Ergonomics

ISSN: 0014-0139 (Print) 1366-5847 (Online) Journal homepage: https://www.tandfonline.com/loi/terg20

The effectiveness of commonly used lifting assessment methods to identify industrial jobs associated with elevated risk of low-back disorders

W. S. MARRAS , L. J. FINE , S. A. FERGUSON & T. R. WATERS

ERGONOMICS, 1999, VOL. 42, NO. 1, 229–245

The effectiveness of commonly used lifting assessment methods to identify industrial jobs associated with elevated risk of low-back disorders

W. S. MARRAS†*, L. J. FINE‡, S. A. FERGUSON† and T. R. WATERS‡

†Biodynamics Laboratory, The Ohio State University, 1971 Neil Avenue, Columbus, OH 43210, USA

‡NIOSH, Cincinnati, Ohio, USA

Keywords: Ergonomics; Low-back disorders; Risk assessment; Risk identification.

Low-back disorders (LBD) continue to be the most costly and common musculoskeletal problem facing society today. Investigators have developed tools or measures that are intended to identify jobs that will probably be associated with an elevated risk of low-back disorders. However, an important and not widely discussed issue associated with these tools and procedures has been that of the validity or effectiveness of the tools. Therefore the objective of this study was to evaluate the validity and effectiveness of two commonly used types of LBD assessment methods in terms of their ability to correctly associate jobs with LBD risk. The 1981 NIOSH *Work Practices Guide for Manual Lifting* and the 1991 NIOSH revised lifting equation, along with psychophysical measures were assessed for their ability to correctly identify high-, medium-, and low-risk (of LBD) jobs. Risk was defined according to a database of 353 industrial jobs representing over 21 million person-hours of exposure. The results indicated that both NIOSH measures were predictive and resulted in odds ratios between 3.1 and 4.6. Higher odds ratios were found when the maximum horizontal distance was used to assess a job compared to the average horizontal distance. Further analyses indicated that the two NIOSH assessment methods classified risk in very different ways. The 1981 NIOSH *Guide* demonstrated good specificity (91%) in that it identified low-risk jobs well but it also displayed low sensitivity by only correctly identifying 10% of the high-risk jobs. The 1993 NIOSH revised lifting equation, on the other hand, had better sensitivity. It correctly identified 73% of the high-risk jobs but did not identify low- and medium-risk jobs well. Using psychophysical criteria it was observed that 60% of the high-risk jobs would be judged to be acceptable, whereas, 64% and 91% of the medium- and low-risk jobs, respectively, would be judged to be acceptable. This study indicates that the different measures have various strengths and weaknesses. When controlling for occupational LBD it should be recognized that a variety of measures exist and that the measure that most appropriately assesses risk depends upon the characteristics of the job.

1. Introduction

Low-back disorders (LBD) continue to be the most costly and common musculoskeletal problem facing society today. Back problems are second only to the common cold as the reason that most people visit physicians. The American

*Author for correspondence.

Society of Orthopaedic Surgeons has identified LBD as one of the most frequently reported as well as costly medical problems facing society today (Praemer *et al.* 1992). Average costs of a LBD in the state of Ohio are now estimated to be over $30 000 per case (Ohio Bureau of Workers' Compensation 1994). There is also a significant amount of evidence that LBDs are related to job demands in the workplace (Kelsey and White 1980, National Safety Council 1989, Pope 1989, Snook 1989, Andersson 1991, Liira *et al.* 1996, Wegman and Fine 1996). It has been generally accepted that the risk of suffering from a LBD can be associated with lifting activities at the workplace.

The large magnitude of work-related low-back disorders have meant that investigators have developed tools or measures that are intended to identify jobs that will probably be associated with an elevated risk of low-back disorders. These tools have several benefits. First, they can be used to identify high-risk jobs. Second, they can be useful for developing solutions for problematic workplace situations. Finally, these tools can be used to evaluate the effectiveness of potential ergonomic solutions considered in a work environment. In addition, these tools may also be useful in identifying which *specific features* of a job are contributing to the potential elevated risk. However, an important and not widely discussed issue associated with these tools and procedures has been that of the validity or effectiveness of the tools.

Validity can be described as the degree to which a measurement measures what it purports to measure. In the case of this paper the predictive value to discriminate between jobs with different risk of occupationally-related low-back disorders is of interest. The predictive validity implies that the estimated level of risk with a particular tool roughly equals the observed level of risk. It is important that the validity of ergonomic tools be established so that the authors can assess the effectiveness of the ergonomic measure.

The accurate identification of high-risk jobs is the first step in an effective control programme. Two approaches are possible. The first relies on the use of health surveillance data to identify jobs with an elevated risk of workers' compensation claims or other reports of low-back disorders. The disadvantage of this approach is that it can only be used retrospectively. One must wait until injury trends have occurred before problem jobs are identified. Thus, the problem with this approach is that injuries must occur before adjustments to the workplace can be made. The second approach is to use a hazard identification approach to evaluate the risk of a job by examining the physical characteristics of the job such as the weight of the object being lifted. This paper will evaluate two classes of tools with respect to their ability to facilitate the second approach. In this paper the authors will compare the ability of two types of widely used tools or measures to identify high-risk jobs for LBD in a database of high- and low-risk manual material handling jobs. The first type of measure consists of the 1981 NIOSH *Work Practices Guide for Manual Lifting* (NIOSH Guide) (NIOSH 1981) and the 1993 NIOSH revised lifting equation (NIOSH lifting equation) (Waters *et al.* 1993). The second type of measure consists of psychophysical criteria used by Liberty Mutual (Snook 1978, Snook and Ciriello 1991). It should also be pointed out that these psychophysical criteria have been partly incorporated into the NIOSH equation. Hence, these measures do have some inherent overlap between them.

In order to address the issue of the effectiveness of a procedure for the assessment of LBD risk a criterion-oriented or predictive validation is necessary. In a criterion-oriented or predictive validation a record of occupationally-related LBDs is maintained and compared to the predictions of the assessment procedure (Cronbach 1970). Only in this manner can the true effectiveness of a tool or model be evaluated. Criterion-oriented validations address the bottom line issue of whether or not the tool or model works under realistic circumstances.

