Applied Ergonomics 74 (2019) 67–73

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Improving the risk assessment capability of the revised NIOSH lifting equation by incorporating personal characteristics

Menekse Salar Barim*, Richard F. Sesek, M. Fehmi Capanoglu, Phil Drinkaus, Mark C. Schall Jr., Sean Gallagher, Gerard A. Davis

Auburn University, Industrial and Systems Engineering Department, Auburn, AL, USA

A R T I C L E I N F O

Keywords:
Revised NIOSH lifting equation (RNLE)
Personal characteristics
BMI
Age
Gender
Low back pain
Intervertebral disc cross sectional area
L5/S1
Risk assessment

A B S T R A C T

The impact of manual material handling such as lifting, lowering, pushing, pulling and awkward postures have been studied, and models using these external demands to assess risk of injury have been developed and employed by safety and health professionals. However, ergonomic models incorporating personal characteristics into a comprehensive model are lacking. This study explores the utility of adding personal characteristics such as the estimated L5/S1 Intervertebral Disc (IVD) cross-sectional area, age, gender and Body Mass Index to the Revised NIOSH Lifting Equation (RNLE) with the goal to improve risk assessment. A dataset with known RNLE Cumulative Lifting Indices (CLIs) and related health outcomes was used to evaluate the impact of personal characteristics on RNLE performance. The dataset included 29 cases and 101 controls selected from a cohort of 1022 subjects performing 667 jobs. RNLE risk assessment was improved by incorporation of personal characteristics. Adding gender and intervertebral disc size multipliers to the RNLE raised the odds ratio for a CLI of 3.0 from 6.71 (CI: 2.2–20.9) to 24.75 (CI: 2.8–215.4). Similarly, performance was either unchanged or improved when some existing multipliers were removed. The most promising RNLE change involved incorporation of a multiplier based on the estimated IVD cross-sectional area (CSA). Results are promising, but confidence intervals are broad and additional, prospective research is warranted to validate findings.

1. Introduction

"Musculoskeletal disorders (MSDs) were recognized as having occupational etiologic factors as early as the beginning of the 18th century. However, it was not until the 1970s that occupational factors were examined using epidemiologic methods, and the work-relatedness of these conditions began to appear regularly in the international scientific literature" (Bruce et al., 1997).

It has been recognized that low back pain (LBP) risk is associated with a combination of personal factors, psychological or psychosocial factors, as well as physical exposures (National Research Council, 2001). da Costa and vieira (2010) conducted a systematic review to evaluate the risk factors for work-related musculoskeletal disorders for the neck, shoulder, wrist/hand, low back, hip, knee, ankle and feet. da Costa et al.'s review supports that heavy physical work, awkward postures, lifting, psychosocial factors, BMI and age all have a strong relationship with LBP. The relationship between occupational LBP and LBP risk factors has been previously investigated primarily in field surveillance studies (Lotters et al., 2003; Marras et al., 1993, 1995a,b; Norman et al., 1998; Punnett et al., 1991; Waters et al., 1999; Bernard, 1997; Hoogendoorn et al., 2000). However, most of these studies have focused almost exclusively on the impact of work demands such as lifting, awkward postures, trunk flexion, heavy weight, force and repetition, static and forceful movements (Marras et al., 1995a,b, 2010a,b; Garg et al., 2013). Several risk assessment tools have been developed to evaluate LBP risk resulting from manual material lifting tasks. The most well-known and widely–used tool among the ergonomics community is the Revised NIOSH Lifting Equation (RNLE) (Dempsey et al., 2005; Waters et al., 1993a,b,c, 1994; Gallagher et al., 2017). However, most ergonomic assessments do not consider personal characteristics directly, rather, they focus on physical factors associated with the job demands.

Changes to the RNLE have been frequently suggested. However, most of these changes have focused on the physical demands of the job. For example, there have been recent efforts to improve risk determination for jobs with varying lifting demands and to estimate risk for an entire, variable work shift (Garg and Kapellusch, 2016; Waters et al., 2007). Despite these techniques demonstrating good estimations for

* Corresponding author. 1090 Tusculum Ave, Cincinnati, OH, USA.
E-mail address: mzs0053@auburn.edu (M.S. Barim).

https://doi.org/10.1016/j.apergo.2018.08.007
Received 19 February 2018; Received in revised form 8 August 2018; Accepted 10 August 2018
Available online 17 August 2018
0003-6870/

LBP risk at the population level, there remains room for improvement regarding individual risk assessment. Indeed, an inherent limitation of these assessment tools is that they only address the work demands, and ignore the capability of the worker performing these tasks. That is, these tools may be able to assess the risk of work activities to the general population of workers, but not the risk to an individual worker. Identifying the causes of LBP is difficult since its causes are multifactorial and involve personal, physical job factors, and workplace psychosocial characteristics (Davis and Heaney, 2000; Lu et al., 2014). It seems reasonable to investigate the risk assessment capabilities of ergonomic tools, which incorporate not only work demands, but also individual characteristics of the worker performing the job.

