

# Accuracy of the Örebro Musculoskeletal Pain Questionnaire and Work Assessment Triage Tool for selecting interventions in workers with spinal conditions

Eloi Gergel<sup>a,b</sup>, Eric Parent<sup>a</sup> and Douglas P. Gross<sup>a,\*</sup>

<sup>a</sup>*Department of Physical Therapy, University of Alberta, Edmonton, AB, Canada*

<sup>b</sup>*Grenoble Alpes University, Grenoble, France*

Received 14 June 2020

Accepted 17 November 2020

## Abstract.

**BACKGROUND:** Accurate clinical decision support tools may help clinicians select appropriate interventions for patients with spinal conditions. The Örebro Musculoskeletal Pain Questionnaire (ÖMPQ) is a screening questionnaire extensively studied as a predictive tool. The Work Assessment Triage Tool (WATT) is a clinical decision support tool developed to help select interventions for injured workers.

**OBJECTIVE:** To compare the classification accuracy of the ÖMPQ and WATT to clinician recommendations for selecting interventions leading to a successful return to work in patients with spinal conditions.

**METHODS:** A secondary analysis was undertaken of data from injured workers with spinal conditions assessed between 2013 and 2016. We considered it a success if the workers did not receive wage replacement benefits 30 days after assessment. Analysis included positive likelihood ratio (LR+) as an indicator of predictive accuracy.

**RESULTS:** Within the database, 2,872 patients had complete data on the ÖMPQ, WATT, and clinician recommendations. At 30 days, the ÖMPQ was most accurate for identifying treatments that lead to successful outcomes with a LR+ = 1.51 (95% Confidence Interval 1.26–1.82) compared to 1.05 (95% Confidence Interval 1.02–1.09) for clinicians, and 0.85 (95% Confidence Interval 0.79–0.91) for the WATT.

**CONCLUSIONS:** All tool recommendations had poor accuracy, however the ÖMPQ demonstrated significantly better results.

Keywords: Spine, decision making, return to work, occupational health, rehabilitation

## 1. Introduction

Spinal conditions, and especially low back pain, have a major public health impact on the working population as they are leading causes of disability and productivity

loss [1,2]. Improved management strategies are needed, especially clinical tools that can help clinicians select effective treatments based on individual worker characteristics (i.e., clinical decision support tools) [3–5]. Accurate and validated clinical screening and decision support tools would allow clinicians to select the most appropriate interventions to facilitate timely recovery and return to work in individuals off work due to spinal conditions [6–8].

The Örebro Musculoskeletal Pain Questionnaire

---

\*Corresponding author: Douglas P. Gross, Department of Physical Therapy, University of Alberta, 2-50 Corbett Hall, Edmonton, Alberta, T6G 2G4, Canada. Tel.: +1 780 492 2690; E-mail: doug.gross@ualberta.ca.

(ÖMPQ) is a screening questionnaire that has been extensively studied as a screening and predictive tool in the work rehabilitation field [9,10]. The ÖMPQ has been shown to have a 89% sensitivity in the prediction of longer-term (more than 30 days) sick leave in patients with back pain as well as high reliability (Intraclass Correlation Coefficient = 0.98) [9,10]. However, its use as a clinical decision support tool has not been studied. Based on the ÖMPQ total score, patients are placed into one of three categories illustrating the risk of chronicity: low, moderate or high. High risk indicates the individual is more likely to experience delayed recovery and return to work, which has been demonstrated in workers with low back pain [11–13]. Theoretically, higher risk patients should warrant more intensive interventions, however, the accuracy and validity of the ÖMPQ as a clinical decision support tool for selecting treatments has not been previously studied.

