data files originally created for the ULMSD Consortium prospective study was conducted for consistency with the definitions of exertions for the current study and when these were not consistent, they were revised. A total of 41 videos were reannotated to be made consistent with the above definitions.

The frequency (*F*) for each video was calculated based on Equation 1 as the number of exertions that occurred during the duration of the video clip of the task divided by the duration of the video clip. The duty cycle (*D*) for each video was calculated based on Equation 2 from the total duration of exertions in the clip (i.e. exertion time) and the difference between the duration of the video clip (i.e. total time) minus the exertion time (i.e. rest time). HAL_F was calculated using Equation 3.

2.6. Computer vision HAL (HAL_C)

Custom tracking software (Chen, Hu, and Radwin 2014) was used for calculating the pixel coordinates of the hand for every video frame. The process for calculating HAL was fully described in Akkas et al. (2015, 2016). The analyst indicated the dominant hand and initialised the region of interest (ROI) to track starting from the first frame that captured the most active hand fully (Figure 1). The analyst started the tracking program and manually corrected the tracked location of the hands when the algorithm failed. The dominant hand was tracked, and the hand location was scaled using hand breadth data. After the hand was tracked, the pixel data was calibrated according to the average male or female hand breadth, as described in Akkas

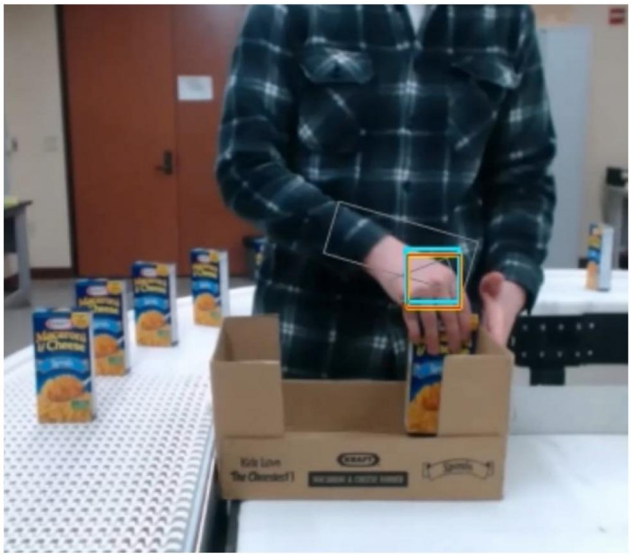

Figure 1. Screen shot of the video tracking software tracking the square box ROI located over the right hand of the worker. The hand location, speed and acceleration are used by the algorithm to calculate the HAL.

et al. (2016). Duty cycle was calculated using the kNN classifier algorithm described in Akkas et al. (2016) based on identifying exertions in the first cycle of the task and then HAL was calculated using Equation 4.

2.7. Statistical analysis

The original HAL scale used observation (Latko et al. 1997) and applied a visual analog scale rating for a mono-cycle task, results in an integer value from 0 to 10. A similar convention was used for recording HAL_O in this study. Since the single-frame (HAL_F) and computer vision (HAL_C) estimates were calculated values, they were not rounded to the nearest integer in the statistical analysis, while observed HAL values were measured as integers. Statistical analyses were performed using RStudio statistical software (Version 1.3.1093) and Microsoft Excel (Version 2207).

Analysis of variance (ANOVA) was used to test the differences between method (HAL_O, HAL_F and HAL_C). To assess the agreement between HAL methods, we calculated percent agreement within ±1-point since it was previously defined as consensus when raters were within a ± 1-point difference (Latko et al. 1997). Additionally, we utilised the Bland–Altman analysis to check agreement between the different HAL measures. The Bland–Altman analysis considers the limits of differences between two non-standard methods to check the agreement between the different HAL measures. The limits of differences were calculated from the means (m) and standard deviations (SD) of differences