The RNLE attempts to assess the risks of LBP resulting from various manual material handling tasks by calculating a recommended weight for specified two-handed, and symmetrical lifting tasks. The RNLE is a job analysis method commonly used to quantify biomechanical stressors to the low back from lifting and lowering of loads in workplaces (Garg et al., 2013). The main objective of the revised equation was to prevent and reduce the occurrence of lifting and lowering overexertion injuries and low back pain among workers (Garg, 1995). An asymmetry (twisting) multiplier (AM) and coupling (grip) multiplier (CM) as well as the concept of a “Lifting Index” (LI) were added to the original (1981) NIOSH Lifting Equation (Waters et al., 1988, Waters et al., 1993a,b,c). In addition to the coupling and asymmetry changes in the revised method, modifications included a 17 kg (37.5 lb) reduction of the load constant, modifications to the horizontal multiplier, modifications to the effect of frequency and replacing multiple limits (the action limit and the maximum permissible limit) by a single limit (recommended weight limit) (Dempsey, 2002).

This equation is accepted as a useful and valuable tool for the design and evaluation of manual lifting impacts to occupational health (Jager and Luttmann, 1999) and it has gained widespread popularity in the United States and internationally as a tool for assessing the physical demands of two-handed manual lifting tasks (Waters et al., 1998). However, variation in the capabilities and limitations of individual workers can render risk assessments inaccurate for many workers. This is particularly true as the workforce changes; more females are entering into traditionally male occupations requiring manual handling and as the US workforce is increasingly obese and aging (Ricci and Chee, 2005). Suggestions have been made on how to modify the equation or multipliers used in the equation to improve its reliability, better estimate stressors faced by varying populations, expand the functionality, or simplify the RNLE (Sesek et al, 2003, 2014). This research explores the potential impact of these factors and proposes several ways to incorporate these characteristics into the Revised NIOSH Lifting Equation. Specifically, multipliers were created to explore age, gender, BMI, and a scaling factor based upon intervertebral disc diameter.

Sesek et al. (2003) explored the idea of simplifying the RNLE to see if its risk assessment ability for determining workers who are at risk of suffering a low back injury could be maintained while requiring less computation. Those findings suggest that risk assessment performance can be maintained while simultaneously simplifying the assessment effort. The goal of this research is to explore both adding and subtracting multipliers to enhance model performance with the aim of minimizing RNLE user computational burden. In that spirit, the new personal characteristic multipliers can be easily integrated *before or after* RNLE computation. Therefore, existing RNLE data can be modified for specific workers without the need to re-analyze the physical job itself. By considering both adding and subtracting multipliers, models can be explored that potentially have fewer or no net difference in multipliers while exhibiting improved performance.

The RNLE provides an empirical method for computing a recommended weight limit (RWL) for manual lifting. The actual weight lifted is divided by the RWL to create a lifting index (LI).

The LI has been used to estimate risk for developing lifting-related LBP (Liles & Mahajan, 1985; Chaffin and park, 1973; Marras et al., 1999a,b; Waters et al., 2011a,b). Higher LIs are associated with higher risk for LBP. LIs can been used to prioritize jobs for hazard abatement indicating which jobs are generally most difficult. However, not all workers will be at the same risk when performing a given set of lifting tasks. The RNLE does not consider personal differences and how these might impact a specific individual's risk for LBP. The RNLE consists of six multipliers (horizontal multiplier (HM), vertical multiplier (VM), Distance Multiplier (DM), asymmetry multiplier (AM), frequency multiplier (FM), and a coupling multiplier (CM)) and a load constant (LC) of 51 pounds. RWL is simply calculated as the product of all multipliers and the load constant:

2. Methodology

This paper modified the RNLE by considering additional multipliers and the elimination or modification of existing multipliers. New multipliers included: age, gender, Body Mass Index (BMI), IVD cross-sectional area (CSA) and a new coupling multiplier with lower coefficients for non-optimal couplings. The vertical, distance, coupling, and asymmetry multipliers were also considered for elimination. A retrospective, case-control methodology was employed to determine the predictive ability of the RNLE and modified RNLE measures.

The database was modified to allow multipliers to be “switched on or off” so that various combinations could be explored. First, multipliers were added individually to determine their impact on the model. Next, multipliers were added in various combinations to determine their impact on model performance as measured by the association of LI to negative health outcomes. Then, existing multipliers were removed individually and in combinations to measure the impact on model performance. Finally, combinations of both adding and subtracting various multipliers were considered. All combinations were evaluated based on odds ratio, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) as compared to baseline (“normal”) RNLE performance with all six original multipliers in place. All outputs are recorded in tables comparing new models to baseline RNLE data.

A database from an epidemiological study involving a large automotive manufacturer was used to explore modifications to the RNLE. The database included historical injury data and symptom interviews (Sesek, 1999). Personal identifiers such name and date of birth were not included in the data set. Researchers in the current study were blinded to all images and potential identifying information and had data on age, height, weight, and gender only. Information regarding low-back related injuries was known for each subject's job, but not whether that specific individual had reported an injury.

2.1. An automotive manufacturing ergonomic field study

The data were collected from six different automotive plants, and consist of 667 manufacturing jobs with 1022 participants as well as job-specific, historical injury data. Well-defined lifting activities meeting the RNLE criteria for analysis (e.g., two-handed lifts that are stable, unconstrained, with good foot/floor coupling, and in favorable environmental conditions) were selected for this study. Administrative jobs or jobs that did not require any lifting tasks or did not have well-defined tasks were not used in this analysis.