The Work Assessment Triage Tool (WATT) is another clinical decision-support tool designed using state-of-the-art machine learning techniques to help select rehabilitation interventions for injured workers [14]. In the original WATT development study conducted on a sample that included a variety of work-related musculoskeletal disorders, classification accuracy of the WATT for selecting successful treatments was found to be higher than recommendations of clinicians [14]. However, in external validation studies [15,16], accuracy of the WATT was more modest and further development and testing of the WATT was recommended

Additional accuracy and validity testing of the ÖMPQ and WATT are needed in workers with spinal conditions to determine if either can classify patients into appropriate intervention groups better than recommendations of clinicians. Therefore, the objective of this study was to compare the classification accuracy of the ÖMPQ and WATT to clinician recommendations for selecting appropriate interventions leading to a successful return to work in patients with spinal conditions.

## 2. Materials and methods

### 2.1. Study design

We conducted a historical comparative cohort study on patients with spinal conditions using data from the administrative and clinical databases of the Workers' Compensation Board – Alberta (WCB-Alberta). This is

a secondary analysis of a previously collected dataset used for evaluating the WATT among workers with a variety of musculoskeletal conditions. Details of the data collection procedures and variables in the dataset are available elsewhere [15]. In brief, all clinical measures in the dataset, including the ÖMPQ and WATT, were collected during routine clinical care as part of a return-to-work assessment conducted to determine patients' ability to return to work or need for further treatment. We followed the Standards for Quality Improvement Reporting Excellence (SQUIRE 2.0) guidelines to report our results [17].

### 2.2. Population

Data were extracted from the WCB-Alberta database on patients considered for rehabilitation between January 2013 and December 2016. We used ICD9 diagnostic codes to identify patients with specific or non-specific spinal conditions. Relevant diagnostic codes that were considered as spinal conditions and the actual number of cases in our study are provided in Supplementary Table 1. We excluded patients who did not have a spinal condition or who did not have a complete ÖMPQ score.

### 2.3. Procedures

This study was approved by the University of Alberta Health Research Ethics Board. No patients were recruited/consented since this study relied on archived program evaluation data from the WCB-Alberta database. More details on the data extraction process are provided elsewhere [15]. From the existing data, we created the intermediate variables needed to perform the analysis.

### 2.4. Measures

#### 2.4.1. ÖMPQ

The ÖMPQ is a 25-item questionnaire that evaluates psychosocial and work-related barriers to recovery [9, 12]. ÖMPQ scores range between 0 and 210, but the overall score is categorized using cut-offs to represent a patient's level of risk of delayed recovery (i.e., low, medium, or high risk). Originally, the ÖMPQ cutoffs were defined as 90 and 105 to discriminate low, medium and high risk of long-term disability [9]. Cut-offs used by the WCB-Alberta are slightly higher because the jurisdiction has found their population typically scores higher on the tool. The values of the WCB-Alberta

Table 1

Clinical decision support tool classification strategies showing risk level for delayed return to work and corresponding treatment recommendation

Low risk	Moderate risk	High risk
Original ÖMPQ score cut-off: < 90	90 to 105	> 105
Alternate Alberta ÖMPQ score cut-off: < 105	105 to 130	> 130
Treatment recommended:		
No further rehabilitation, community single service provider (physical therapy, chiropractic, other)	Functional restoration, workplace-based intervention, hybrid of these two programs	Complex interdisciplinary biopsychosocial rehabilitation

cutoffs are 105 and 130 to discriminate low, medium and high risks. We conducted our analysis with both the original and the WCB-Alberta cutoffs.

For use as a clinical decision support tool, the different ÖMPQ risk categories lead to different treatment recommendations. Within the WCB-Alberta system [18], various rehabilitation interventions were available including: 1) single service rehabilitation provider (i.e., physiotherapy, chiropractic, other provider); 2) functional restoration; 3) workplace-based intervention; 4) hybrid functional restoration/workplace-based intervention; 5) complex interdisciplinary biopsychosocial rehabilitation; or 6) no further rehabilitation. These programs have been described in greater detail elsewhere [14,18]. According to the WCB-Alberta, the complex interdisciplinary biopsychosocial program is designed for patients with “significant barriers to return to work” [19]. We considered this complex interdisciplinary program as the only intervention adequate for patients in the ÖMPQ ‘high risk’ category. Functional restoration, workplace-based intervention, hybrid functional restoration/workplace-based intervention have been described as interventions used “to address the medical, functional, musculoskeletal, psychosocial, and vocational needs of the worker” [20]. We considered these interventions adequate for patients in the ÖMPQ moderate risk category. We considered the “single service interventions” and “no treatment” options as adequate for patients in the ÖMPQ low risk category. The two ÖMPQ risk level cut-offs used and corresponding treatment recommendations are shown in Table 1.

