Expanded Review of the Literature: Low Back Pain Prognostic Factors, Psychometric Properites of the Oswestry Disability Index, Multivariate Logistic Regression Analysis, and Low Back Pain Clinical Prediction Rules

Chris Green

This expanded review of the literature serves to inform our group's analysis of data from a single outpatient physical therapy clinic. Our dataset included prognostic factors of interest, which have been studied elsewhere in prognostic low back pain studies. A variety of prognostic factors have been reviewed below, including demographic information, baseline physical factors of the patients low back pain, and other possible influences like total visits or prescribed drug use. The demographic factors considered are age and gender. Physical factors include disability at baseline as measured by the Oswestry Disability Index (ODI), chronicity level or duration of symptoms, and the presence of widespread or radicular pain. A brief discussion of other common prognostic factors that were not available in this dataset yet occur frequently in the literature is included.

A review of the psychometric properties of the ODI, one of the primary outcome measures used in the data analysis, has been included. A brief explanation of the multinomial logistic regression modeling used for our data analysis follows the psychometrics section. Trichotimization of the ODI into 3 groups of minimally important clinical difference (MCID) is discussed in terms of the appropriateness of this approach to the data analysis. