Original article:

DEVELOPMENT AND VALIDATION OF A NEW SCALE FOR PREDICTION OF LOW BACK PAIN OCCURRENCE AMONG NURSES

Mohammad Javad Jafari, Forough Doshman Fana Yazdi*, Yadollah Mehrabi, Sakineh Rakhshanderou, Mahnaz Saremi

School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

* Corresponding author: School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran. E-mail: forfarvardin@gmail.com, Tel: +98(2122432041), +98(9131564257), Fax: +98(2122432037)

http://dx.doi.org/10.17179/excli2019-1167

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).

ABSTRACT

There are several scales for prediction of low back pain (LBP) occurrence, but most of them only consider occupational aspect. This study aimed to develop and validate a new biopsychosocial scale for the LBP prediction among nurses. In this mixed-method study, a scale was developed by integrating the findings from the literature review and the semi-structured interviews. The qualitative and quantitative face and content validation were then performed. The construct validation was performed based on the hypothesis testing by independent-samples t-test using the SPSS in a case study with 241 nurses. The reliability of the scale was also tested through 15-day interval test-retest reliability by Intra Class Correlation Coefficient (ICC). Then the Minimum Detectable Changes (MDC) and MDC % was calculated. The results showed that the three dimensions (occupational, psychosocial and individual), consisted of 40 items, predict LBP occurrence. Both the scale and the three sub-scales could differentiate known groups of nurses in terms of LBP. These groups were nurses: with/without LBP during the past 12 years, with a high/low occurrence of LBP, with/without co-morbidity, being female/male, with/without night shift, and with high/low repetition of loads/patients handling. The average measure ICC of the scale was 0.866 (P <0.001). The MDC95 (MDC %) was 14.86 (15.65 %). We concluded that the proposed scale is a simple and trustworthy tool which supports the multidimensional nature of LBP.

Keywords: Low back pain, ergonomic assessment, nursing personnel, biopsychosocial factors, scale development

INTRODUCTION

Low Back Pain is one of the most common disorders among workforces (Piravneyseh et al., 2016; Choobineh et al., 2013; Motamedzade et al., 2013; Mohammadi et al., 2013) especially among nurses, which can lead to a decreased productivity. The decreased quality of services delivered to the patients, medication errors, etc., also the increased rate of absenteeism, as well as, direct and indirect costs are some examples of the agents affect on the productivity (Yip, 2001; Asadi et al., 2016). According to several recent studies, the prevalence of LBP among nurses in Iran is much higher (over 50 %) (Asadi et al., 2016; Sadeghian et al., 2014; Mehrdad et al., 2010) than the general population (29.3 %) (Biglarian et al., 2012). It is noteworthy that it is not possible to introduce only one single agent as the cause of LBP occurrence, because of the multidimensional
nature of LBP, and the variability in the tasks performed at the workplaces. From the ergonomics point of view, the treatment of LBP has been often unsuccessful (Bakker et al., 2009; Deyo et al., 2014) and costly. So, prevention would be a more effective approach to LBP management (Bakker et al., 2009; Cohen et al., 2008).

To manage LBP in the workplace more effectively, we need LBP predictive scales which consider all of the effective aspects (Koes et al., 2010). Several ergonomics scales are available for LBP assessment and prediction, such as The American Conference of Governmental Industrial Hygienists method (ACGIH, 2017), National Institute for Occupational Safety and Health equation (NIOSH) (Waters et al., 1993), etc. (David, 2005). The common weakness of these scales is that they only consider the occupational aspect without addressing other effective aspects such as psychosocial ones (David, 2005). Moreover, it is believed that psychosocial risk factors influence LBP occurrence through interaction with the work environment and individual characteristics (Abedini et al., 2015; Dehghany et al., 2012). Therefore, it is highly recommended that the researchers pay more attention to biopsychosocial factors (Mitchell et al., 2009; George et al., 2012; Dunn et al., 2013; Dagenais et al., 2010) in designing the screening scales (Pincus et al., 2002).

In a biopsychosocial model, none of the individual, occupational and psychosocial dimensions can explain LBP singly (Mitchell et al., 2009), and for a comprehensive explanation of LBP occurrence, the interaction within and between these dimensions should also be considered (Marras, 2005). To the best of our knowledge, no such study is available in Iran. Hence, this study is intended to develop and validate a new biopsychosocial scale for the LBP prediction among nurses.

**MATERIAL AND METHODS**

This mixed-method study was conducted on nursing staff from June 2017 to August 2018 in Yazd, Iran, through the qualitative and quantitative methods as follows:

**Literature review**

The LBP risk assessment models and scales, as well as, the entire body risk assessment scales in the low back section were studied. For this purpose, published papers from 1990 to 2017 were included in this study. Relevant papers were explored by searching Google Scholar, Science Direct, PubMed, NCBI, and Scopus databases. The LBP related specific terms (Low back pain, or occupational low back pain with nursing, nurses, prevalence, occurrence, predictor, incidence, prognosis, first episode, first onset, the risk factor) were used to explore all databases. Two independent researchers explored the databases to find the eligible papers. Once the qualified papers were identified, the abstract and the full-text of papers were reviewed by the researchers. The extracted factors were classified into categories according to their similarities.

**Qualitative study**

A qualitative study was carried out to confirm the agents obtained from the literature review.

**Participants**

The participants were experienced nurses in the nursing care fields. A purposive sampling technique was employed to choose the subjects. We communicated with the hospitals to introduce the volunteer nurses in order to identify their expertise. Subsequently, the main criteria to choose the eligible experts included:

1. Having at least 5 years of job tenure in the nursing care activities
2. Having a history of Low back pain during the last year
3. Having the willingness to participate in the interview
4. Having a negative history of LBP due to a specific causation such as trauma, tumor, skeletal anomalies, spine surgery, or pregnancy during the last year.
From 43 nurses, 29 subjects who met the inclusion criteria were invited to participate in the study. Details of the study, time and location of the interview were set with each expert.

**Data gathering**

A semi-structured interview approach was employed in order to gather the experience of experts. The interview was performed based on the guide obtained from the advisers for a good interview. An interview guide was provided with several questions and supplemented with the complementary questions during the interview sessions. Each interview session lasted 15 to 30 minutes. The informed consent was obtained from the interviewees for recording the interviews. After each interview, the recorded contents were listened thoroughly and the sentences were transcribed. Subsequently, the contents were derived, coded, and classified by careful studying of the transcripts. According to the basic theory of the study, coding and classifying the extracted concepts were carried out inductively. Data saturation was obtained after analyzing the 26 interviews which continued for all 29 subjects. The extracted codes were classified into the categories based on their similarities.

**Construction of the scale**

**Aggregating the qualitative study and literature review findings**

The findings of the qualitative study and the literature review were aggregated into a draft model. We also checked the draft to remove duplicated factors.

**Psychometric properties of the questionnaire**

- **Validity analysis**

  Validity means the degree to which a scale can correctly measure the target (Jafari et al., 2017). The validity of the proposed questionnaire was assessed as follows:

  - **Face validity**

    Face validity was measured qualitatively by sending a Persian version of the questionnaire to 45 experts (5 experts at the academic level and 40 nurses) and receiving their overall conception in responding to the statement content for all of the items.

    For quantitative face validity, an impact score was computed through a 5-point Likert scale, in which, the "always" scale received the score 5 and score 1 belonged to the "never" scale. To obtain the impact score, frequency (the percent of responses with the important score of 4 or 5) and item’s importance (importance of each item on a 5-point Likert scale) should be calculated. Item impact score was obtained by multiplying the item’s frequency and importance. The cutoff point to select the eligible items was calculated to be 1.5 (means that an item has the frequency of 50% and the importance of 3 on the Likert scale). The values below 1.5 were removed from the questionnaire (Zamanzadeh et al., 2015).

  - **Content validity**

    The questions in the prepared draft were categorized into the three sub-scales: individual, occupational, and psychosocial.

    According to the Lawshe table (Lawshe, 1975), 5 through 40 experts should be chosen based on the accessibility to the experts. In this study, fifteen experts related to each sub-scale (a total of 45 experts) were called to administer the content validation forms. Every specialist rated the ‘necessity of each item’ by selecting one of the three options ‘essential’, ‘useful but not essential’, or ‘not essential’. Content validity ratio (CVR) for each item was calculated based on the ratings, according to Equation 1:

    \[
    CVR = \frac{n_E - n}{\frac{n}{2}}
    \]

    where \(n\) represents the number of raters, and \(n_E\), represents the number of raters who considered the item “Essential”.

    Based on the Lawshe table (Lawshe, 1975) the acceptable CVR score was 0.49. Therefore, the items with scores lower than the cut point 0.49 were discarded. The scale’s content validity index (S-CVI) was determined by calculating the mean CVR for the total items remained in the scale.
(Lawshe, 1975). The values higher than 0.8 were accepted (Polit and Beck, 2010). The clarity and simplicity of the scale were investigated as well.

- **Construct validity**

  The scale was distributed between nurses of all 5 public Hospitals in Yazd province, Iran. The stratified random sampling approach was used for data sampling. It means that samples were taken from all wards, based on the nursing population occupied in each ward. The sample size was estimated to be at least 5 people per variable (Ebadi et al., 2017). Thus, the nurses who met inclusion criteria, whether having or not having LBP during the past 12 years formed the sample population.

  After collecting the distributed scales, the data were entered into SPSS software. It should be mentioned that the convergent and discriminant validity through explanatory factor analysis were obtained undesirable. Therefore, the explanatory factor analysis can't be a proper option to test the construct validity. In this situation, a scale should be tested by the proper hypotheses to approve the proven facts. The scores were calculated at both levels: the whole scale and the sub-scales. The validity of this scale was checked by hypothesis testing. Since the scale scores should be significantly different between the groups with different levels of LBP risk, the hypothesis testing should be able to reject the null hypothesis on the equality of means. The independent-samples t-test was used by SPSS software to compare means. The following groups were compared: the nurses with/without low back pain during the past twelve months, the nurses with high/low frequency of LBP, the nurses with/without comorbidity, female/male nurses, and night/day shift nurses.

- **Reliability testing**

  The reliability testing for the whole scale was performed using a test-retest method. The developed scale was administered by 30 nurses with a 15-day interval between test and retest. For every participant, the whole scale score was calculated at both test and retest stage by summing the scores of all items. Then the Intra Class Correlation Coefficient (ICC) (Dagenais et al., 2010) was estimated for the two scores by means of the SPSS (Version 20) in two-way mixed mode for absolute agreement. The results were interpreted based on the following criteria: 0.0–0.2 (low), 0.21–0.40 (fair), 0.41–0.60 (moderate), 0.61–0.80 (substantial), and 0.81–1.0 (almost perfect) (Sharif Nia et al., 2013).

  The absolute reliability of the scale was tested by the equations 2 and 3, respectively:

  \[
  \text{SEM} = SD_{\text{all testing scores}} \times \sqrt{1 - r} \quad \text{Eq. 2}
  \]

  \[
  MDC = 1.96 \times \sqrt{2} \times SEM \quad \text{Eq. 3}
  \]

  Here, SD is the standard deviation of all testing scores, and r is the coefficient of the test-retest reliability (ICC) (Lee et al., 2017).

  The MDC % was also calculated by equation 4:

  \[
  \text{MDC} = \text{MDC/mean} \times 100\% \quad \text{Eq. 4}
  \]

  Here, mean is the mean score of all trials. An MDC % of 30 % or less was considered acceptable (Azadi et al., 2015).

## RESULTS

Table 1 illustrates the demographic information about the participants in the cross-sectional study. The participants were mostly female (80.9 %) and married (81.7 %). The mean age of the participants was 35.7 ± 6.3 (range: 25-55). The mean job tenure of the participants was 11 years, ranging from 5 to 29 with an interquartile range of 5-29.

A total of 86 variables from the literature review and 36 variables from the qualitative study were identified. After removing duplicated items, 86 items remained in draft scale.

The results of quantitative face validity indicated that "impact scores" for all of the items were higher than 1.5. Hence, all of the items remained for the following steps. Most of the experts stated that they had no difficulty in reading and understanding the questionnaire items. According to the participated nurses, a few items needed to be modified to
enhance the face validity, so all of which were corrected.

Based on the Lawshe table, forty-six items were removed due to CVR scores lower than 0.49. The overall scale’s content validity (S-CVI) was measured to be 0.81.

Forty items remained in the questionnaire consisted of 9 psychosocial items, 12 occupational items, and 19 individual items. The minimum acceptable sample size was obtained 220 (By considering 5 samples per each item). But a total of 241 nurses was included in this study, which returned the questionnaires.

The null hypothesis on the equality of means of the whole scores was rejected between some known groups (Table 2).

In addition, the null hypothesis on the equality of means of the sub-scale scores was rejected between the known groups by the individual, occupational, and psychosocial sub-scales respectively (Tables 3, 4, and 5).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>Number</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Woman</td>
<td>195</td>
<td>80.9</td>
</tr>
<tr>
<td></td>
<td>Man</td>
<td>46</td>
<td>19.1</td>
</tr>
<tr>
<td>Marital status</td>
<td>Single</td>
<td>42</td>
<td>17.5</td>
</tr>
<tr>
<td></td>
<td>Married</td>
<td>196</td>
<td>81.7</td>
</tr>
<tr>
<td></td>
<td>Widow</td>
<td>2.0</td>
<td>0.8</td>
</tr>
<tr>
<td>Educational level</td>
<td>Technician</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Bachelor of science</td>
<td>224</td>
<td>92.0</td>
</tr>
<tr>
<td></td>
<td>Master of science</td>
<td>12</td>
<td>5.8</td>
</tr>
<tr>
<td>Employment status</td>
<td>Permanent job (Formal employment)</td>
<td>111</td>
<td>47.4</td>
</tr>
<tr>
<td></td>
<td>Semipermanent</td>
<td>22</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td>Contractual</td>
<td>58</td>
<td>24.8</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td>43</td>
<td>18.4</td>
</tr>
<tr>
<td>Job position</td>
<td>Supervisor</td>
<td>4</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Head nurse</td>
<td>21</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>Nurse</td>
<td>214</td>
<td>89.5</td>
</tr>
</tbody>
</table>