#### 2.4.2. WATT

WATT is an 18-item computerized clinical decision-support tool that provides rehabilitation recommendations based on an algorithm developed using machine learning [14]. The tool provides recommendations using individual worker-level characteristics for different rehabilitation program possibilities available within the WCB-Alberta jurisdiction (listed above). Validity of the WATT has been studied in injured workers with a variety of musculoskeletal conditions.

In the original development study, the WATT outperformed recommendations made by clinicians (Receiver Operating Characteristic accuracy of 0.94 for the WATT versus 0.86 for clinicians) [14]. However, external validity demonstrated more modest accuracy, and further development and testing of the WATT is needed especially within specific patient groups such as those with spinal conditions where it has not been evaluated [15,16].

#### 2.4.3. Clinician recommendations

The dataset also contained information on the treatment recommendations made by clinicians conducting the assessments of workers with spinal conditions. These clinicians were trained physical therapists, occupational therapists or exercise therapists using a standardized battery of clinical tools that was consistent across workers as it was conducted in the context of a formal functional capacity evaluation. Therapists could recommend any rehabilitation program available within the jurisdiction (see Table 1) based on their clinical assessment findings. The clinicians did not have access to recommendations of the WATT, but they completed and scored the ÖMPQ as part of their standardized battery of assessments. However, they were not trained on using the ÖMPQ to inform treatment decisions but used it for prediction purposes only. Additionally, they made decisions using information from the entire clinical encounter and battery of assessments performed, of which the ÖMPQ was only one tool.

#### 2.4.4. ‘Matching’ strategy

For each of the decision making strategies (WATT, ÖMPQ, clinician recommendations) we considered that the recommendation and the intervention ‘matched’ if a low risk intervention (single service interventions and no treatment) was recommended and actually undertaken, if a moderate risk intervention (functional restoration program, workplace-based intervention, hybrid functional restoration/workplace-based intervention) was recommended and actually undertaken, or if a high risk intervention (complex, psychological) was recommended and actually undertaken. We considered

that the recommendation and intervention undertaken were ‘unmatched’ if the recommendation and intervention undertaken were not in the same risk level (i.e., high risk intervention was recommended but low risk intervention undertaken).

## 2.5. Outcome

Our reference criterion was based on whether the intervention actually undertaken led to a successful complete or partial return to work. Return to work was measured using a surrogate indicator, reception of wage replacement benefits, and was available on 100% of our sample. Workers who were not receiving a full day of wage replacement benefits 30 days following assessment were considered to have returned to work. This was our main indicator of treatment success since wage replacement benefits are a commonly used measure within workers’ compensation jurisdictions and was used to evaluate the WATT tool previously [14,21]. We also considered other potential wage replacement outcomes including whether the workers received wage replacement benefits (full or partial benefits) for up to 180 days after assessment in the follow-up year.

## 2.6. Analysis

The sample was described using appropriate descriptive statistics for nominal, ordinal, and continuous variables. We used available demographic data to describe the patients with spinal conditions and a complete  MPQ score.