Personal characteristic variables investigated for this study included height, weight, age and gender (used to estimate the lower lumbar spinal geometry and compute BMI) and self-reported ratings of perceived discomfort.

Subjects were asked to report their LBP discomfort on the day they were interviewed as well as to report any LBP symptoms for the previous year. In addition, data were available regarding which jobs had one or more LBP-related medical visits during the previous year. Injuries on those jobs may or may not have been to subjects working on those jobs during the data collection. Cases were defined as subjects who had both LBP symptoms in the previous year and whose job had one or more LBP-related medical visits in the previous year.

The automotive database contained numerous instances of jobs involving several different tasks (up to six), for which RNLE was calculated using the cumulative lifting index (CLI). Cases and controls were those subjects meeting case-control definitions for whom all data were available. For example, some subjects did not report height, weight or gender and were therefore excluded.

Cases were defined as subjects who reported LBP related symptoms in the previous year and whose job had one or more reports to medical regarding LBP in the previous year. The reports were defined as "first time office visits" (FTOV) related to low back pain and may or may not have been related specifically to the subject studied. Controls were subjects who had no LBP symptoms in the previous year and whose job did not have any reports to medical regarding LBP in the previous year. Ergonomic analyses were blinded to subject symptoms and job health outcomes. Occupational health nurses collected health and symptoms data and ergonomists and engineering graduate students collected ergonomic analysis data.

There were 130 subjects meeting all inclusion criteria: 29 cases and 101 controls. The subject population was composed of 101 males and 29 females aged 23–65 (mean 42 years, SD ± 11.2), heights from 59 to 76 inches (mean 69.5, SD ± 3.6), weights from 115 to 350 pounds (mean 191, SD ± 45.1), and Body Mass Index from 17.0 to 54.8 (mean 27.6, SD ± 5.6). The prevalence of low back pain for this population was 22% (29/130).

2.2. A morphometric study of low back geometry using MRI technology

Previous research has yielded a regression equation to predict the size of an individual's intervertebral disc (IVD) cross-sectional area (Tang, 2013; Sesek et al., 2014). That study used subjects without current or chronic episodes of LBP and examined them using a whole body 3T Magnetic Resonance Imaging machine (Siemens Verio open-bore). The IVD cross-sectional area used for this study was the L5/S1 IVD measured at its center (see Line "B" in Fig. 1 below).

L5/S1 IVD CSAs = [-16.959 + 0.179*Height*2.54 + 1.7*Gender] cm^2

(Gender (G) = 0 for females and 1 for males).

IVD area was used to scale risk up or down for smaller and larger subjects, respectively. A 50th percentile female IVD area was used to normalize risk. Subjects with smaller estimated IVD areas were considered at higher risk and those with larger IVD areas were considered to be at lower risk. Normalizing to a 50th percentile female was selected

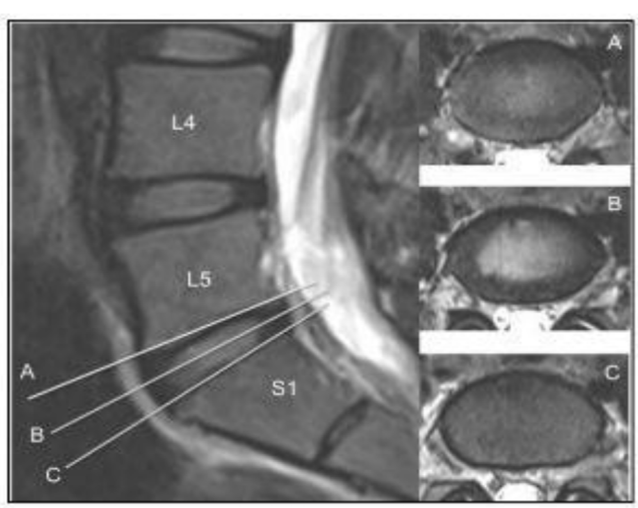

Fig. 1. Sample of MRI scan in sagittal and transverse planes.

because the RNLE is a relatively "conservative" test and produces a significant number of false positives (e.g., indicates a job is hazardous when the job does not result in symptoms and/or injury). In this way, the IVD multiplier can *decrease* estimated risk from baseline for those with larger IVD areas, therefore, this multiplier can assume values greater than 1.0. The IVD multiplier could have been normalized to any size disc, but was targeted to a smaller than average size to account for false positives common with the RNLE.

3. Experimental design

RWL = LC x HM x VM x DM x AM x FM x CM

The actual weight lifted is divided by the computed recommended weight limit to create the lifting index (LI). When the LI is ≥ 3.0, the lift is considered to pose risk for "nearly all workers" according to NIOSH (Waters et al., 1993a,b,c).

LI = Actual Weight / RWL

Modifications to the RNLE were proposed and several novel multipliers that describe fundamental characteristics of the subject were selected for evaluation. These multipliers are gender (GM), body mass index (BMIM), age (AGEM), an approximation of the low back intervertebral disc (IVD) size (IVDM) as a scaling factor to adjust risk based on a subject's specific anthropometry. The IVDM was intended to normalize the risk as a function of subject size. Individual subjects were normalized by dividing their estimated IVD area by that of a 50th percentile female. Note: while traditional RNLE multipliers can be no greater than 1.0, the IVDM can be greater than 1.0 suggesting that an individual with a larger IVD area would be at less risk than their smaller counterparts.