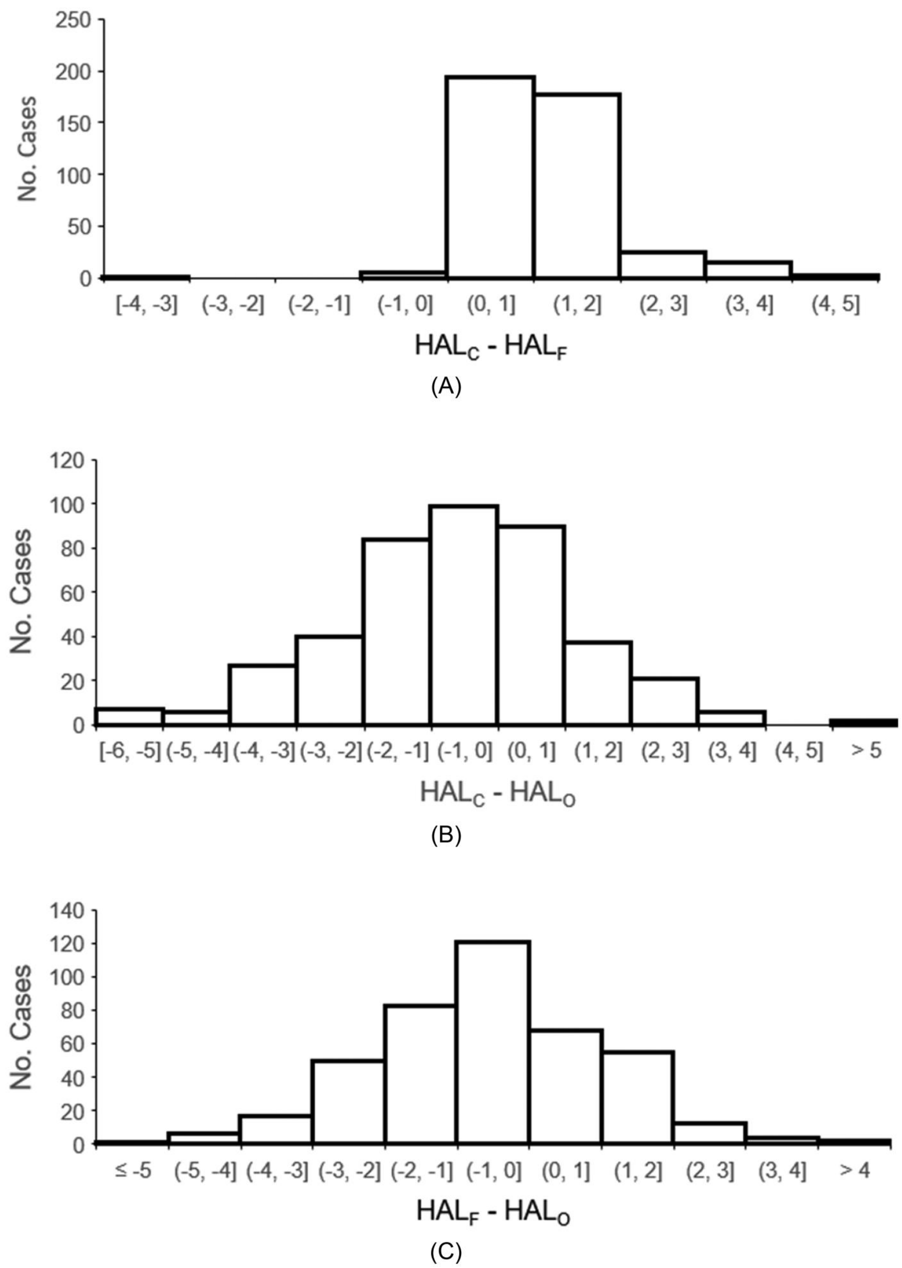

Figure 2. Histograms of the differences between the three HAL methods (A) HAL$_C$-HAL$_F$, (B) HAL$_C$-HAL$_O$, (C) HAL$_F$-HAL$_O$. Bin ranges are greater than the lower limit, left parentheses, and less than or equal to the upper limit, right bracket ($N = 419$).

between two methods. An acceptable difference was considered (±1.96 SD) (Giavarina 2015).

3. Results

The HAL values were statistically different among the three methods (HAL$_C$, HAL$_F$ and HAL$_O$) for calculating HAL ($F(2, 1251) = 15.71$, $p < .001$), where the videos were treated as a random variable. The contrast between HAL$_C$–HAL$_F$ was 0.35, and the 95% Tukey multiple family-wise confidence interval ranged from 0.05 to 0.74 ($p = .02$). The average absolute difference between HAL$_C$–HAL$_F$ was 1.20 (SD = 0.83, Max = 4.89, 95% CI = 0.11). The contrasts between HAL$_O$–HAL$_F$ was 0.56, and the 95% Tukey multiple family-wise confidence interval was 0.27–0.86 ($p < .001$). The average absolute difference between HAL$_O$–HAL$_F$ was 1.31 (SD = 1.02, Max = 5.75, 95% CI = 0.12). There was no statistically significant difference in the contrasts between HAL$_C$–HAL$_O$ with an overall difference of −0.220, and the 95% Tukey multiple family-wise confidence interval was −0.52 to 0.08 ($p = .19$). The average absolute difference between HAL$_C$–HAL$_O$ was 1.38 (SD = 1.14, Max = 7.16, 95% CI = 0.14).

Histograms for the differences between HAL$_C$–HAL$_F$, HAL$_C$–HAL$_O$, and HAL$_F$–HAL$_O$ are plotted in Figure 2.