To assess the matching of risk-based recommended treatments to those actually undertaken versus whether that treatment was successful, we used the contingency table shown in Table 2. Columns reported whether the patients had successfully returned to work or not within 30 days and the rows reported the frequencies of whether the actual treatment received matched the recommended treatment or not. Sensitivity was calculated as the number of cases where the recommendation and the actual intervention matched and led to a successful return to work divided by the total number of cases with successful return to work. Specificity was calculated as the number of cases where the recommendation and the intervention mismatched and did not lead to a successful return to work divided by the total number of cases without successful return to work. Sensitivity and specificity were compared between decision tools and clinician recommendations using 95% confidence intervals (95%CI). The overall accuracy for each recommenda-

Table 2  
Contingency table used to calculate accuracy parameters of the clinical decision support tool recommendations

	Successful intervention	Unsuccessful intervention
Matched	True positive (TP)	False positive (FP)
Not matched	False negative (FN)	True negative (TN)

Sensitivity =  $TP/(TP + FN)$ , Specificity =  $TN/(FP + TN)$ , Overall accuracy =  $(TP + TN)/(TP + TN + FP + FN)$ , Positive likelihood ratio =  $sensitivity/(1-specificity)$ , Negative likelihood ratio =  $(1-sensitivity)/specificity$ .

tion was calculated as the number of patients correctly classified divided by the total number of patients. Receiver Operating Characteristic (ROC) values were also calculated. These parameters were calculated overall for the  MPQ, WATT, and clinician recommendations, as well as separately within each risk level.

Our main comparison between recommendation sources was based on the positive likelihood ratio (LR+). LR+ was considered useful for identifying how the decision-making tools improved the initial probability of return to work (i.e., prevalence). It is interpreted as how importantly does following a recommendation change the initial probability (prevalence) of return to work. The size of the LR+ is interpreted as follows: 1 no difference, 2 small, 5 moderate, 10 large [22]. In addition, this parameter takes into account both sensitivity and specificity. It is calculated by the formula  $LR+ = Sensitivity/(1-Specificity)$ . We are interested in the LR+ because we want to know how much we increase the probability of return to work by using the appropriate intervention (i.e., by following a tool’s recommendation). The LR– is a secondary outcome because we are less interested in knowing the probability of return to work when we don’t follow a tool’s recommendation. The 95% CI was used to determine if the different tools had significantly different results (no superposition/overlapping).

In addition to completing analyses on the entire sample, we conducted sub-group analyses within relevant occupational and diagnostic groups. We examined results within the occupational group of ‘Trades, Transport and Equipment Operators and Related Occupations’ since this group had adequate sample size and has traditionally been at high risk of experiencing spinal conditions. We also examined results separately for workers with neck, back, non-specific, and specific diagnoses (see Supplementary Table 1 for relevant ICD9 codes representing these groups). Analyses were performed using IBM<sup> </sup> SPSS<sup> </sup> Statistics version 26 (Armonk, NY, USA).

Table 3  
 Characteristics of patients with spinal conditions and complete data on the Örebro Musculoskeletal Pain Questionnaire ( $n = 2872$ )

	<i>n</i> (%)
Worker has a job to return to (yes)	2592 (90.3%)
Occupational category	
Management occupations	89 (3.1%)
Business, finance and administration occupations	159 (5.5%)
Natural and applied sciences and related occupations	69 (2.4%)
Health occupations	366 (12.7%)
Occupations in education, law and social, community and government	143 (5.0%)
Occupations in art, culture, recreation and sport	22 (0.8%)
Sales and service occupations	627 (21.8%)
Trades, transport and equipment operators and related occupations	1179 (41.1%)
Natural resources, agriculture and related production occupations	67 (2.3%)
Occupations in manufacturing and utilities	150 (5.2%)
Occupation unknown	1 (0.0%)
Currently working in some capacity	1424 (49.6%)
Educational level	
Grade 8 or less	32 (1.1%)
Partial high school	191 (6.7%)
High school diploma	354 (12.3%)
Partial technical school	92 (3.2%)
Technical school	267 (9.3%)
Partial university	69 (2.4%)
University degree	132 (4.6%)
Education not specified	1735 (60.4%)
Comorbid injury	466 (16.2%)
Rural (residence)	789 (27.5%)
Self-employed	23 (0.8%)
Concurrent employment	17 (0.6%)
Sex (male)	1749 (60.9%)
Did the worker receive any benefits 30 days after assessment (% yes)	455 (15.8%)
Did the worker receive any benefits more than 180 days in the follow-up year	97 (3.4%)
Spinal area affected	
Back	1463 (50.9%)
Neck	1407 (49.0%)
Spine (region unspecified)	2 (0.1%)
Specific spinal pathology (fracture, dislocation, spondylolisthesis, stenosis, etc.)	498 (17.3%)
Nonspecific pathology (back pain, neck pain, sacroiliac pain)	2374 (82.7%)
	Mean (SD)
Duration of injury – number of calendar days from injury to assessment	144.9 (696.7)
Age (years)	41.5 (12.5)
Number of prior compensation claims	3.6 (4.7)
Pain Disability Index out of 100 (203 had missing data)	52.3 (21.6)
Visual Analogue Pain Scale out of 10 (377 had missing data)	5.4 (2.5)