Several assessment tools or models are commonly used for the assessment and correction of lifting situations in the workplace. One of the better recognized tools has been the 1981 NIOSH *Guide*. This *Guide* was established based upon four criteria. These criteria consist of epidemiological, biomechanical, physiological, and psychophysical knowledge. The 1981 *Guide* was intended to control musculoskeletal injuries in general not only back injuries. This *Guide* establishes a safe limit for lifting called an 'action limit' as well as an upper limit for lifting, which is three times the action limit and called the 'maximum permissible limit'. According to the guide, 'musculoskeletal injury rate and severity rates have been shown to increase moderately in populations exposed to lifting conditions described by the action limit' and 'musculoskeletal injury rates and severity rates have been shown to increase significantly in populations when work is performed above the maximum permissible limit' (NIOSH 1981:124). The two limits are defined in terms of a lifting equation consisting of a load constant representing the greatest weight a person could lift that is then discounted by several discounting factors or 'multipliers' that reduce the allowable weight of the object as a function of the workplace features. The multipliers in the 1981 *Guide* consisted of: (1) the horizontal distance of the load from the worker, (2) the vertical location of the load at the origin of the lift, (3) the vertical distance travelled by the load, and (4) the average frequency of lifting. There were also several assumptions associated with the use of this *Guide*. These assumptions consisted of: (1) the lift was smooth, (2) the lift was two-handed and occurring in the sagittal plane, (3) the load was of moderate width, (4) the lifting posture was unrestricted, (5) good coupling, and (6) a favourable ambient environment. The idea behind this model was to define the action limit and maximum permissible limit given the requirements of the workplace and to compare the load lifted by the worker to these limits. Depending upon whether the load was below the action limit or above the maximum limit the task was judged to be either safe or to place the worker at risk. The 1981 *Guide* has been widely used as a design tool by occupational health professionals and ergonomists.

In 1993 a revision to the 1981 *Guide* was introduced. This 1993 *Revised NIOSH Lifting Equation* was designed to work in the same manner as the 1981 *Guide* in that there was a load constant that was mediated by several multipliers (Waters *et al.* 1993). However, in the 1993 *Lifting Equation* the load constant as well as the form of the multipliers was changed and two additional multipliers were included. The additional factors consisted of an asymmetry and a coupling multiplier. The remaining assumptions (other than asymmetry and coupling) from the 1981 *Guide* still applied. The concepts of an action limit and maximum permissible limit were also revised. The 1993 *Lifting Equation* predicted a 'recommended weight limit' that was compared to the weight of lift in the task of interest. The quotient of the weight lifted compared to the recommended weight limit yielded a 'lifting index'. The revised lifting equation (Waters *et al.* 1993:768) states that 'it is likely that lifting tasks with a lifting index greater than 1.0 pose an increased risk for lifting related low back pain for some fraction of the workforce' and 'many workers will be at elevated risk if the lifting index exceeds 3.0'. The 1993 lifting equation serves as an assessment tool as well as a design tool. This *Equation* has also rapidly gained popularity in the industrial and occupational health community. However, few studies have evaluated the ability of these two measures to identify high-risk jobs.

Another LBD control measure that has traditionally been used to design workplaces has been the psychophysical assessment of manual materials handling tasks. In the psychophysical approach, workers are asked to select maximum acceptable weights for a range of specific lifting tasks. Each task involves lifting an object at a set rate, height, distance and other conditions set by the experimenter. In general, weaker workers select lower weights than stronger workers. Snook (1978) has reported the results of six psychophysical studies where the task parameters were varied and the subjects were asked to adjust the load weight to a level that was acceptable to them throughout the workday. Based upon these studies, tables were developed that report the maximum acceptable weight that one could lift given specific workplace parameters. Snook *et al.* (1978) also validated this model against 191 back injury records and reported that this technique could reduce back injuries by up to one-third. In this study jobs that were acceptable to at least 75% of the workforce were compared to jobs that were not acceptable to 75% of the workforce. The latter jobs had higher low-back injury rates. In 1991 Snook and Ciriello reported revisions of these psychophysical tables based upon the additional testing. They suggest that these revised tables could be used by industry to assist 'in the evaluation and design of manual handling tasks and thereby contribute to the reduction of disability from LBD' (Snook and Ciriello 1991:1212). However, the values reported in these tables are different from those reported in 1978. Some values are lower while other values are higher. As with the NIOSH measures the authors have been unable to find any criterion-oriented validation of the 1993 revised tables.

Even though these evaluation methods are widely used there is a void in the body of knowledge in that these approaches have not been evaluated for their effectiveness or validity in controlling LBD in the workplace. The objective of this study was to evaluate the validity and effectiveness of these commonly used tools in terms of their ability to correctly associate jobs with LBD risk. An objective criterion-oriented database was used for this assessment. An additional goal of this study was to assess the sensitivity and specificity of these tools. In other words, if these tools or models did misidentify jobs according to risk, were they overestimating the risk or underestimating the risk?

2. Methods

This study has utilized a surveillance database that has been developed over a 6-year period and has been used in several other studies (Marras *et al.* 1993, 1995). The specific approach used in this study involved: (1) identification of industries involved with repetitive MMH work; (2) examination of the company medical records as well as the health and safety records to identify those repetitive MMH jobs that were associated, historically, with low-, medium-, or high-risk of occupationally-related LBD; (3) quantitative monitoring of workplace factors associated with each of these jobs; (4) application of the collected data to each of the three ergonomic tools so that a measure of LBD injury risk prediction could be determined; and (5) the evaluation of the historical risk data compared to the ergonomic tool prediction so that the relationship between the evaluation tool and the prediction of LBD injury could be assessed.

2.1. *Study database*

The database consisted of a cross-sectional evaluation of 353 industrial jobs from 48 manufacturing companies throughout the midwestern USA. The types of industries sampled included automobile assembly, machine products manufacturing, vehicle parts accessory assembly, rubber and plastic product production, truck assembly, food processing, electrical and electronic assembly, chemical production, printing, paper production, lumber production, food processing, metal fabrication and glass production. Only highly repetitive materials handling jobs that did not involve job rotation were used in the database. Jobs were reviewed to ensure that the material handling activities associated with the work have remained relatively unchanged over the years. Jobs examined in this database were divided into three groups, low-, medium-, and high-risk of LBD, based upon examination of the injury and medical records. Whenever possible, company medical reports were used to categorize risk. In some cases only injury logs (OSHA 200 logs) were available. All records were scrutinized before entry into the database. Quality control checks were employed to ensure that the database was as accurate as possible. These checks included comparing the medical records to the reporting logs and reviewing questionable cases with company personnel. When errors were identified corrections were made prior to entry into the database.

The independent variable in this study consisted of three levels of job-related LBD risk categories based upon observed LBD incidence rates derived from a much larger surveillance database (over 600 jobs) collected by the Biodynamics Laboratory over the past 15 years. All incidence rates were based upon the availability of an average of over 30 person-years of exposure data for jobs where the lifting requirements did not change during the recorded exposure time. It is recognized that the definition of LBD risk can be extremely problematic. However, for the purposes of this assessment three operational definitions of LBD risk were established. Low-risk jobs were operationally defined as those jobs with at least 3 years of records showing no LBD injuries and no turnover and represent the 25th percentile of risk (from the larger database). Medium- and high-risk jobs were also defined based on the distribution of the larger database. Medium-risk jobs were defined as those jobs that had greater than zero LBD incidences but incidence rates of less than 12 per 200 000 h of exposure (75th percentile of the distribution of the data). This 75th percentile break point uses similar logic to that of the psychophysical tools. High-risk jobs were those jobs that had at least 12 incidences of LBD per 200 000 h of exposure. The mean incidence rate of the high-risk group was 26.4. Of the 353 jobs examined, 124 of the jobs were considered to be low-risk, 118 were considered to be medium-risk, and 111 were considered to be high-risk. This database represents over 10 688 person-years (well over 21 million person-hours of exposure) of exposure data.

The dependent variables in this study consisted of the workplace characteristics associated with each job that are used by the 1981 NIOSH *Guide*, the 1993 NIOSH lifting equation, and the psychophysical method of Snook and Ciriello (1991). The workplace characteristics consisted of the average values of the variables typically considered in current workplace guidelines for materials handling (NIOSH 1981, Waters *et al*. 1994). Specifically, these variables were: (1) the maximum horizontal distance of the load from the spine; (2) the weight of the object lifted; (3) the height of the load at origin of the lift; (4) the height of the load at the destination of the lift; (5) the frequency of lifting (lift rate); and (6) the asymmetric angle of the lift (as defined by NIOSH 1993). It should be noted that coupling was not available for this analysis. A statistical description of the dependent variables in this database may be found in Marras *et al*. (1993, 1995).

2.2. *Industrial surveillance protocol*

Each job of interest was first reviewed to ensure that monitoring could be done safely without interfering with production. Subjects provided consent for participation and then answered a questionnaire about their health and employment history. Information about past low-back strain injuries or injury symptoms in other parts of the body was collected. Information about subject experience with the job of interest was also recorded. The subject was then asked to return to work where data were collected for at least 10 job cycles.

2.3. *Analyses*

Several validity measures were used to assess the predictiveness of the ergonomic tools of interest. First, the relationship of each workplace variable (multipliers in the NIOSH *Guide*) to the risk groups was examined. This analysis included a simple logistic regression model fit for each variable (Hosmer and Lemeshow 1989). The fitted logistic regression provides an equation that estimates the probability that a job is high-risk versus low-risk as a function of the workplace variable considered. Some jobs were represented in the database by more exposure data than others. This could lead to a misinterpretation (bias) of risk. In order to minimize this type of bias each job was weighted proportionally to the number of person-hours from which the injury rates were derived. Thus, jobs with more exposure data were weighted heavier in the analysis. Since the variables have very different scales, a useful summary of the fitted model is an odds ratio.

In order to be clear as to the meaning of the odds ratio in this study, an explicit definition of odds was necessary. In this case the odds is the probability of a high-risk job over the probability of a low-risk job. In general, the odds ratio is a statistical method for comparing the odds of any two risk situations (Hosmer and Lemeshow 1989). In this case, the authors report the odds ratios comparing the odds of mean high-risk jobs versus the odds of mean low-risk jobs for several individual workplace measures. This type of analysis was used because in the authors' opinion the dichotomous risk classification is more relevant to workplaces than the exact values of injury rate. In addition, the descriptive statistics of the injury rate showed that this variable was so skewed that ordinary regression analyses using injury rate would not be appropriate.

Second, a similar validity measure was derived for the combination of workplace variables used by each of the NIOSH *Guides*. Multiple logistic regression was used to estimate the odds ratio for the odds of mean high-risk versus the odds of the mean low-risk, as a function of the values of several workplace factors. A multiple logistic regression model relies on the hypothesis that the logarithm of the odds that a job is high-risk versus low-risk is a linear function of all the workplace variables.

Next, the 1993 NIOSH lifting index was calculated for each of the jobs in the database. For comparison purposes a lifting index for the 1981 *Guide* was created by considering the weight lifted relative to the action limit. When the weight lifted was equal to the action limit the index had a value of 1.0 and when the weight equalled the maximum permissible limit the index assumed a value of 3.0. Univariate analysis was performed on the low-, medium- and high-risk historical data to determine the number of jobs with lifting index < 1, 1 – 3, and > 3 in each historical risk group. In order to see how well the two NIOSH equations identified risk of LBD for the lifting jobs relative to the 'historical' database, the authors compared jobs with lifting index values < 1 to the low-risk group, jobs with lifting index values from 1 – 3 to the medium-risk group, and jobs with a lifting index > 3 to the high-risk group. Thus, the results of this analysis would show how well the two NIOSH lifting equations identified risk as defined by this database.

Finally, psychophysically assessed risk was evaluated by comparing the percentage of jobs that this evaluative tool would judge to be acceptable to 75% of the females as a function of each risk category. It was assumed that as risk increases fewer jobs would be acceptable from a psychophysical standpoint.

3. Results

3.1. *NIOSH Guides*

The ability of the individual workplace factor multipliers used by the two NIOSH measures to account for high-risk versus low-risk jobs was assessed first. The results of the simple logistic analysis are shown in table 1 for the factors used in both the 1981 and 1993 NIOSH measures as well as for other potentially predictive workplace factors. Of the factors included in these *Guides* the average weight of the object lifted and vertical destination of the lift (used to compute vertical travel distance) produced a statistically significant odds ratio. Many of the factors that have been generally

Table 1. Single workplace factor odds ratios between high-risk versus low-risk of developing LBD using the 1983 NIOSH *guide* and 1993 NIOSH lifting equation.

NIOSH guide variables 1981	NIOSH guide variables 1991	Workplace factor odds ratios: Workplace factor	Mean high-risk (SD)	Mean low-risk (SD)	Odds ratio	Coefficient	Standard error of coefficient	95% Confidence interval
X	X	Lift rate	187.2 (210.0)	183.3 (310.6)	1.00	0.00005	0.0004	0.99 – 1.01
X	X	Vertical origin (m)	0.993 (0.19)	1.031 (0.27)	1.02	– 0.6748	0.5636	0.98 – 1.07
		Vertical destination (m)	**1.021 (0.22)**	**1.129 (0.26)**	**1.23**	**– 1.8747**	**0.5817**	**1.08 – 1.39**
X	X	Vertical distance (m)	0.240 (0.16)	0.275 (0.22)	1.03	– 0.8702	0.6742	0.99 – 1.07
		Average weight (N)	**97.4 (81.3)**	**30.6 (57.2)**	**2.76**	**0.0152**	**0.0027**	**1.94 – 3.93**
X	X	Average horizontal distance	0.664 (0.11)	0.646 (0.17)	1.01	0.7808	0.8838	0.99 – 1.02
		Average moment	**62.3 (50.7)**	**17.3 (29.5)**	**4.08**	**0.0313**	**0.0050**	**2.62 – 6.34**
	X	Asymmetry	31.6 (28.9)	35.3 (25.4)	1.02	– 0.005	0.00497	0.98 – 1.06
	X	Coupling	0.90 (0.0)	0.90 (0.0)	1.00	0	0	1.00 – 1.00

Bold numbers indicate statistically significant odds ratios.

accepted as risk indicators did not result in significant odds ratios in this analysis. Interestingly, the mean high-risk value for vertical destination is lower than the mean low-risk value, which is reflected in a negative logistic regression coefficient for several of the workplace factors.

Table 2 shows the ability or potential of the factors used by the 1981 *Guide* to be associated with or identify high-risk versus low-risk situations when the variables used in the *Guide* were considered collectively. This table indicates that a statistically significant odds ratio of 3.5 is achieved when all five variables comprising the 1981 guide are considered collectively. When these analyses were performed with maximum horizontal distance instead of the average horizontal distance the odds ratio increased to 4.6. It is remarkable to note that this model yields odds ratios that are statistically similar to those produced when considering load moment alone. Table 1 indicated that load moment by itself yields an odds ratio of 4.08 (5.17 if maximum horizontal distance is used).

Table 3 shows a similar analysis for the 1993 *Revised Equation*. When the factors in the 1993 *Revised Equation* were considered collectively the odds ratio was found to be 3.1. This odds ratio improved to 4.3 when the maximum horizontal distance was

Table 2. Multiple logistic regression model odds ratio for high-risk versus low-risk group of developing LBD using the 1981 NIOSH *guide*.

Variables	Coefficient	Standard error of coefficient	Wald Score*
Constant	– 1.4549	1.3197	– 1.102
Lift rate	0.0005	0.0005	1.000
Average box weight	0.0169	0.0028	6.030
Average horizontal distance	2.4790	1.1255	2.202
Vertical load distance	– 1.6244	0.8428	– 1.927
Vertical start height	– 0.9972	0.7841	– 1.271
Odds ratio	**3.5**		
Confidence interval	(2.8 – 4.3)		

*Coefficient/standard of coefficient.

Table 3. Multiple logistic regression model odds ratio for high-risk versus low-risk group of developing LBD using the 1993 NIOSH lifting equation.

Variables	Coefficient	Standard error of coefficient	Wald Score*
Constant	– 1.4876	1.3214	– 1.125
Lift rate	0.0003	0.0006	0.500
Average box weight	0.0147	0.0027	5.444
Average horizontal distance	3.0419	1.2319	2.469
Vertical load distance	– 1.7669	0.8524	– 2.072
Vertical start height	– 1.1585	0.7883	– 1.469
Asymmetry	– 0.0041	0.0062	– 0.661
Odds ratio	**3.1**		
Confidence interval	(2.6 – 3.8)		

*Coefficient/standard of coefficient.

used in the analysis instead of the average horizontal distance for the task. Tables 2 and 3 show that the difference between the 1981 and 1993 *Guides* evaluated was the addition of the asymmetry variable to the 1993 *Guide*. To evaluate whether or not the addition of the asymmetry significantly influenced the model, the authors examined the Wald Score of the asymmetry variable. According to Hosmer and Lemeshow (1989) the Wald Score should be at least 2.0 in order for the variable to be significant at the 0.05 level. Table 3 shows that the Wald Score for the asymmetry variable is – 0.661, therefore the asymmetry variable does not significantly contribute to the model. Thus, one could conclude that the effectiveness of these two NIOSH measures is approximately equivalent. It is interesting to note that the only variables that significantly contribute to both the 1981 and 1993 models using the Wald Score criteria are the average box weight and average horizontal distance. These two variables are combined in average moment, which had the highest odds ratio in table 1. It should also be noted that handle condition was not evaluated in this analysis. However, it is felt that this would not alter the odds ratios considerably since the discounting factor for the handle factor is small.

Odds ratios indicate the power of the factor or combination of factors to identify high-risk (for LBD) versus low-risk manual materials handling situations. However, they do not provide insight as to the types of situations where the measures perform well compared to the situations in which the measures do not perform well. Therefore in order to assess the ability of the NIOSH measures to correctly identify low- or high-risk situations a univariate analysis was performed. The ability of the 1981 *Guide* to correctly identify 353 high-, medium-, and low-risk manual materials handling jobs based on the lifting index is shown in table 4. The lifting index row in this table relates to a lifting index similar to that found in the 1993 *Lifting Equation*. Table 4 shows that the 1981 *Guide* did an excellent job of identifying the low-risk jobs. Over 91% of the low-risk jobs were correctly identified indicating high specificity. However, only 10% of the high-risk jobs were correctly classified indicating low sensitivity. Most of the high-risk jobs were identified as low-risk with the lifting index. In fact 57% of the high-risk jobs were identified as low-risk in the 1981 *Guide*. This *Guide* also correctly identified 43% of the medium-risk jobs. Here again the majority of jobs (52%) were misidentified as low-risk. Thus, when the 1981 *Guide* misidentified jobs it did so by identifying the job as low-risk.

Table 4. 1981 NIOSH *guide* lifting index risk versus historical risk.

	NIOSH lifting index		
Historical risk	Lifting index < 1	1 < lifting index < 3	lifting index > 3
Low risk (124)	**91% (113)**	9% (11)	0% (0)
	LI= 0.19	LI= 1.85	LI= ***
	Risk= 0	Risk= 0	Risk= 0
Medium risk (118)	52% (61)	**43% (51)**	5% (6)
	LI= 0.41	**LI= 1.79**	LI= 4.8
	Risk= 3.9	**Risk= 6.4**	Risk= 4.0
High risk (111)	57% (63)	33% (37)	**10% (11)**
	LI= 0.41	LI= 1.67	**LI= 7.1**
	Risk= 26.7	Risk= 25.3	**Risk= 27.2**

***LI not calculated due to sample size.
Bold type indicates correct identification.

A similar univariate analysis was performed for the 1993 *Lifting Equation* and is shown in table 5. This analysis indicated that the *Lifting Equation* did a reasonable job of identifying high-risk jobs indicating increased sensitivity. A total of 73% of the high-risk jobs were correctly classified as such. However, 45% of the low-risk jobs were misidentified indicating marginal specificity. Of particular interest was the fact that nearly one-quarter of the jobs that had never experienced a low-back disorder incident were identified as high-risk by this lifting equation. In fact as shown in table 5 the average lifting index for these misidentified jobs was 9.9. Another 22% of the low-risk group were labelled as medium-risk. Of the medium-risk jobs the lifting equation classified 21% of those correctly. Over two-thirds of the medium-risk jobs were identified as high-risk jobs with the lifting equation. As shown in table 5, very few jobs were misidentified into a lower risk category. Thus, even though the ability of the 1993 lifting equation to predict risk was roughly equivalent to that of the 1981 *Guide* as evidenced from the odds ratios, the two measures misidentified jobs in very different ways. The 1981 *Guide* achieved high specificity because it was more liberal in the assessment of risk by misidentifying most jobs as safe. The 1993 lifting equation, on the other hand, achieved high sensitivity because it is more conservative and identified most jobs as being risky.

3.2. *Psychophysically determined risk*

Since the logic behind psychophysically determined job design is significantly different from that of the NIOSH measures it was not possible to evaluate the power of the model components via odds ratios as was done with the NIOSH measures. According to psychophysical logic, jobs should be designed so that 75% of the females should consider the task acceptable to them. Therefore, another means to assess the ability of psychophysical methods to correctly assess risk is to determine whether this acceptable lift criteria distinguished between the high-, medium-, and low-risk jobs. Table 6 shows the results of such an analysis. According to this table, 60% of the high-risk jobs would be acceptable to 75% of the females, whereas 64% and 91% of the medium- and low-risk jobs, respectively, would be acceptable to 75% of the females. Thus, there appear to be two levels of acceptability. High- and medium-risk jobs appear to be acceptable to almost two-thirds of women whereas low-risk jobs would be acceptable to about 90% of women. Hence there appears to

Table 5. 1993 NIOSH *guide* lifting index risk versus historical risk.

	NIOSH lifting index		
Historical risk	Lifting index < 1	1 < lifting index < 3	lifting index > 3
Low risk (124)	**55% (68)**	22% (27)	23% (29)
	LI = 0.45	LI = 1.67	LI = 9.9
	Risk = 0	Risk = 0	Risk = 0
Medium risk (118)	11% (13)	**21% (25)**	68% (80)
	LI = 0.38	**LI = 1.90**	LI = 10.9
	Risk = 4.2	**Risk = 3.3**	Risk = 5.6
High risk (111)	8% (9)	19% (21)	**73% (81)**
	LI = 0.69	LI = 1.72	**LI = 11.3**
	Risk = 34.2	Risk = 24.9	**Risk = 25.9**

Bold type indicates correct identification.

Table 6. Percentage of population finding jobs acceptable in high-, medium- and low-risk groups based upon psychophysical criteria.

Percentage of population that can perform job	Number of jobs	Cumulative distribution
High risk		
90	57 (51%)	57 (51%)
75	10 (9%)	67 (60%)
50	13 (12%)	80 (72%)
25	8 (7%)	88 (79%)
10	23 (21%)	111 (100%)
Medium risk		
90	62 (52%)	62 (52%)
75	14 (12%)	76 (64%)
50	13 (11%)	89 (75%)
25	9 (8%)	98 (83%)
10	20 (17%)	118 (100%)
Low risk		
90	108 (87%)	108 (87%)
75	5 (4%)	113 (91%)
50	5 (4%)	118 (95%)
25	0 (0%)	118 (95%)
10	6 (5%)	124 (100%)

be little distinction between the high- and low-risk acceptability. It is also interesting to note that nearly two-thirds of the high- and medium-risk jobs would be acceptable by psychophysical criteria yet there were significant rates of injuries associated with these jobs. Thus, the psychophysical criteria appear to be similar to the 1981 NIOSH *Guide* in the way that it misidentifies risk. It misidentifies risk in such a manner that it errs on the side of calling a job low-risk.

4. Discussion

The analyses used in this investigation have employed an independent data set to evaluate the efficacy of several lifting assessment methods. Use of an independent data set that is sufficiently large, such as this set, and different from any data sets used to develop the lifting assessment measure ensures an unbiased estimate of the usefulness of the tools. These analyses have indicated that all three tools do indeed have predictive power to identify jobs associated with high-risk of occupationally-related low-back disorders. Both the 1981 and 1993 NIOSH measures were found to have odds ratios for high risk versus low risk of LBD between 3.1 and 3.5 when average moment arms were considered and between 4.3 and 4.6 when maximum horizontal moments are considered. According to Waters *et al.* (1993), 'many' workers would be at risk of LBD if the lifting index was above 3.0 compared to if the index was less than 1.0. Thus, these analyses indicate that the NIOSH approaches to indeed have predictive power for identifying jobs that place workers at risk of devloping LBD.

For comparison purposes, the predictive power of the load moment supported by the trunk was evaluated. This factor produced an odds ratio for high-risk versus low-risk of LBD of 4.08 (C.I. 2.62 – 6.34) using average moment arm distance (horizontal distance) and 5.17 (C.I. 3.19 – 8.38) when maximum horizontal distance was used to compute moment. These values were significantly greater than any other single measure. In addition, this odds ratio was greater than the odds ratio produced by either the 1981 or 1993 NIOSH *Guide* models. For the case of the jobs observed in this study knowledge of this simple factor alone would enable one to identify those jobs that were high risk versus low risk of LBD as well as the more computationally intensive measures. However, the authors should caution that this may not hold true for jobs with work parameters that are outside the range of those observed in this study. For example, the database did not involve jobs where lifting was performed from an extremely low or high vertical location. Thus, the authors can not evaluate how well the NIOSH approach would work in identifying risk in those positions. However, it is expected that it would perform much better than simple load moment in these situations.

This analysis has also indicated how the two NIOSH measures behave when they do not correctly identify work-related risk according to our risk criteria. Figures 1 and 2 show the relationship between the lifting index and back injury incidence rate for the 1981 *Guide* and 1993 *Lifting Equation*, respectively. As shown in figure 1 by the large cluster of points associated with the low lifting index jobs, the 1981 *Guide* misidentifies risk by predicting that most jobs will be low risk, thereby under-estimating the risk. Figure 2 shows a high density of jobs with a low incidence rate that are spread throughout the lifting index range indicating that the 1993 lifting equation misidentifies risk by predicting that jobs will be high-risk, hence, overestimating risk. Consequently, these two methods misclassify in exactly opposite directions.

A possible explanation for why the NIOSH 1993 LI values above 3 might be classified as members of the 'low-risk' group might have to do with the density of

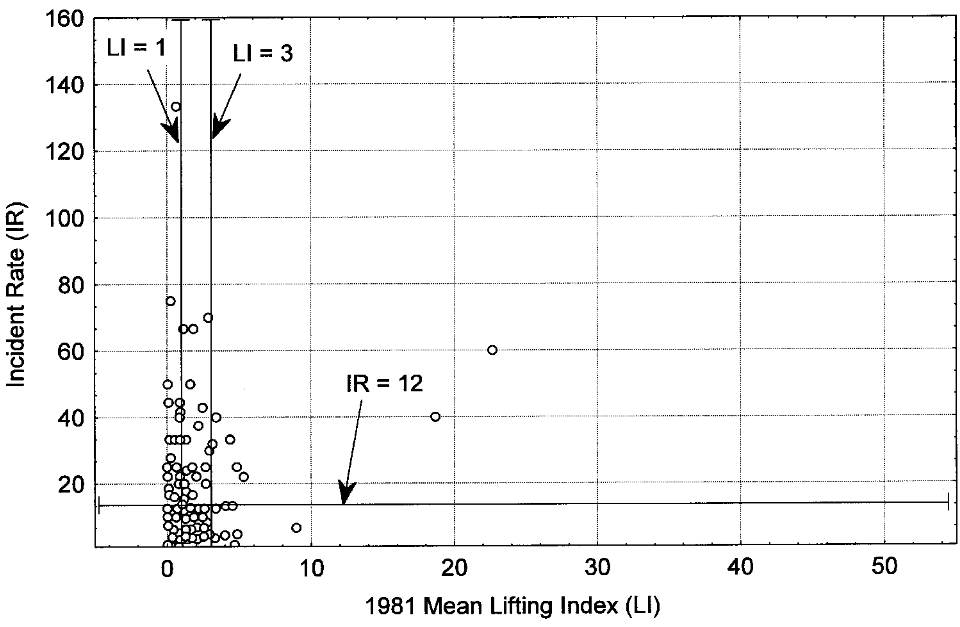

Figure 1. 1981 NIOSH *Guide* lifting index compared to historical incident rate.

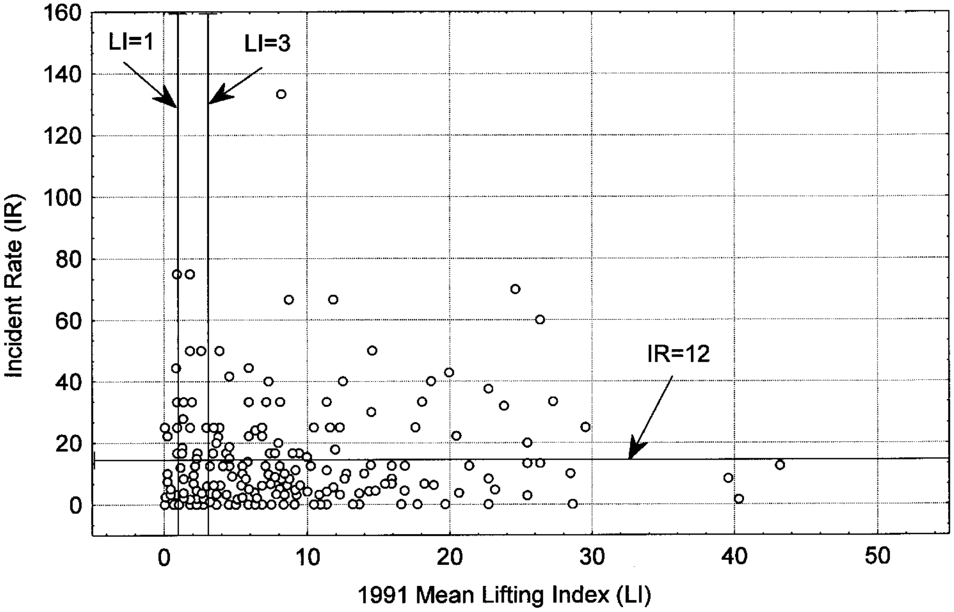

Figure 2. 1993 NIOSH lifting equation lifting index compared to historical incident rate.

exposure data available for a given job. It is possible that 'survivors' in a job might mask the potential risk for the group of workers exposed unless very large groups of workers exposed to the same exact job can be observed. However, extremely large exposure to a particular job is often a rare event for industrial work. This type of misclassification was less likely in the multivariate analysis because the rates were weighted by the number of person years of observations for each job. Another possible explanation for the mismatch of the 1993 NIOSH equation might be due to the interpretation of an LI value of 1. The original intent of an LI of 1 was to protect the vast majority of potential workers. However, realistically, one would expect that workers with above average risk of back injuries are probably not placed in highly risky jobs. This situation would suggest a selection process for some jobs.

None the less, one can gain insight as to the nature of these systematic misclassifications in an attempt to identify how future forms of such assessment methods might be improved. If one considers the form of the lifting equations used in these two NIOSH measures one can gain insight as to why the predictions behave as they do. Table 1 demonstrated that the odds ratio indicating the ability of the average load moment alone to identify a high-risk job was greater than the predicted odds ratio for either the 1981 NIOSH *Guide* or the 1993 NIOSH *Lifting Equation*. To be precise, the Wald Score in tables 2 and 3 indicate that lift rate, vertical load, vertical start height, and asymmetry do not significantly contribute to the model. The average weight of the box and average horizontal distance are the only two variables that significantly contribute to both the NIOSH models using the Wald Score criteria. The load moment is embedded into both NIOSH tools in the form of a load constant and a horizontal multiplier. Thus, the strength of the predictions could all be considered equivalent. However, this situation begs the question as to why lift rate, vertical load, vertical start height, and asymmetry do not significantly contribute to the predictive power of the NIOSH model.

There may be two explanations for this situation. First, examination of the database used in this analysis indicated that there was not a great difference in the magnitude of the vertical travel distances and asymmetries between the high- and low-risk workplaces. Therefore, these variables did not distinguish well between the high- and low-risk situations. Had jobs with more extreme vertical and asymmetric motions been observed, these factors may have indeed become significant in the analysis. However, in the development of this database the authors did not attempt to ensure that all extremes of the various workplace factors were represented in the database. The authors simply observed whatever workplace situations were present in the repetitive jobs in which they were able to collect data. This situation might imply that the NIOSH tools would be sensitive in identifying the more extreme lifting situations.

Second, it may be the case that the form of the functional multipliers may need improvement. The logic for including these multipliers cannot be disputed. For example, the inclusion of a biomechanical length-strength modifier or an asymmetry modifier to reduce the allowable load is certainly justified from a biomechanical, physiological, or epidemiologic standpoint. However, it can be submitted that the characterization of the multiplier can be improved as described in a work situation. It has been well established that when a muscle is at its resting length the muscle can produce the greatest amount of force and when the muscle is lengthened or shortened the amount of force a muscle can produce decreases monotonically with an increase or decrease in muscle length. This trend is represented primarily by way of the vertical multiplier, which is similar to an inverted 'V' thereby approximating the physiologic descriptions shown in the literature. This relationship has been described with isolated joints such as the elbow, knee, etc. However, under realistic manual materials handling conditions it is suspected that the trunk does not behave as an isolated joint and the form of the length-strength relationship is not described well by the inverted 'V' function. Perhaps when one lifts from a low (or high) vertical position the natural behaviour of a person is to rock upon the femoral bones, thereby bending at the hip joint, as opposed to bending the trunk and lengthening the back muscles. If this is true then the vertical multiplier's magnitude and shape should be represented by something other than an inverted 'V' shape. This inverted 'V' function is probably overestimating the risk. Similar arguments could be made for the asymmetry multiplier. When workers twist, this action is most likely performed using the hip and knee joints for the first several degrees of twist as opposed to the twisting motion occurring at the back (Ferguson 1990). Hence, the 1993 lifting equation might place too severe a discounting penalty on a twisting motion since the magnitude of the twist in the back has been over-predicted. Such an error could explain why the 1993 lifting equation over-predicts the risk to a worker performing a manual materials handling task. Hence, there appears to be a potential for the NIOSH measures to become more predictive if they better describe the functional nature of the multipliers. More focused studies are needed to determine which one of these explanations is correct.

It is interesting to note that the psychophysical assessment of the jobs show a similar trend to that of the NIOSH measures. Hidalgo *et al.* (1995) have also acknowledged this association. The present analyses indicated that as the risk increases fewer workers would consider the jobs to be acceptable based upon the psychophysical criteria. Hence the psychophysically established guides do also appear to be associated with risk of occupationally-related low-back disorder. However, as with the NIOSH measures the nature of the relationships do appear to deviate significantly from what would be expected. For example, why would 60% of the jobs with a documented history of high-risk for low-back disorder be judged as acceptable to 75% of the female population if such measures were true assessment methods? It is contended that the psychophysical relationships are indicative of potential back problems, however, the authors also contend that the sensitivity of such a method may not be adequate to account for the majority of occupationally-related low-back disorders.

Several potential limitations of this study must also be considered. First, as mentioned previously, a database can always mischaracterize risk. The definition of LBD risk itself is controversial. Different results might have been observed if different outcomes had been used to define risk, such as self-reports of LBD, workers' compensation claims, disability etc. In addition, different definitions of risk might have led to different results. In addition, the cut-off values of what one might consider high-, medium-, or low-risk might vary significantly among various organizations and researchers. The authors chose a definition that had meaning (relative to the percentile of risk) in their larger industrial database. It should be acknowledged that this methodology might be inefficient in that the definition of risk is dependent upon the larger database mixture of jobs from which the risk percentile was derived, and, therefore might be considered somewhat arbitrary. However, it can easily be argued that the cut-off of zero for low-risk is certainly appropriate for low-risk since no injuries occurred in this group. An incident rate of 12 per 200 000 h of exposure not only matches the 75th percentile of risk from the ongoing database but also represents a value that is three times greater than the US national average for lost time in industry according to the Bureau of Labor Statistics (as of 1990). Thus, it is likely that this value also has meaning to industry as a high risk cut-off point.

Next, as discussed earlier, there is a potential for exposure bias from several sources in these types of studies. For example, if the number of person-years of exposure for a job is not sufficiently large it may bias the characterization of risk. However, in order to minimize this possibility the authors weighted the analyses as a function of the richness of exposure data to lend more credibility to the larger more robust job histories. Third, the risk characterization might be affected by the healthy worker effect (Andersson 1991) or the fact that some jobs might employ survivors who don't get injured or simply don't report injuries. This might be an alternative explanation of why some jobs had no history of LBD yet had large lifting indices. Another possible limitation of this data set would be the reliance on injury logs compared to active surveillance. However, in order to minimize these types of recording errors company records were scrutinized. In addition, it is believed that active surveillance does not have the ability to perform any better since worker memories fade over the years and LBD is not always easily documented. All of these possible biases should be minimized in this study since the database was very large. There are over 21 million person-hours of exposure data represented in this database. Certainly, isolated cases of misclassification of risk would not lead to erroneous conclusions in this situation.