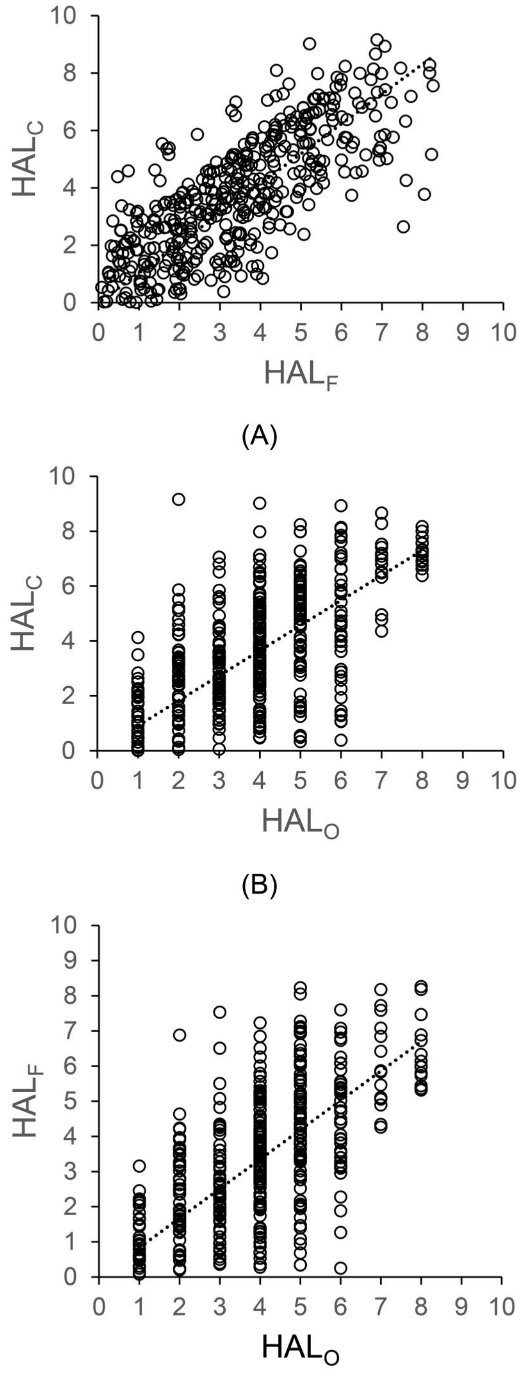

Figure 3. Scatter plots of (A)HAL$_C$ v. HAL$_F$, (B) HAL$_C$ v. HAL$_O$, (C) HAL$_F$ v. HAL$_O$.

Since the HAL scale requires an integer value between 0 and 10 (ACGIH 2021), a difference of one HAL rounded to the nearest integer between $-1.5 < HAL_C–HAL_F < 1.5$ was observed for 68% ($N = 284$ cases), and between $-2.0 < HAL_C–HAL_F < 2.0$ was observed for 89% (N = 374 cases) of the data, as shown in Figure 2(A). A difference between $-1.5 < HAL_C–HAL_O < 1.5$ was observed for 64% ($N = 269$ cases), and between and $-2.0 < HAL_C–HAL_O < 2.0$ was observed for 77% ($N = 324$ cases) of the data of the data, as shown in Figure 2(B). A difference between $-1.5 < HAL_F–HAL_O < 1.5$ was observed for 65% ($N = 271$ cases), and between $-2.0 < HAL_F–HAL_O < 2.0$ was observed for 78% (327 cases) of the data, as shown in Figure 2(C).

Scatter plots for HAL$_C$ v. HAL$_F$, HAL$_C$ v. HAL$_O$, and HAL$_F$ v. HAL$_O$ are provided in Figure 3. Linear regression between HAL$_C$ v. HAL$_F$ (Figure 3(A)) had a coefficient of 1.03 for a coefficient of determination $R^2 = 0.89$ when the intercept was set to zero, and a coefficient of 0.83 and an intercept of 0.91 for a coefficient of determination $R^2 = 0.57$. Linear regression between HAL$_C$ v. HAL$_O$ (Figure 3(B)) had a coefficient of 0.91 for a coefficient of determination $R^2 = 0.84$ when the intercept was set to zero, and a coefficient of 0.73 and an intercept of 0.85 for a coefficient of determination $R^2 = 0.34$. Linear regression between HAL$_F$ v. HAL$_O$ (Figure 3(C)) had a coefficient of 0.84 for a coefficient of determination $R^2 = 0.86$ when the intercept was set to zero, and a coefficient of 0.72 and an intercept of 0.57 for a coefficient of determination $R^2 = 0.40$.

Bland–Altman plots for HAL$_C$–HAL$_F$, HAL$_C$–HAL$_O$, and HAL$_F$–HAL$_O$ are provided in Figure 4. The plot for the HAL$_C$–HAL$_F$ had 95% of the data between the upper limit of 3.13 and the lower limit of −2.44, for a range of 5.57 (Figure 4(A)). The plot for the HAL$_C$–HAL$_O$ had 95% of the data between the upper limit of 3.26 and the lower limit of −3.72, for a range of 6.12 (Figure 4(B)). The plot for the HAL$_F$–HAL$_O$ had 95% of the data between the upper limit of 2.49 and the lower limit of −3.63, for a range of 6.98 (Figure 4(C)).