### 3. Results

#### 3.1. Patient characteristics

In the full database, 8,747 injured workers had a spinal condition. All of these had data on the clinician and WATT recommendations. However, only 2,872 (33%) had complete ÖMPQ data. In 32.8% of the cases, patients had low back pain (i.e. ICD9 diagnosis of lumbago, backache, or lumbar sprain). In 82.7% of cases, patients had a non-specific diagnosis. For those with spinal conditions and complete ÖMPQ data, descriptive characteristics are presented in Table 3.

#### 3.2. Treatment recommendations

The tools' and clinicians' recommendations cross-tabulated against the level of risk assigned to the interventions actually undertaken by each worker are shown in Table 4. Clinicians rarely recommended interventions considered suitable for 'high risk' workers (1.4% of recommendations made), whereas the WATT rarely recommended interventions considered suitable for 'low risk' workers (0.2% of recommendations made). The ÖMPQ was more balanced across the three risk levels, with some differences seen between the original and WCB-Alberta classifications.

Table 4  
Cross-tabulation of the risk levels identified by each source compared to levels of risk for programs actually undertaken

Source of the recommendation	Recommendation categories	Level of risk corresponding to the rehabilitation program actually undertaken			Total
		Low risk	Moderate risk	High risk	
WATT	Low risk	6	0	1	7
	Moderate risk	1147	1133	11	2291
	High risk	273	284	17	574
Clinician recommendation	Low risk	861	152	6	1019
	Moderate risk	546	1263	5	1814
	High risk	19	2	18	39
�MPQ (original cut-offs)	Low risk	487	420	1	908
	Moderate risk	328	373	2	703
	High risk	611	624	26	1261
�MPQ (WCB-Alberta cut-offs)	Low risk	793	771	3	1567
	Moderate risk	474	511	13	998
	High risk	159	135	13	307

Table 5  
Diagnostic accuracy statistics for whether matched recommendations/programs led to successful return to work 30 days after rehabilitation

Tool	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio	Overall accuracy	Area under the ROC
Clinician recommendation	0.75 (0.73–0.76)	0.26 (0.22–0.30)	1.01 (0.95–1.07)	0.97 (0.82–1.14)	0.67	0.50
WATT recommendation	0.40 (0.38–0.42)	0.57 (0.53–0.62)	0.93 (0.83–1.05)	1.05 (0.96–1.14)	0.43	0.49
�MPQ original cut-offs	0.33 (0.31–0.34)	0.78 (0.75–0.82)	1.51 (1.26–1.82)	0.86 (0.81–0.91)	0.40	0.56
�MPQ WCB-Alberta cut-offs	0.47 (0.45–0.49)	0.59 (0.54–0.63)	1.13 (1.01–1.27)	0.91 (0.83–0.99)	0.49	0.53

WATT = Work Assessment Triage Tool,  MPQ =  rebro Musculoskeletal Pain Questionnaire, ROC = Receiver Operating Characteristic. For each parameter the numbers represent: Value (95% confidence interval).