IVDM = Subject L5/S1 IVD Area / 50th percentile female L5/S1 IVD Area

A new, more conservative CM was also proposed and tested. The RNLE uses the following multipliers "good coupling" = 1.0, "fair coupling" = 0.95, and "poor coupling" = 0.90. The proposed new coupling multiplier (NCM) uses 1.0, 0.80, and 0.70 for good, fair, and poor couplings, respectively. The NCM was considered to determine if a stiffer penalty for poor coupling could improve risk assessment since the current multiplier decreases the RWL by only 10% (1.0–0.9) for good vs. poor coupling. A gender multiplier (GM) of 2/3 was applied to female subjects. Males were assigned 1.0 for GM. This multiplier was determined based on guidance in the "Applications Manual for the Revised NIOSH Lifting Equation" (Waters et al., 1994) that suggests that women typically have 60% of the upper body strength of males and 70% of male aerobic capacity on average. It has previously been demonstrated to improve the RNLE predictive ability (Sesek et al., 2014; Waters et al., 2011a,b). A BMI multiplier (BMIM) was applied to penalize subjects whose BMI was greater than 30. The BMIM consisted of 30/BMI for BMIs > 30 and 1.0 for BMIs less than or equal to 30. An age multiplier (AGEM) to account for strength losses expected from aging was also tested. The age multiplier was 1.0 for subjects under the age of 40 and decreased by 1% (0.01) for each year of age beyond 40.

To evaluate RNLE multipliers, a LI of 3.0 was used to classify jobs as more or less risky. Odds ratios were computed for models with various combinations of old and new multipliers. All new multipliers were tested individually and in groups to see if predictions could be improved for the RNLE CLI. Similarly, existing multipliers were removed to determine their overall contribution to risk estimation. The new multipliers work just as the original multipliers and can be simply included in the RWL calculations as shown in below:

RWL = LC x HM x VM x DM x AM x FM x CM x GM x BMIM x AGEM x IVDM

However, since these factors are intrinsic to the subject, they do not have to be computed task-by-task at the RWL level and can be used to modify an existing CLI directly. One simply computes the CLI as per NIOSH guidelines and then modifies the output as shown in below;

$$CLI_{mod} = CLI/(GM \times BMIM \times AGEM \times IVDM)$$

It is recommended that personal modifiers be accounted for at the CLI level to simplify the personalization of RNLE results for multiple workers on the same job. Modifications to the RNLE were proposed to account for an increasingly diverse, aging, and obese population of workers. For example, a given job may present more risk to an elderly and obese worker than a young and fit worker. Direct comparisons are made between the predictions of the unmodified CLI and the proposed CLI modifications. For all analyses, a CLI cut-point of 3.0 was used to differentiate high and low risk jobs.

Risk assessment was evaluated using 2 × 2 outcome matrices that compared tool predicted outcomes low or high risk ($LI < 3.0$ and $LI \geq 3.0$, respectively) to actual case/control status of each subject. The matrices were therefore composed of true positives, false positives, true negatives, and false negatives. These matrices were used to compute the Odds Ratios (ORs) for the association of negative health outcomes with CLIs. In addition, sensitivity, specificity, positive predictive value, and negative predictive values were computed for all RNLE combinations of multipliers, new and old and are presented in the results.

4. Results

Table 1 shows the impact of adding personal multipliers to the RNLE as compared to the baseline (original) RNLE CLI. Each column represents the addition of a single multiplier to the baseline RNLE. Please note that it is possible for individual subjects to change classifications with no net change in overall model performance. For example, the same number of false positives may be converted to true negatives as true negatives are converted to false positives after the inclusion of a new multiplier. Therefore, there were sometimes no net changes to models after the inclusion of new or removal of old multipliers. Also, performance statistics are identical when matrix cell values are the same. Hence, several explored models have identical statistical performance.

The IVD multiplier (IVDM) had the greatest impact on the RNLE, significantly improving overall odds ratio. However, the sensitivity dropped substantially (from 0.30 to 0.17). The GM modestly improved the odds ratio and overall model performance. The addition of an AGEM or BMIM did not alter the RNLE's predictions of high or low risk jobs. Next, multipliers were added in combinations to see if RNLE's odds ratio could be further increased. Table 2 illustrates impact of various combinations of proposed multipliers, ranging from lower to higher odds ratios.

RNLE odds ratio was maximized by adding both the IVDM and GM. The IVDM had the greatest impact and combinations that included it performed the best. An odds ratio near 25 was achieved by combining the IVDM and the GM. It should be noted, however, that sensitivity remained significantly lower than baseline with odds ratio improvements coming from the reclassification of false positives as true negatives along with some misclassification of true positives as false negatives. The confidence interval, while significant, is very large (2.8–215.4). Unlike all of the other multipliers, the IVDM can actually reduce estimated risk; reducing some over-estimation of risk.

Multipliers were also removed. As with the new multipliers, existing multipliers were investigated individually. Table 3 shows the impact of removing individual multipliers from the RNLE.

The removal of the distance multiplier (DM) slightly decreased RNLE associations with negative health outcomes. Removing the coupling (CM) and asymmetry (AM) multipliers did not change RNLE odds ratios. Interestingly, removing the vertical multiplier (VM) actually increased the RNLE odds ratio. This was achieved by reducing false positives. It should be noted, however, that sensitivity decreased because some true positives were also erroneously reclassified as true negatives. Next, combinations of RNLE multipliers were removed. Table 4 shows these combinations.

Removing some combinations of multipliers decreased performance or modestly increased performance. However, removing the VM along with other multipliers including removing all multipliers other than the horizontal multiplier (HM) and the frequency multiplier (FM) actually significantly improved overall model performance increasing the odds ratio to nearly 25. However, sensitivity was decreased from 0.30 to 0.20 and the confidence interval is broad (2.8–215.4). This is a function of a relatively small sample size. Similar to the results obtained from adding new multipliers, these results suggest that the RNLE may be overly conservative; therefore, reducing some of the factors that increase the model's predicted risk will eliminate false positives. Unfortunately, some true positives were also reclassified as false negatives.

Finally, combinations of both adding and subtracting multipliers were explored to determine if RNLE odds ratios could be further improved. Table 5 shows these combinations.

While it was not possible to improve the odds ratio further by both adding and subtracting multipliers, it is possible to achieve the best improved overall performance with no net change in number of multipliers. The replacement of the coupling multiplier (CM) with the new coupling multiplier (NCM) a showed modest improvement in odds ratio suggesting that stiffer penalties for poor coupling may be warranted. The NCM, however, was not present in any of the highest performing combinations. Overall, RNLE performance was greatly enhanced by adding the IVDM and GM while eliminating the CM and AM. However, sensitivity decreased (from 0.30 to 0.20). Positive and negative predictive values were robust at 0.86 and 0.80, respectively.

While model performance was significantly enhanced by incorporating personal characteristics, model sensitivity (detecting cases) was relatively low and was, in fact, lower than the baseline sensitivity in the best performing models (sensitivity reduced from 0.3 to 0.2). While positive predictive value (PPV) was relatively high, with 86% of subjects with CLIs over 3.0 properly identified as cases, only 1 in 5 cases, however (0.2), were identified using the new RNLE model (+IVDM, +GM). However, ergonomists can also alter decision points to impact sensitivity. For example, Table 6 below shows the impact of reducing the CLI decision cut-point from 3.0 to 2.5.

Sensitivity returned to 0.30 (baseline) along with modest improvements to specificity and PPV. The practice of ergonomics often requires tradeoffs since models are imperfect. If a company has the resources to investigate and improve more jobs, then they may opt for models or decision cut-points with superior sensitivity.

Table 1
Addition of new personal multipliers.

	CLI (Baseline)	+BMIM (No change)	+AGEM (No change)	+GM (Increase)	+IVDM (Increase)
Odds Ratio (95% CI)	6.71 (2.2–20.9)	6.71 (2.2–20.9)	6.71 (2.2–20.9)	7.83 (2.6–24.0)	19.8 (2.2–177.2)
p-value	0.0057	0.0057	0.0057	0.0015	0.0073
Sensitivity	0.30	0.30	0.30	0.33	0.17
Specificity	0.94	0.94	0.94	0.94	0.99
PPV	0.60	0.60	0.60	0.63	0.83
NPV	0.82	0.82	0.82	0.82	0.80

5. Limitations

Specific injury outcomes were unknown for subjects in the original study. Therefore, the true impact of personal characteristics cannot be fully assessed from this study. A prospective study including personal characteristics with corresponding individual subject injury outcomes is

Table 2
Combinations of new multipliers.

	CLI (Baseline)	+BMIM +AGEM (Decrease)	+BMIM +AGEM +GM (Decrease)	+GM +AGEM *or* + BMIM +GM (Decrease)	+IVDM +AGEM (Increase)
Odds Ratio	6.71	4.33	4.5	6.64	9.8
(95% CI)	(2.2–20.9)	(1.5–12.2)	(1.7–12.3)	(2.3–19.6)	(1.8–53.5)
P-value	0.0010	0.0057	0.0032	0.0006	0.0084
Sensitivity	0.30	0.30	0.33	0.33	0.17
Specificity	0.94	0.91	0.90	0.93	0.98
PPV	0.60	0.50	0.50	0.59	0.71
NPV	0.82	0.81	0.82	0.82	0.80

	CLI (Baseline)	+GM +AGEM +BMIM +IVDM (Increase)	+BMIM +IVDM +GM *or* + BMIM +IVD +AGEM *or* + IVDM +AGEM +GM (Increase)	+BMIM +IVDM (Increase)	+IVDM +GM (Increase)
Odds Ratio	6.71	9.84	12.25	19.80	24.75
(95% CI)	(2.2–20.9)	(2.4–41.0)	(2.3–64.5)	(2.2–177.2)	(2.8–215.4)
P-value	0.0010	0.0017	0.0031	0.0076	0.0036
Sensitivity	0.30	0.23	0.20	0.17	0.20
Specificity	0.94	0.97	0.98	0.99	0.99
PPV	0.60	0.70	0.75	0.83	0.86
NPV	0.82	0.81	0.80	0.80	0.80

Table 3
Removing existing multipliers.