Table 1: Details of the participants involved in the cross-sectional study (n=241)

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean ± SD</th>
<th>P-value (2-tailed)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without LBP during past 12 months</td>
<td>89.21 ± 14.67</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>With LBP during past 12 months</td>
<td>97.83 ± 13.11</td>
<td></td>
</tr>
<tr>
<td>Low frequency of LBP (occurrence period longer than one month)</td>
<td>88.75 ± 14.98</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>High frequency (occurrence period shorter than or equal one month)</td>
<td>97.05 ± 13.13</td>
<td></td>
</tr>
<tr>
<td>Without co-morbidity</td>
<td>92.06 ± 13.75</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>With co-morbidity</td>
<td>99.15 ± 12.30</td>
<td></td>
</tr>
<tr>
<td>Without night shift</td>
<td>89.48 ± 14.61</td>
<td>0.001</td>
</tr>
<tr>
<td>With night shift</td>
<td>96.17 ± 14.35</td>
<td></td>
</tr>
</tbody>
</table>

* The significance level was considered at P <0.05
Table 3: The results of the hypothesis testing with the Individual sub-scale scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean ± SD</th>
<th>P-value (2-tailed)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without LBP during past 12 months</td>
<td>42.71 ± 6.51</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>With LBP during past 12 months</td>
<td>47.13 ± 6.13</td>
<td></td>
</tr>
<tr>
<td>Low frequency of LBP (occurrence period longer than one month) **</td>
<td>42.90 ± 6.43</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>High frequency (occurrence period shorter than or equal one month)</td>
<td>46.26 ± 6.43</td>
<td></td>
</tr>
<tr>
<td>Without co-morbidity</td>
<td>44.33 ± 6.4</td>
<td>0.004</td>
</tr>
<tr>
<td>With co-morbidity</td>
<td>47.08 ± 5.91</td>
<td></td>
</tr>
<tr>
<td>Without night shift</td>
<td>43.56 ± 6.61</td>
<td>0.014</td>
</tr>
<tr>
<td>With night shift</td>
<td>45.73 ± 6.65</td>
<td></td>
</tr>
</tbody>
</table>

* The significance level was considered at p<0.05
** The range was ‘Never’ through “Daily”

Table 4: The results of the hypothesis testing with the Occupational sub-scale scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean ± SD</th>
<th>P-value (2-tailed)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handling frequency (lower than or equal to 6 times) **</td>
<td>25.24 ± 5.93</td>
<td>0.033</td>
</tr>
<tr>
<td>Handling frequency (higher than 6 times)</td>
<td>31.00 ± 4.24</td>
<td></td>
</tr>
<tr>
<td>Without co-morbidity</td>
<td>24.56 ± 6.71</td>
<td>0.04</td>
</tr>
<tr>
<td>With co-morbidity</td>
<td>26.51 ± 5.58</td>
<td></td>
</tr>
<tr>
<td>Without night shift</td>
<td>23.31 ± 6.65</td>
<td>0.001</td>
</tr>
<tr>
<td>With night shift</td>
<td>26.12 ± 6.65</td>
<td></td>
</tr>
</tbody>
</table>

* The significance level was considered at p<0.05
** The frequency of handling loads/patients ranged between 0-12 times