### 3.3. Tools' accuracy

At 30 days after the assessment, all sensitivities and specificities were modest for ability to recommend interventions that lead to a successful outcome (see Table 5). The clinicians' recommendations were the most sensitive (0.75, 95% CI 0.73–0.76) while the  MPQ with original cut-offs was the most specific (0.78, 95% CI 0.75–0.82). The LR+ (primary outcome) were small for each tool, but the LR+ of the  MPQ with the original cut-offs (1.51, 95% CI 1.26–1.82) was significantly higher than the WATT (0.93, 95% CI 0.83–1.05), the recommendations of clinicians (1.01, 95% CI 0.95–1.07) or the  MPQ with the alternate WCB-Alberta cut-offs (1.13, 95% CI 1.01–1.27) (Table 5). Matching interventions with low risk categories recommendations were much more accurate in predicting successful outcomes than the matching in the moderate and high risk categories. For example, for the  MPQ with the original cut-offs the LR+ were 3.80 (95% CI 2.53–5.7), 0.96 (95% CI 0.75–1.24), 0.16 (95% CI 0.08–0.35), respectively, for matching interventions to the low, moderate and high risk recommendation categories.

Results did not change substantially on any of the subgroup analyses with the exception of slightly improved sensitivity of clinician recommendations within

the back pain group (sensitivity increased from 0.75 to 0.83). However, all Likelihood Ratios as well as the  MPQ and WATT results did not change meaningfully (less than 10% change).

When we defined return to work success as receiving less than 180 days of compensation in the follow up year, results were not substantially different (results not shown) however there were no longer statistically significant differences between the tools. Overall, in most of the analyses the clinicians were the most sensitive (when return to work is a success clinical recommendations often matched actual treatments undertaken) and the  MPQ with the original cut-offs was the most specific (when recommended treatments don't match those undertaken, the  MPQ had the highest number of failures).

## 4. Discussion

### 4.1. Main findings

Previous research indicates that screening questionnaires such as the  MPQ may help for predicting delayed return to work [9,12,23]. In this paper, we evaluated clinical decisions made based on the  MPQ risk

recommendations against a computer-based decision support tool (WATT) and clinician treatment recommendations made following a thorough assessment. Matching risk recommendations to the actual treatments undertaken resulted in overall poor accuracy, however the  MPQ had significantly better results than the WATT and clinician recommendations. Therefore,  MPQ may also be helpful for selecting appropriate interventions. In fact, based on our observed likelihood ratios, following the  MPQ's recommendations could increase the return to work probability and not following it could decrease the return to work probability. The original cut-off showed the best accuracy.

To our knowledge this is one of the first times that the  MPQ has been tested as a clinical decision support tool. Clinicians knew the  MPQ results and may have used it in their decisions, however the  MPQ alone still resulted in higher accuracy than clinician recommendations. Further research is needed to explore optimal cut-offs and what additional information or measures could improve the accuracy of treatment recommendations.

Overall, the difference between the different recommendation sources is small and the LR+ were often low. However matching treatments provided by the  MPQ recommendations with the original cutoffs offered significantly better prediction of treatment success than the clinician recommendations alone (current standard practice). Even if the differences between tools are low, the consequences of using the tool are important. For example, if we consider a return to work prevalence of 68.2% [24] at 30 days, following the clinicians' recommendation would increase the rate of return to work by only 1% (95% CI = 0.5–1.8%), whereas following the  MPQ with the original cutoffs would improve the likelihood of return to work by 8% (95% CI = 4.7–11.4%). When applied to an entire population of workers, this may result in sizeable cost savings through earlier return to work.