	CLI (Baseline)	-DM (Decrease)	-CM (No change)	-AM (No change)	-VM (Increase)
Odds Ratio	6.71	5.7	6.71	6.71	12.25
(95% CI)	(2.2–20.9)	(1.8–18.1)	(2.2–20.9)	(2.2–20.9)	(2.3–64.5)
p-value	0.0010	0.0032	0.0010	0.0010	0.0031
Sensitivity	0.30	0.27	0.30	0.30	0.20
Specificity	0.94	0.94	0.94	0.94	0.98
PPV	0.60	0.57	0.60	0.60	0.75
NPV	0.82	0.81	0.82	0.82	0.80

Table 4
Combinations of removing existing multipliers.

	CLI (Baseline)	-AM, -DM (Decrease)	-CM -DM (Increase)	-CM -AM (Increase)	-DM -VM (Increase)	-CM -VM *or* -AM -VM *or* -CM -VM -AM *or* -CM -VM -AM -DM (Increase)
Odds Ratio	6.71	5.7	6.91	8.14	12.25	24.75
(95% CI)	(2.2–20.9)	(1.8–18.1)	(2.1–23.2)	(2.5–26.8)	(2.3–64.5)	(2.8–215.4)
p-value	0.0010	0.0032	0.0017	0.0006	0.0031	0.0036
Sensitivity	0.30	0.27	0.27	0.30	0.20	0.20
Specificity	0.94	0.94	0.95	0.95	0.98	0.99
PPV	0.60	0.57	0.62	0.64	0.75	0.86
NPV	0.82	0.81	0.81	0.82	0.80	0.80

warranted. Asymmetry in the original study was not precisely captured (it was recorded in discrete categories rather than continuously), perhaps impacting the performance of this multiplier. However, in practice the asymmetry measure is difficult to estimate, hence the discrete categories selected in the original study. Perhaps a simplified means of assessing the impact of asymmetry should be studied further. Some subjects may not have been on their current job for 1 year or more (employees were on their current jobs for an average of 3.3 years). Therefore, subjects with symptoms in the previous year, but related to previous jobs may be misclassified in this study as cases. This misclassification, however, should present no systematic bias towards improving results and should only result in misclassification noise. This is particularly likely since the union environment in which the data were collected typically resulted in persons transferring to less difficult jobs as they earned seniority. This may also explain why the age multiplier was not effective since older workers tend to have more seniority and can preferentially select less demanding jobs.

Future work should better control subject inclusion criteria and,

Table 5
Combinations of both adding new and subtracting existing multipliers.

	CLI (Baseline)	-CM +NCM (Increase)	+IVDM +GM -AM (Increase)	+GM +IVDM -CM (Increase)	+GM +IVDM *or* + GM +IVDM -CM -AM (Increase)
Odds Ratio	6.71	7.83	24.75	24.75	24.75
(95% CI)	(2.2–20.9)	(2.6–24.0)	(2.8–215.4)	(2.8–215.4)	(2.8–215.4)
P-value	0.0010	0.0003	0.0036	0.0036	0.0036
Sensitivity	0.30	0.33	0.20	0.20	0.20
Specificity	0.94	0.94	0.99	0.99	0.99
PPV	0.60	0.63	0.86	0.86	0.86
NPV	0.82	0.82	0.80	0.80	0.80

Table 6
RNLE model with new multipliers and reduced CLI cut-point of 2.5.

	CLI (Baseline)	+IVDM +GM (Increase)
Odds Ratio	6.71	10.29
(95% CI)	(2.2–20.9)	(2.9–36.6)
P-value	0.0010	0.0003
Sensitivity	0.30	0.30
Specificity	0.94	0.96
PPV	0.60	0.69
NPV	0.82	0.82

ideally, follow subjects prospectively, thereby minimizing potential recall bias associated with retrospective symptoms questions. Most importantly, sample size was relatively lowly. While all odds ratios were statistically significant, the confidence intervals were relatively wide. A larger sample size with more power could provide better approximations of true odds ratios. Results, however, are promising and suggest that the RNLE is a robust tool, capable of assessing relative risk even when heavily modified. This implies that the primary mechanism of risk is a function of the horizontal distance (present in all models tested) and the frequency of lifting. However, it appears that further improvements can be made.

6. Discussion

This research indicates that personal characteristics can be successfully and simply factored into ergonomic assessment tools such as the RNLE to improve odds ratios and model performance. Further, some factors may be removed from tools without a decrement in performance. In the case of the RNLE, personal characteristics may even be integrated after job level data collection to improve risk estimation for individuals. In fact, for the RNLE it may be easier to do so after computing the CLI; simply dividing the CLI by these multipliers to account for individual differences. Further study is underway to explore additional personal characteristic driven multipliers and to revisit the unsuccessful ones studied here (BMI and Age).

This study demonstrates that model performance cannot solely be assessed by univariate analyses. Various combinations of multipliers should be explored to determine the best performing models. This is particularly true for the traditional multipliers, all of which can hold maximum values of 1.0. In other words, risk estimates increase (or stay the same) when these multipliers are employed. The IVDM, on the other hand, can increase or decrease risk since it can have values both less than and greater than 1.0 (suggesting that an individual may be more or less susceptible than other workers). Future work should consider other multipliers that can hold values greater than 1.0 and/or consider modifying existing multipliers to allow values above 1.0. Multipliers exceeding 1.0 may especially help to minimize false positive classifications.