4. Discussion

The objective of the current study was to compare the different methods for estimating HAL. Although the differences in HAL for the three methods were relatively small on average (<1), the differences were statistically significant ($p < .001$). On average the ratings for HAL$_C$ and HAL$_F$ were the closest. A difference within ±1 point on the HAL scale was considered equivalent. The HAL$_C$ and HAL$_F$ ratings were the most consistent; more than half (68%) of all the task ratings were within that range. Most of the data for HAL$_C$ and HAL$_F$ (89%) had differences less than ±2 (Figure 1(A)). The absolute difference between HAL$_C$–HAL$_F$ was smallest on average (mean = 1.20, SD = 0.83, 95% CI = 0.11). The Bland–Altman analysis had similar findings, where a majority of the HAL values agreed

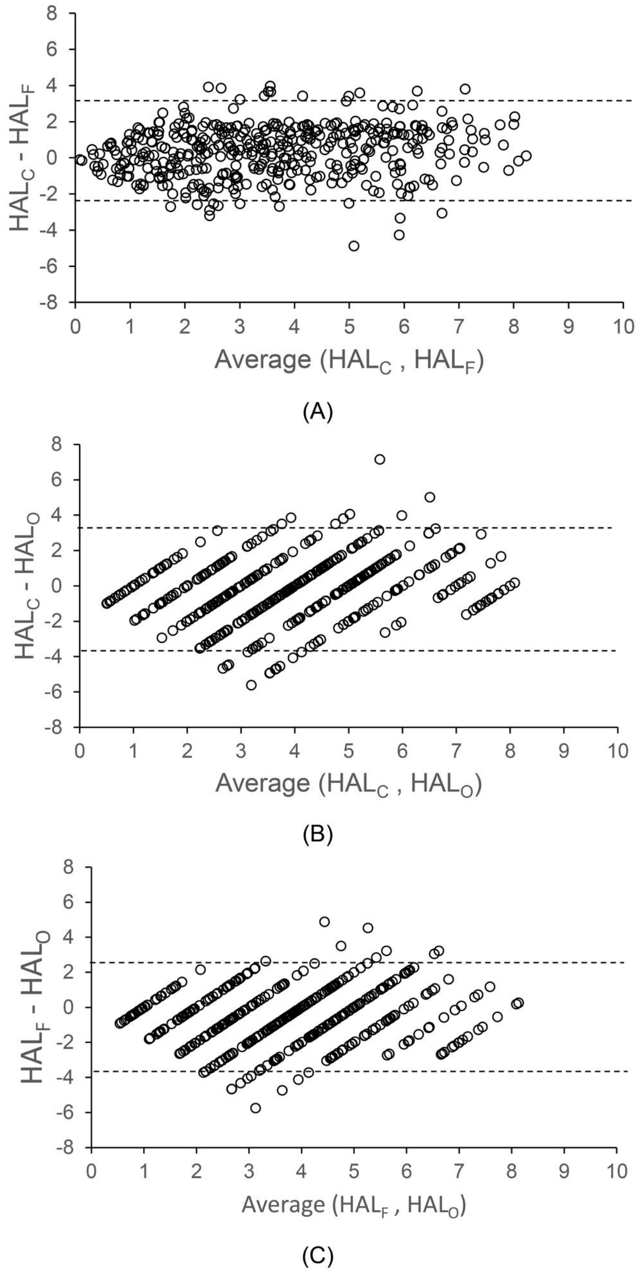

Figure 4. Bland–Altman plots of (A) HAL_C–HAL_F, (B) HAL_C–HAL_O, (C) HAL_F–HAL_O.

within a similar margin. HAL_C and HAL_F had the greatest correlation ($R^2 = .89$).

All three methods in the current study had better correlations among the three methods than those reported among individual analysts for observed HAL ratings by Spielholz et al. (2008). They compared HAL_O ratings among expert raters and had a Spearman *r* correlation of 0.67 while the correlation among experts and novice raters was 0.39. These results in the current study indicate that computed HAL values (HAL_C and HAL_F) were the most reliable and better correlated than when multiple observers rated the same videos.

Wurzelbacher et al. (2010) compared observer rated HAL (HAL_O) and calculated HAL based on single-frame video analysis and the ACGIH TLV integer HAL look-up table (HAL_T). They reported that the Spearman correlation *r* correlation for HAL_O and HAL_T was 0.49. The percent of exact agreement was 22%, while agreement within ±1 point on the HAL scale was 61%.

The consistency between definition of exertions for obtaining the frequency (F) was likely an important factor in gaining consistency between the calculated HAL measures (HAL_C and HAL_F). Bao, Howard, et al. (2006) found that different definitions of repetitive exertions produced different measures of repetitiveness and that under those circumstances their correlations were poor. Although on average the differences between the three methods for estimating HAL were small, the impact of averages on studies involving large cohorts may be less important than for HAL values used by practitioners where unreliable data for a single subject might be more consequential.