Table 5: The results of the hypothesis testing with the Psychosocial sub-scale scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean ± SD</th>
<th>P-value (2-tailed)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without LBP during past 12 months</td>
<td>22.95 ± 5.24</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>With LBP during past 12 months</td>
<td>25.65 ± 5.83</td>
<td></td>
</tr>
<tr>
<td>Low frequency of LBP (occurrence period longer than one month)</td>
<td>22.84 ± 5.5</td>
<td>0.001</td>
</tr>
<tr>
<td>High frequency (occurrence period shorter than or equal one month)</td>
<td>25.46 ± 5.66</td>
<td></td>
</tr>
<tr>
<td>Without co-morbidity</td>
<td>23.73 ± 5.32</td>
<td>0.009</td>
</tr>
<tr>
<td>With co-morbidity</td>
<td>25.98 ± 6.41</td>
<td></td>
</tr>
<tr>
<td>Without night shift</td>
<td>23.26 ± 5.54</td>
<td>0.013</td>
</tr>
<tr>
<td>With night shift</td>
<td>25.13 ± 5.69</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>22.52 ± 4.51</td>
<td>0.025</td>
</tr>
<tr>
<td>Females</td>
<td>24.6 ± 5.84</td>
<td></td>
</tr>
</tbody>
</table>

* The significance level was considered at p<0.05
On the reliability testing stage, the average measure ICC was 0.866 with a 95% confidence interval from 0.687 to 0.943 (F = 7.38, P < 0.001). The SEM and MDC95 were 5.47, and 15.16 respectively. The MDC % was equal to 15.97 %.

The final scale for the prediction of LBP occurrence among nurses is shown in the Appendix (Supplementary material).

DISCUSSION

This study is intended to develop and validate a novel scale for the prediction of LBP occurrence among nurses. The developed scale consisted of the three sub-scales including individual, occupational, and psychosocial. The structural validity demonstrated that the scale could predict the risk of LBP well, because it was able to distinguish the known groups with different levels of LBP risk. Further, the scale is able to distinguish those who had LBP in the past 12 months and those who had not experienced LBP in the recent 12 months. The former group obtained higher scores than the latter group. The scale also distinguishes the nurses with high and low frequency of LBP (the former group received a higher score). Other groups that the scale is able to differentiate are presented in the following paragraphs.

Women versus men

The relationship between LBP and woman gender has been shown in various studies (Troussier et al., 1999; Schneider et al., 2006; Bejia et al., 2005; Wang et al., 2016; Yang et al., 2016). Wang et al. (2018) indicated that among all age groups, the prevalence of LBP is higher in women compared with men (Wang et al., 2018). Similarly, this developed scale is able to differentiate these two groups by giving higher scores of women. Wang et al. also identified the psychological factors as one of the possible causes for the higher prevalence of LBP in women compared with men (Wang et al., 2018). In this study, it was the psychosocial sub-scale that revealed the difference in prevalence of LBP between women and men.

Co-morbidity versus absence of co-morbidity

Different studies have shown the association between co-morbidity and low back pain (Hestbaek et al., 2004; Schneider et al., 2007; de Luca et al., 2017). The concurrent diseases can be musculoskeletal disorders, such as rheumatoid arthritis, osteoarthritis, and osteoporosis (Schneider et al., 2007), or diseases such as diabetes, cardiovascular or pulmonary diseases (de Luca et al., 2017) or headache and asthma (Hestbaek et al., 2004). The scale gives higher scores in case of co-morbidity.

Working in night shifts versus day shifts

Studies have shown the relationship between shift working and the prevalence of LBP (Zhao et al., 2012; June et al., 2011). In the present study, the scores were higher in nurses who worked in night shifts than those who did not.

High frequent versus low frequent patients/loads handling

More frequent lifting during a shift increases the likelihood of LBP occurrence. Even lifting the light loads with high frequency can contribute to the occurrence of LBP. For example, if a person lifts 11.3 Kg weight 25 times a day, the risk of acute LBP increases by 3 times (Yip, 2001; Yasobant and Rajkumar, 2014). The occupational sub-scale could well differentiate these two groups by giving higher scores to nurses with more frequent load/patient handling.

The present study was an attempt to propose and validate a new scale which supports the multidimensional nature of LBP. The final scale consisted of the individual, occupational, and psychosocial dimensions. Furthermore, the proposed scale is a simple, reliable and validated scale to predict LBP in nursing activities.

Conflict of interest

The authors declare that there are no conflicts of interest.
Acknowledgments
This paper was submitted as part of a thesis for the Ph.D. degree in Occupational Hygiene, School of Health and Safety, Shahriv Beheshti University of Medical Sciences, Tehran, Iran. We deeply appreciate the guidance and collaboration of Dr. Hamid Sharif Nia from The Mazandaran University of Medical Sciences, Dr. Heidar Mohammadi from The Larestan University of Medical Sciences, and Dr. Yalda Hashempour from The Mazandaran University of Medical Sciences.

Funding
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

ACGIH. Threshold limit values for chemical substances and physical agents and biological exposure indices. Cincinnati, OH: American Conference of Governmental Industrial Hygienists, 2017.


285