One of the best performing models included one that simply eliminated 4 of the RNLE's current multipliers. While one may be tempted to simply remove these multipliers and use this "more efficient" tool, caution is advised due to the study limitations described above. However, it may be possible to pre-screen all jobs using such a tool (with HM and FM only) and a conservative decision cut-point to increase sensitivity. A second, more complex tool with additional multipliers could be applied to those pre-screened jobs identified as potentially risky to better assess their risk. The current RNLE requires a relatively significant time commitment to analyze jobs, particularly jobs with multiple lifting tasks. Such a hybrid 2-stage approach may be attractive for rapid screening with a subsequent "deep dive" analysis (e.g., including personal characteristics and additional multipliers) for tasks quickly identified as potentially hazardous. A high negative predictive value for the first stage is prerequisite for such an approach. A multi-stage analysis methodology is currently being investigated for the fatigue failure based LiFFT tool (Gallagher et al., 2017).

7. Conclusion

Personal characteristics appear to drive a significant proportion of manual material handling (MMH) risk and should be considered when assessing MMH risk. Models incorporating a subject's estimated intervertebral disc size were the most promising and should be explored further. This study demonstrated the potential value of including these personal characteristics on diverse set of subjects and lifting tasks from 6 different automotive manufacturing sites. The subjects included a wide range of ages, BMIs, and were comprised of 22% female workers. Likewise, future work should also include subject populations that are as diverse as possible, particularly since the workforce is aging and increasingly obese. Identifying the contributions of obesity to MMH risk may further demonstrate the value of wellness programs aimed at assisting workers in maintaining healthy lifestyles and physical conditions.

References

Bernard, B., 1997. A Critical Review of Epidemiologic Evidence for Work-related Musculoskeletal Disorders of the Neck, Upper Extremity, and Low Back. U.S. Department of Health and Human Services, Center for Disease Control and Prevention, National Institute for Occupational Safety and Health, Washington, DC.

Bruce, P.B., et al., 1997. Musculoskeletal Disorders and Workplace Factors, a Critical Review of Epidemiologic Evidence for Work-related Musculoskeletal Disorders of the Neck, Upper Extremity, and Low Back. U.S. Department of Health and Human Services – source by. https://www.cdc.gov/niosh/docs/97-141/pdfs/97-141.pdf.

Chaffin, D.B., Park, K.S., 1973. A longitudinal study of low back pain as associated with occupational lifting factors. Am. Ind. Hyg. Assoc. J. 34, 513–525.

Da Costa, B.R., Vieira, E.R., 2010. Risk factors for work-related musculoskeletal disorders: a systematic review of recent longitudinal studies. Am. J. Ind. Med. 53, 285–323.

Davis, K.G., Heaney, C.A., 2000. The relationship between psychosocial work characteristics and low back pain: underlying methodological issues. Clin. BioMech. 15, 389–406.

Dempsey, P.G., 2002. Usability of the revised NIOSH lifting equation, liberty mutual research center for safety & health, 71 Frankland Road, Hopkinton, Massachusetts 01748, USA. Ergonomics *45* (12), 817–828.