#### 4.2. Strengths and limitations

Relying on historical archived data from WCB-Alberta allowed us to achieve a large sample size, however it also had some limitations and resulted in a large amount of missing data. We had no control over what variables could be measured and were limited to the variables in the existing dataset. In addition, the WATT and the clinicians normally have 6 treatment categories but in order to make it comparable with the  MPQ risk levels, we merged treatment options into only 3 categories. This increases the accuracy parameters of

matching the recommendations but decreases precision (it recommends a group of interventions, not a single one). Additionally, the  MPQ treatment classifications used have not been previously studied and different classification strategies may perform better. The generalization of results is limited due to the restricted location of the population. The analysis has been restricted only to the 33% of the patients with spinal pain who had a complete  MPQ, which could lead to a selection bias. On the other hand, the study has a relatively high number of participants that could improve external validity of results. Another limitation is the use of wage replacement benefits as our main outcome measure. Reception of benefits is a surrogate indicator of actual return to work and is influenced by several individual and contextual factors. However, this outcome is commonly used in studies of injured workers and is an important outcome within workers' compensation jurisdictions. Instead of wage replacement outcomes, future prospective studies in this area could include patient-reported outcome measures that were not available in our dataset.

## 5. Conclusion

To conclude, it is too early to recommend the use of the  MPQ by clinicians as a decision support tool. However, the  MPQ had significantly better accuracy than the clinician recommendations or the WATT. The  MPQ should be tested in a randomized controlled trial comparing outcomes between current practice and when making decisions informed by the  MPQ.

## Acknowledgments

The Workers' Compensation Board of Alberta provided funding and assisted with data collection for the original study from which data came.

## Conflict of interest

The authors have no conflict of interest to report.

## Supplementary data

The supplementary files are available to download from <http://dx.doi.org/10.3233/BMR-200169>.