Values for HAL_F and HAL_C had good agreement despite that HAL_F was calculated based on frequency (F) in Equation 3 while HAL_C was calculated based on hand speed (S) in Equation 4. The HAL scale is anchored by hand speed and proportion of exertions, which is better aligned with the speed/duty cycle relationship described in Equation 4 for HAL_C, rather than the frequency/duty cycle relationship described in Equation 3 for HAL_F. This might have occurred because the origin of Equation 3 (Radwin et al. 2015) was developed by aligning its values with the TLV HAL table, which is based on F and D. This agreement provides further evidence that frequency measures used in Equation 3 for HAL_F, and speed measures used for HAL_O and HAL_C in Equation 4 are consistent measures. The large variations in HAL_O indicate that the observation method is less reliable than the single-frame or computer vision measures. Comparable agreement might be expected for HAL_T and HAL_C although lookup table values are discrete and lack the precision of an equation.

There are several qualifications that limit these findings. Since the video database was created for a different purpose, many of the videos were unsuitable for computer vision analysis. Although 419 videos out of 1653 video clips (25%) were available for study, there is no indication that the excluded videos were systematically related to the HAL values. Exclusions were based on independent circumstances such as the videographer's location, camera movements such as panning or zooming, or extraneous objects interfering with a full video of the workers' hands. All HAL_O

values were evaluated by a single observer, rather than multiple observers. Although the observer may have had a bias, it is assumed the bias was consistent among all HAL_O estimates.

Considering the excluded videos, it is important that in practice, recordings be made appropriately for computer vision analysis. Videos should be focussed on the hands from a vantage point perpendicular, or within ±30° of perpendicular to the plane of greatest motion (Chen et al. 2015) to minimise parallax error. Image motion should be kept at a minimum by fixating on the hands, avoiding panning or zooming, and using fixed focus lenses. The full video frame should be used for recording the hand motion, if possible. Obstructed views of the hands should also be avoided.

Since HAL is related to muscle exertions (muscle contractions), manual single-frame HAL_F can sometimes be considered more accurate than computer vision HAL_C. This is because HAL is related to muscle exertions (muscle contractions) that can occur with movement (dynamic muscle contractions) or without movement (isometric muscle contractions). In some cases, such as when operating a power hand tool, although there is no observable movement during the exertions, their occurrence may be associated with corresponding events such as squeezing a trigger, or the sound made by the tool. In these cases, a human analyst may be able to detect the exertions, where computer vision may not. HAL_F was demonstrated reliable (Wurzelbacher et al. 2010) and utilised in the ULMSDC prospective studies that demonstrated the predictivity of the TLV® for HAL (Garg et al. 2012; Yung et al. 2019; Kapellusch, Gerr, et al. 2014; Yung et al. 2019; Burt et al. 2013; Kapellusch, Garg, et al. 2014), and thus is the most validated method. Since HAL_C is based on a computer algorithm, its re-test reliability should not be an issue, although it should be formally tested in follow-on studies.

While computer vision was utilised in the current study, other researchers have incorporated continuous measurement methods such as inertial measurement unit sensors (Thamsuwan et al. 2020) which may be an alternative method for making direct HAL measures. These types of measures should be compared with computer vision in future investigations.

Although manual single-frame MVTA analysis is considered more time consuming than automated computer vision analysis, we are not aware of any benchmarks for which to compare times and analysis time was not recorded for the current study. The time to do a single-frame analysis is related to the frequency of exertions and the cycle time. The more frequent exertions occur, or the longer the cycle time, the more time the analyst needs to take to code them. The analysis time also depends on the required accuracy where the video has to pause longer and more often when a more accurate measure is necessary.

5. Conclusions

The results of the current study suggest that the computer vision methodology mostly yielded comparable results as single-frame video analysis. The potential advantages of an automated method include objectivity and that computer vision analysis has the potential to require considerably less analyst time than single-frame video analysis.

Given that there is external validity for the ACGIH TLV® for HAL, having more objective and efficient methods of applying the TLV should be useful in practice. Since there is good epidemiologic evidence of an association between the TLV® for HAL and incident CTS, computer vision shows promise for increasing efficiency and objectivity for assigning HAL ratings, which could improve adoption of the method by practitioners.

This study concludes that computed methods for calculating HAL are more consistent and reliable than observation. Computer vision methods for estimating HAL had better agreement with single-frame methods than with observation.

Acknowledgements

The authors wish to thank John Wallner and Callahan Manuel for their assistance with data analysis.

Disclosure statement

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Institute for Occupational Safety and Health, Centres for Disease Control and Prevention.

Funding

This work was supported, in part, by a grant from the National Institute for Occupational Safety and Health (NIOSH/CDC) grant number R01-OH-011024 (Radwin). Additional support came from NIOSH/CDC grant numbers R01-OH-007914 and R01-OH-009712 (Rempel).

ORCID

Robert G. Radwin http://orcid.org/0000-0002-7973-0641
Stephen Bao http://orcid.org/0000-0001-7507-2402
Carisa Harris-Adamson http://orcid.org/0000-0003-4031-5034
Jia-Hua Lin http://orcid.org/0000-0001-8908-9328
Alysha R. Meyers http://orcid.org/0000-0002-6402-0145
David Rempel http://orcid.org/0000-0002-6472-6500

References

Akkas, O., D.P. Azari, C-H. Chen, Y.H. Hu, S.S. Ulin, T.J. Armstrong, D. Rempel, and R.G. Radwin. 2015. "A Hand Speed and Duty Cycle Equation for Estimating the ACGIH Hand Activity Level Rating." *Ergonomics* 58 (2): 184–194. doi:10.1080/00140139.2014.966155.

Akkas, O., C.H. Lee, Y.H. Hu, C. Harris-Adamson, D. Rempel, and R.G. Radwin. 2017. "Measuring Exertion Time and Duty Cycle and Hand Activity Level for Industrial Tasks Using Computer Vision." *Ergonomics* 60 (12): 1730–1738. doi:10.1080/00140139.2017.1346208.

Akkas, O., C.H. Lee, Y.H. Hu, T.Y. Yen, and R.G. Radwin. 2016. "Measuring Elemental Time and Duty Cycle Using Automated Video Processing." *Ergonomics* 59 (11): 1514–1525. doi:10.1080/00140139.2016.1146347.

American Conference of Governmental Industrial Hygienists (ACGIH). 2021. *Threshold Limit Values for Chemical Substances and Physical Agents & Biological Exposure Indices*. Cincinnati, OH: ACGIH.

Bao, S., N. Howard, P. Spielholz, and B. Silverstein. 2006. "Quantifying Repetitive Hand Activity for Epidemiological Research on Musculoskeletal disorders-Part II: Comparison of Different Methods of Measuring Force Level and Repetitiveness." *Ergonomics* 49 (4): 381–392. doi:10.1080/00140130600555938.

Bao, S., P. Spielholz, N. Howard, and B. Silverstein. 2006. "Quantifying Repetitive Hand Activity for Epidemiological Research on Musculoskeletal Disorders–Part I: Individual Exposure Assessment." *Ergonomics* 49 (4): 361–380. doi:10.1080/00140130500520214.

Bao, S., P. Spielholz, N. Howard, and B. Silverstein. 2009. "Application of the Strain Index in multiple task jobs." *Applied Ergonomics* 40(1): 56–68.

Bao, Stephen S., Jay M. Kapellusch, Arun Garg, Barbara A. Silverstein, Carisa Harris-Adamson, Susan E. Burt, Ann Marie Dale, Bradley A. Evanoff, Frederic E. Gerr, Kurt T. Hegmann, Linda A. Merlino, Matthew S. Thiese, and David M. Rempel. 2015. "Developing a Pooled Job Physical Exposure Data Set from Multiple Independent Studies: An Example of a Consortium Study of Carpal Tunnel Syndrome." *Occupational and Environmental Medicine* 72 (2): 130–137.

Bao, Stephen S., Jay M. Kapellusch, Andrew S. Merryweather, Matthew S. Thiese, Arun Garg, Kurt T. Hegmann, Barbara A. Silverstein, Jennifer L. Marcum, and Ruoliang Tang. 2016. "Impact of Work Organizational Factors on Carpal Tunnel Syndrome and Epicondylitis." *Journal of Occupational and Environmental Medicine* 58 (8): 760–764.

Burt, S., K. Crombie, Y. Jin, S. Wurzelbacher, J. Ramsey, and J. Deddens. 2011. "Workplace and Individual Risk Factors for Carpal Tunnel Syndrome." *Occupational and Environmental Medicine* 68 (12): 928–933.

Burt, S., J.A. Deddens, K. Crombie, Y. Jin, S. Wurzelbacher, and J. Ramsey. 2013. "A Prospective Study of Carpal Tunnel Syndrome: Workplace and Individual Risk Factors." *Occupational and Environmental Medicine* 70 (8): 568–574.

Chen, C.H., Y.H. Hu, and R.G. Radwin. 2014. "A Motion Tracking System for Hand Activity Assessment."2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP), 320–324. IEEE. doi:10.1109/ChinaSIP.2014.6889256.

Chen, C.H., D.P. Azari, Y.H. Hu, M.J. Lindstrom, D. Thelen, T.Y. Yen, and R.G. Radwin. 2015. "The Accuracy of Conventional 2D Video for Quantifying Upper Limb Kinematics in Repetitive Motion Occupational Tasks." *Ergonomics 58* (12): 2057–2066.

Chen, C-H.E., Y.H. Hu, T.Y. Yen, and R.G. Radwin. 2013. "Automated Video Exposure Assessment of Repetitive Hand Activity Level for a Load Transfer Task." *Human Factors* 55 (2): 298–308. doi:10.1177/0018720812458121.