Dempsey, P.G., McGorry, R.W., Maynard, W.S., 2005. A survey of tools and methods used by certified professional ergonomists. Appl. Ergon. 36, 489–503.
Garg, A., 1995. Revised NIOSH equation for manual lifting: a method for job evaluation. AAOHN J. 43 (4), 211–216.
Garg, A., Kapellusch, J.M., 2016. The cumulative lifting index (CULI) for the revised NIOSH lifting equation quantifying risk for workers with job rotation. Hum. Factors: The Journal of the Human Factors and Ergonomics Society 58 (5), 683–694.
Garg, A., Boda, S., Hegmann, K.T., Moore, J.S., Kapellusch, J.M., Bhoyar, P., Thiese, M., Merryweather, A., Deckow-Schaefer, G., Bloswick, D., Malloy, E.J., 2013. The NIOSH lifting equation and low-back pain, Part 1, association with low-back pain in the backworks prospective cohort study. Human factors: J. Human Factors Ergon. Soc. 56 (1), 6–28.
Hoogendoorn, W.E., Bongers, P.M., de Vet, H.C., Douwes, M., Koes, B.W., Miedema, M.C., Ariens, G.A., Bouter, L.M., 2000. Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: results of a prospective cohort study. Spine 25, 3087–3092.
Jager, M., Luttmann, A., 1999. Critical survey on the biomechanical criterion in the NIOSH method for the design and evaluation of manual lifting tasks. Int. J. Ind. Ergon. 23 (4), 331–337.
Liles, D.H., Mahajan, P., 1985. Using NIOSH lifting guide decreases risks of back injuries. Occup. Health Saf 54 (2), 57–60.
Lotters, F., Burdorf, A., Kuiper, J., Miedema, H., 2003. Model for the work-relatedness of low-back pain. Scand. J. Work. Environ. Health 29 (6), 431–440.
Lu, M.L., 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. Hum. Factors 56 (1), 73–85.
Marras, W.S., Lavender, S.A., Leurgans, S.E., Rajulu, S.L., Allread, W.G., Fathallah, F.A., Ferguson, S.A., 1993. 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 risk of injury. Spine 18, 617–628.
Marras, W.S., Lavender, S.A., Leurgans, S.E., Fathallah, F.A., Ferguson, S.A., Allread, W.G., Rajulu, S.L., 1995a. Biomechanical risk factors for occupationally related low back disorders. Ergonomics 38, 377–410.
Marras, W.S., Parnianpour, M., Ferguson, S.A., Kim, J.Y., Crowell, R.R., Bose, S., Simon, S.R., 1995b. The classification of anatomic- and symptom-based low back disorders using motion measure models. Spine 20, 2531–2546.
Marras, W.S., Davis, K.G., Kirking, B.C., Granata, K.P., 1999a. Spine loading and trunk kinematics during team lifting. Ergonomics 42, 1258–1273.
Marras, W.S., Fine, L.J., Ferguson, S.A., Waters, T.R., 1999b. The effectiveness of commonly used lifting assessment methods to identify industrial jobs associated with elevated risk of low-back disorders. Ergonomics 42 (1), 229–245.
Marras, W.S., Lavender, S.A., Ferguson, S.A., Splittstoesser, R.E., Yang, G., 2010a. Quantitative biomechanical workplace exposure measures: distribution centers. J. Electromyogr. Kinesiol. 20, 813–822.
Marras, W.S., Lavender, S.A., Splittstoesser, R.E., Gang, Y., 2010b. Quantitative dynamic measures of physical exposure predict low back functional impairment. Spine 35, 914–923.
Institute of Medicine and National Research Council, 2001. Musculoskeletal Disorders and the Workplace. The National Academies Press, Washington, DC.
Norman, R., Wells, R., Neumann, P., Frank, J., Shannon, H., Kerr, M., 1998. A comparison of peak vs cumulative physical work exposure risk factors for the reporting of low back pain in the automotive industry. Clin. BioMech. 13, 561–573.
Punnett, L., Fine, L.J., Keyserling, W.M., Herrin, G.D., Chaffin, D.B., 1991. Back disorders and nonneutral trunk postures of automobile assembly workers. Scand. J. Work. Environ. Health 17 (5), 337–346.
Ricci, J.A., Chee, E., 2005. Lost productive time associated with excess weight in the U.S. workforce. JOEM (J. Occup. Environ. Med.) 47 (12), 1227–1234.
Sesek, R.F., 1999. Evaluation and Refinement of Ergonomic Survey Tools to Evaluate Worker Risk of Cumulative Trauma Disorders (Dissertation). .
Sesek, R., Gilkey, D., Drinkaus, P., Bloswick, D.S., Herron, R., 2003. Evaluation and qualification of manual materials handling risk factors. Int. J. Occup. Saf. Ergon. 9 (3), 271–287.
Sesek, R., Tang, R., Gungor, C., Gallagher, S., Davis, G.A., Foreman, K.B., 2014. Using MRI-derived spinal geometry to compute back compressive stress (BCS): a new measure of low back pain risk. In: Duffy, V. (Ed.), Proceedings of the 5th AHFE Conference, pp. 13–18.
Tang, R., 2013. Morphometric Analysis of the Human Lower Lumbar Intervertebral Discs and Vertebral Endplates: Experimental Approach and Regression Models. Doctoral Dissertation. Auburn University, Auburn, AL.
Waters, T.R., Baron, S.L., Kemmlert, K., 1988. Accuracy of measurements for the revised NIOSH lifting equation. Appl. Ergon. 29 (6), 433–438.
Waters, T.R., Putz-Anderson, V., Garg, A., Fine, L., 1993a. Revised NIOSH equation for design and evaluation of manual lifting tasks. Ergonomics 36 (7), 446–749.
Waters, T.R., Putz-Anderson, V., Garg, A., 1993b. Applications Manual for the Revised NIOSH Lifting Equation. US Department of Health and Human Services, Cincinnati.
Waters, T.R., Putz-Anderson, V., Garg, A., Fine, L.J., 1993c. Revised NIOSH equation for the design and evaluation of manual lifting tasks. Ergonomics 36, 749–776.
Waters, T.R., Putz-Anderson, V., Garg, A., 1994. 'Applications Manual for the Revised NIOSH Lifting Equation' DHHS (NIOSH) Pub. No. 94–110. U. S. Department of Health and Human Services, *National Institute for Occupational Safety and Health*, Cincinnati, OH.
Waters, T.R., Baron, S.L., Piacitelli, L.A., Anderson, V.P., Skov, T., Haring-Sweeney, M., Wall, D.K., Fine, L.J., 1999. Evaluation of the revised NIOSH lifting equation. A cross-sectional epidemiologic study. Spine 24, 386–394.
Waters, T.R., Lu, M.L., Occhipinti, E., 2007. New procedure for assessing sequential manual lifting jobs using the revised NIOSH lifting equation. Ergonomics 50 (11), 1761–1770.
Waters, T.R., Lu, M., Piacitelli, L.A., Werren, D., Deddens, J.A., 2011a. Efficacy of the revised NIOSH lifting equation to predict risk of low back pain due to manual lifting: expanded cross-sectional analysis. J. Occup. Environ. Med. 53 (9), 1061–1067.
Waters, T.R., Dick, R.B., Krieg, E.F., 2011b. Trends in work related musculoskeletal disorders: a comparison of risk factors for symptoms using quality of work life data from the 2002 and 2006 General Social Survey. J. Occup. Environ. Med. 53, 1013–1024.