## References

- [1] Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. *Lancet*. 2018; 391(10137): 2356–67. doi: 10.1016/S0140-6736(18)30480-X.
- [2] Global Burden of Disease Injury, Incidence, Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. 2016; 388(10053): 1545–602. doi: 10.1016/S0140-6736(16)31678-6.
- [3] Gross DP, Armijo-Olivo S, Shaw WS, Williams-Whitt K, Shaw NT, Hartvigsen J, et al. Clinical decision support tools for selecting interventions for patients with disabling musculoskeletal disorders: a scoping review. *J Occup Rehabil*. 2016; 26(3): 286–318. doi: 10.1007/s10926-015-9614-1.
- [4] Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008; 59(5): 632–41. doi: 10.1002/art.23563.
- [5] Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*. 2011; 378(9802): 1560–71. doi: 10.1016/S0140-6736(11)60937-9.
- [6] Nicholas MK, Linton SJ, Watson PJ, Main CJ, Decade of the Flags, Working G. Early identification and management of psychological risk factors (“yellow flags”) in patients with low back pain: a reappraisal. *Phys Ther*. 2011; 91(5): 737–53. doi: 10.2522/ptj.20100224.
- [7] Shaw WS, Linton SJ, Pransky G. Reducing sickness absence from work due to low back pain: how well do intervention strategies match modifiable risk factors? *J Occup Rehabil*. 2006; 16(4): 591–605. doi: 10.1007/s10926-006-9061-0.
- [8] Six Dijkstra M, Siebrand E, Dorrestijn S, Salomons EL, Reneman MF, Oosterveld FGJ, et al. Ethical considerations of using machine learning for decision support in occupational health: an example involving periodic workers’ health assessments. *J Occup Rehabil*. 2020; 30(3): 343–53. doi: 10.1007/s10926-020-09895-x.
- [9] Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Örebro Musculoskeletal Pain Questionnaire. *Clin J Pain*. 2003; 19(2): 80–6.
- [10] Gabel CP, Melloh M, Yelland M, Burkett B, Roiko A. Predictive ability of a modified Örebro Musculoskeletal Pain Questionnaire in an acute/subacute low back pain working population. *Eur Spine J*. 2011; 20(3): 449–57. doi: 10.1007/s00586-010-1509-4.
- [11] Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *Clin J Pain*. 1998; 14(3): 209–15.
- [12] Dagfinrud H, Storheim K, Magnussen LH, Odegaard T, Hof-taniska I, Larsen LG, et al. The predictive validity of the Örebro Musculoskeletal Pain Questionnaire and the clinicians’ prognostic assessment following manual therapy treatment of patients with LBP and neck pain. *Manual Therapy*. 2013; 18(2): 124–9. doi: 10.1016/j.math.2012.08.002.
- [13] Hill JC, Dunn KM, Main CJ, Hay EM. Subgrouping low back pain: a comparison of the STarT Back Tool with the Örebro Musculoskeletal Pain Screening Questionnaire. *Eur J Pain*. 2010; 14(1): 83–9. doi: S1090-3801(09)00006-8 [pii] 10.1016/j.ejpain.2009.01.003.
- [14] Gross DP, Zhang J, Steenstra I, Barnsley S, Haws C, Amell T, et al. Development of a computer-based clinical decision support tool for selecting appropriate rehabilitation interventions for injured workers. *J Occup Rehabil*. 2013; 23(4): 597–609. doi: 10.1007/s10926-013-9430-4.
- [15] Gross DP, Steenstra IA, Shaw W, Yousefi P, Bellinger C, Zai-ane O. Validity of the work assessment triage tool for selecting rehabilitation interventions for workers’ compensation claimants with musculoskeletal conditions. *J Occup Rehabil*. 2020; 30(3): 318–30. doi: 10.1007/s10926-019-09843-4.
- [16] Qin Z, Armijo-Olivo S, Woodhouse LJ, Gross DP. An investigation of the validity of the work assessment triage tool clinical decision support tool for selecting optimal rehabilitation interventions for workers with musculoskeletal injuries. *Clin Rehabil*. 2016; 30(3): 277–87. doi: 10.1177/0269215515578696.
- [17] Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for Quality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. *BMJ Qual Saf*. 2016; 25(12): 986–92. doi: 10.1136/bmjqs-2015-004411.
- [18] Stephens B, Gross DP. The influence of a continuum of care model on the rehabilitation of compensation claimants with soft tissue disorders. *Spine (Phila Pa 1976)*. 2007; 32(25): 2898–904. doi: 10.1097/BRS.0b013e31815b64b6.
- [19] Return to Work Services Schedule 4B: Complex Programs. Workers’ Compensation Board of Alberta. 2020. [https://www.wcb.ab.ca/assets/pdfs/providers/RTW/Schedule\\_4B\\_Complex\\_Programs.pdf](https://www.wcb.ab.ca/assets/pdfs/providers/RTW/Schedule_4B_Complex_Programs.pdf). Accessed October 14 2020.
- [20] Return to Work Services Schedule 3: Provider-Site Based, Worksite-Based & Hybrid Programs. 2020. [https://www.wcb.ab.ca/assets/pdfs/providers/RTW/Schedule\\_3\\_PSB\\_WSB\\_Hybrid\\_Programs.pdf](https://www.wcb.ab.ca/assets/pdfs/providers/RTW/Schedule_3_PSB_WSB_Hybrid_Programs.pdf). Accessed October 14 2020.
- [21] Fan JK, Macpherson RA, Smith PM, Harris MA, Gignac MAM, McLeod CB. Age Differences in Work-Disability Duration Across Canada: Examining Variations by Follow-Up Time and Context. *J Occup Rehabil*. 2020. doi: 10.1007/s10926-020-09922-x.
- [22] Raslich MA, Markert RJ, Stutes SA. Selecting and interpreting diagnostic tests. *Biochem Medica*. 2007; 17(2): 151–61. doi: 10.11613/BM.2007.014.
- [23] Sattelmayer M, Lorenz T, Roder C, Hilfiker R. Predictive value of the Acute Low Back Pain Screening Questionnaire and the Örebro Musculoskeletal Pain Screening Questionnaire for persisting problems. *Eur Spine J*. 2011. doi: 10.1007/s00586-011-1910-7.
- [24] Wynne-Jones G, Cowen J, Jordan JL, Uthman O, Main CJ, Glozier N, et al. Absence from work and return to work in people with back pain: a systematic review and meta-analysis. *Occup Environ Med*. 2014 Jun; 71(6): 448–56. doi: 10.1136/oemed-2013-101571.