A concern of this study might be that different methods were used to measure some of the workplace variables than those recommended by NIOSH. In particular, horizontal distance was measured from the estimated point of the spine instead of from a point bisecting the ankles. However, since these were constant differences they should not affect the analyses. In addition, these two different measures would be expected to be quite similar since the NIOSH method is intended to approximate spine location.

A final issue to be discussed is that of the part of the body affected by the work. As mentioned earlier, the 1981 NIOSH *Guide* as well as the psychophysical analysis methods are intended to control for musculoskeletal disorders in general whether they are backs, shoulder, lower extremities, or wrists. The focus of this analysis was to determine how well these analysis tools do in identifying risk to the low back specifically. Thus, the intent of these tools may be different. Since back problems are the most costly musculoskeletal problem facing society the authors were specifically concerned with how well the back might be protected via these tools.

It should be noted that other promising lifting analysis and assessment methods have been developed in recent years. Marras *et al.* (1993) developed a multiple logistic regression risk model that includes the effects of trunk motion and workplace factors on risk and has reported an odds ratio of 10.7 (CI = 4.9– 23.6) using this same database. Subsequent studies using an independent data set have indicated that classification results are excellent (Marras 1997). However, it is imperative that such techniques be validated by independent studies as was done in the present study. However, this risk analysis requires intrumentation that may not be available to one performing the job assessment.

Collectively this study implies that there are situations where the various low-back assessment methods have various strengths and weaknesses. Some of the more straightforward tools such as the NIOSH *Guide* or lifting equation have the advantage of being easy to use. These measures might perform well in situations where the tool's assumptions are not violated. For example, they may be most appropriate when the lift is static or slow and significant three-dimensional dynamic motion is not involved in the task. However, if the job is dynamic and motion is suspected of being a problem then tools such as the LMM risk model (Marras *et al*. 1993, 1995) might be necessary. However, this technique requires intrumentation of the worker.

These analyses have shown that it is imperative to examine the validity and effectiveness of the tools that are used to evaluate and control the risk of low-back disorders associated with the workplace. It is important that independent validations be performed and that one critically evaluates lifting control tools so that one facilitates an understanding of how future analyses and controls might be altered so that they become even more powerful means by which one controls the risk of occupationally-related injuries.

Acknowledgements

Partial funding for this study has been supplied by the National Institute for Disability and Rehabilitation Research.

References

Andersson, G. B. J. 1991, The epidemiology of spinal disorders, in J. W. Frymoyer, T. B. Ducker, N. M. Hadler, J. P. Kostuik, J. N. Weinstein and T. S. Whitecloud (eds), *The Adult Spine* (New York: Raven Press), 107 – 146.

Cronbach, L. J. 1970, *Essentials of Psychological Testing*, 3rd edn (New York: Harper & Row).

Ferguson, S. 1990, Quantification of back motion during asymmetric manual material handling tasks. Thesis, The Ohio State University.

Hidalgo, J., Genaidy, A., Karwowski, W., Christensen, D., Huston, R. and Stambough, J. 1995, A cross-validation of the NIOSH limits for manual lifting, *Ergonomics* **38,** 2455 – 2464.

Hosmer, D. W. and Lemeshow, S. 1989, *Applied Logistic Regression* (New York: Wiley).

Kelsey, J. L. and White, A. A. III, 1980, Epidemiology and impact on low back pain, *Spine* **5**(2), 133 – 142.

Liira, J. P., Shannon, H. S., Chambers, L. W. and Haines, T. A. 1996, Long-term back problems and physical work exposures in the 1990 Ontario Health Survey, *American Journal of Public Health*, **86,** 382 – 387.

Marras, W. S. 1997, The validity of the LMM risk model. Paper presented at the International Ergonomics Association meeting. Tampere, Finland, July.

Marras, W. S., Lavender, S. A., Leurgans, S., Fathallah, F., Allread, W. G., Ferguson, S. A. and Rajulu, S. 1995, Biomechanical risk factors for occupationally related low-back disorder risk, *Ergonomics*, **38,** 377 – 410.

Marras, W. S., Lavender, S. A., Leurgans, S. E., Rajulu, S. L., Allread, W. G., Fathallah, F. A. and 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.

National Safety Council 1989, *Accident Facts 1989* (Chicago, IL: National Safety Council).

NIOSH 1981, *Work Practices Guide for Manual Lifting*, DHSS (NIOSH) Publication no. 81 – 122. (Cincinnati, OH: NIOSH).

Ohio Bureau of Workers' Compensation, 1994, Personal communication, 13 May 1994. Occupational Safety and Health Research, Division of Safety and Hygiene.

Pope, M. H. 1989, Risk indicators in low back pain, *Annals of Medicine*, **21,** 387 – 392.

Praemer, A., Furner, S. and Rice, D. P. 1992, *Musculoskeletal Conditions in the United States* (Park Ridge, IL: American Academy of Orthopaedic Surgeons).

Snook, S. H. 1978, The design of manual handling tasks, *Ergonomics*, **21,** 963 – 985.

Snook, S. H. 1989, The control of low back disability: the role of management, *Manual Materials Handling: Understanding and Preventing Back Trauma* (Akron, OH: American Industrial Hygiene Association).

Snook, S. H. and Ciriello, V. M. 1991, The design of manual handling tasks: revised tables of maximum acceptable weights and forces, *Ergonomics*, **34,** 1197 – 1213.

Snook, S. H., Campanelli, R. A. and Hart, J. W. 1978, A study of three preventive approaches to low back injury, *Journal of Occupational Medicine*, **20,** 478 – 481.

Waters, T. R., Putz-Anderson, V., Garg, A. and Fine, L. J. 1993, Revised NIOSH equation for the design and evaluation of manual lifting tasks, *Ergonomics*, **36,** 749 – 776.

Waters, T. R., Putz-Anderson, V., Garg, A. and Fine, L. J. 1994, *Applications Manual for the Revised NIOSH Lifting Equation*, DHHS (NIOSH) Publication No. 94 – 110. (Cincinnati, OH: NIOSH).

Wegman, D. H. and Fine, L. J. 1996, Occupational and environmental medicine, *JAMA*, **275,** 1831 – 1832.