David, G.C. 2005. "Ergonomic Methods for Assessing Exposure to Risk Factors for Work-Related Musculoskeletal Disorders." *Occupational Medicine* 55 (3): 190–199. doi:10.1093/occmed/kqi082.

Ebersole, M.L, and T.J. Armstrong. 2002. "Inter-Rater Reliability for Hand Activity Level (HAL) and Force Metrics." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 46 (13) 1037–1040. doi:10.1177/154193120204601307.

Fan, Z Joyce, Carisa Harris-Adamson, Fred Gerr, Ellen A. Eisen, Kurt T. Hegmann, Stephen Bao, Barbara Silverstein, Bradley Evanoff, Ann Marie Dale, Matthew S. Thiese, Arun Garg, Jay Kapellusch, Susan Burt, Linda Merlino, and David Rempel. 2015. "Associations between Workplace Factors and Carpal Tunnel Syndrome: A Multi-Site Cross Sectional Study." *American Journal of Industrial Medicine* 58 (5): 509–518.

Franzblau, A., T.J. Armstrong, R.A. Werner, and S.S. Ulin. 2005. "A Cross-Sectional Assessment of the ACGIH TLV for Hand Activity Level." *Journal of Occupational Rehabilitation* 15 (1): 57–67. doi:10.1007/s10926-005-0874-z.

Garg, A., J. Kapellusch, K. Hegmann, J. Wertsch, A. Merryweather, G. Deckow-Schaefer, and E.J. Malloy. 2012. "The Strain Index (SI) and Threshold Limit Value (TLV) for Hand Activity Level (HAL): Risk of Carpal Tunnel Syndrome (CTS) in a Prospective Cohort." *Ergonomics* 55 (4): 396–414. doi:10.1080/00140139.2011.644328.

Giavarina, D. 2015. "Understanding Bland Altman Analysis." *Biochemia Medica* 25 (2): 141–151. doi:10.11613/BM.2015.015.

Harris, C., E.A. Eisen, R. Goldberg, N. Krause, and D. Rempel. 2011. "PREMUS Best Paper Competition: Workplace and Individual Factors in Wrist Tendinosis among Blue-Collar Workers–the San Francisco Study." *Scandinavian Journal of Work, Environment & Health 37* (2): 85–98. doi:10.5271/sjweh.3147.

Harris-Adamson, Carisa, Ellen A. Eisen, Ann Marie Dale, Bradley Evanoff, Kurt T. Hegmann, Matthew S. Thiese, Jay M. Kapellusch, Arun Garg, Susan Burt, Stephen Bao, Barbara Silverstein, Fred Gerr, Linda Merlino, and David Rempel. 2013. "Personal and Workplace Psychosocial Risk Factors for Carpal Tunnel Syndrome: A Pooled Study

Cohort." *Occupational and Environmental Medicine* 70 (8): 529–537. doi:10.1136/oemed-2013-101365.

Harris-Adamson, C., E.A. Eisen, J. Kapellusch, A. Garg, K.T. Hegmann, M.S. Thiese, A.M. Dale, B. Evanoff, S. Burt, S. Bao, B. Silverstein, L. Merlino, F. Gerr, and D. Rempel. 2014. "Biomechanical Risk Factors for Carpal Tunnel Syndrome: A Pooled Study of 2474 Worker." *Occupational and Environmental Medicine* 72 (1): 33–41.

Harris-Adamson, C., D. You, E.A. Eisen, R. Goldberg, and D. Rempel. 2013. "The Impact of Posture on Wrist Tendinosis among Blue-Collar Workers the San Francisco Study." *Human Factors* 56 (1): 143–150. 0018720813502807. doi:10.1177/0018720813502807.

Joshi, M, and V. Deshpande. 2019. "A Systematic Review of Comparative Studies on Ergonomic Assessment Techniques." *International Journal of Industrial Ergonomics 74*: 102865. doi:10.1016/j.ergon.2019.102865.

Kapellusch, Jay M., Arun Garg, Stephen S. Bao, Barbara A. Silverstein, Susan E. Burt, Ann Marie Dale, Bradley A. Evanoff, Frederic E. Gerr, Carisa Harris-Adamson, Kurt T. Hegmann, Linda A. Merlino, and David M. Rempel. 2013. "Pooling Job Physical Exposure Data from Multiple Independent Studies in a Consortium Study of Carpal Tunnel Syndrome." *Ergonomics* 56 (6): 1021–1037. doi:10.1080/00140139.2013.797112.

Kapellusch, J.M., A. Garg, K.T. Hegmann, M.S. Thiese, and E.J. Malloy. 2014. "The Strain Index and ACGIH TLV for HAL Risk of Trigger Digit in the WISTAH Prospective Cohort." *Human Factors* 56 (1): 98–111. doi:10.1177/0018720813493115.

Kapellusch, J.M., Frederic E. Gerr, Elizabeth J. Malloy, Arun Garg, Carisa Harris-Adamson, Stephen S. Bao, Susan E. Burt, Ann Marie Dale, Ellen A. Eisen, Bradley A. Evanoff, Kurt T. Hegmann, Barbara A. Silverstein, Matthew S. Theise, and David M. Rempel. 2014. "Exposure–Response Relationships for the ACGIH Threshold Limit Value for Hand-Activity Level: Results from a Pooled Data Study of Carpal Tunnel Syndrome." *Scandinavian Journal of Work, Environment & Health* 40 (6): 610–620. doi:10.5271/sjweh.3456.

Latko, W.A., T.J. Armstrong, J.A. Foulke, G.D. Herrin, R.A. Rabourn, and S.S. Ulin. 1997. "Development and Evaluation of an Observation Method for Assessing Repetition in Hand Tasks." *American Industrial Hygiene Association Journal* 58 (4): 278–285. doi:10.1080/15428119791012793.

Lowe, B.D, and E.F. Krieg. 2009. "Relationships between Observational Estimates and Physical Measurements of Upper Limb Activity." *Ergonomics* 52 (5): 569–583. doi:10.1080/00140130802449682.

National Research Council. 2001. *Musculoskeletal Disorders and the Workplace: Low Back and Upper Extremities*. Washington, DC: National Academies Press.

Radwin, R.G., D.P. Azari, M.J. Lindstrom, S.S. Ulin, T.J. Armstrong, and D. Rempel. 2015. "A Frequency-Duty Cycle Equation for the ACGIH Hand Activity Level." *Ergonomics* 58 (2): 173–183. doi:10.1080/00140139.2014.966154.

Spielholz, P., S. Bao, N. Howard, B. Silverstein, J. Fan, C. Smith, and C. Salazar. 2008. "Reliability and Validity Assessment of the Hand Activity Level Threshold Limit Value and Strain Index Using Expert Ratings of Mono-Task Jobs." *Journal of Occupational and Environmental Hygiene* 5 (4): 250–257.

Spielholz, P., B. Silverstein, M. Morgan, H. Checkoway, and J. Kaufman. 2001. "Comparison of Self-Report, Video Observation and Direct Measurement Methods for Upper Extremity Musculoskeletal Disorder Physical Risk Factors." *Ergonomics* 44 (6): 588–613.

Takala, Esa-Pekka, Irmeli Pehkonen, Mikael Forsman, Gert-Åke Hansson, Svend Erik Mathiassen, W Patrick Neumann, Gisela Sjøgaard, Kaj Bo Veiersted, Rolf H. Westgaard, and Jørgen Winkel. 2010. "Systematic Evaluation of Observational Methods Assessing Biomechanical Exposures at Work." *Scandinavian Journal of Work, Environment & Health* 36 (1): 3–24. doi:10.5271/sjweh.2876.

Thamsuwan, O., K. Galvin, M. Tchong-French, L. Aulck, L.N. Boyle, R.P. Ching, K.J. McQuade, and P.W. Johnson. 2020. "Comparisons of Physical Exposure between Workers Harvesting Apples on Mobile Orchard Platforms and Ladders, Part 2: Repetitive Upper Arm Motions." *Applied Ergonomics 89*: 103192. doi:10.1016/j.apergo.2020.103192.

Werner, R., A. Franzblau, N. Gell, S. Ulin, and T. Armstrong. 2005. "A Longitudinal Study of Industrial and Clerical Workers: Predictors of Upper Extremity Tendonitis." *Journal of Occupational Rehabilitation* 15 (1): 37–46. doi:10.1007/s10926-005-0872-1.

Wurzelbacher, Steve, Susan Burt, Ken Crombie, Jessica Ramsey, Lian Luo, Steve Allee, and Yan Jin. 2010. "A Comparison of Assessment Methods of Hand Activity and Force for Use in Calculating the ACGIH(R) hand Activity Level (HAL) TLV(R)." *Journal of Occupational and Environmental Hygiene* 7 (7): 407–416. doi:10.1080/15459624.2010.481171.

Yen, T.Y, and R.G. Radwin. 1995. "A Video-Based System for Acquiring Biomechanical Data Synchronized with Arbitrary Events and Activities." *IEEE Transactions on Bio-Medical Engineering* 42 (9): 944–948. doi:10.1109/10.412663.

Yung, Marcus, Ann Marie Dale, Jay Kapellusch, Stephen Bao, Carisa Harris-Adamson, Alysha R. Meyers, Kurt T. Hegmann, David Rempel, and Bradley A. Evanoff. 2019. "Modeling the Effect of the 2018 Revised ACGIH® Hand Activity Threshold Limit Value®(TLV) at Reducing Risk for Carpal Tunnel Syndrome." *Journal of Occupational and Environmental Hygiene 16* (9): 628–633. doi:10.1080/15459624.2019.1640366.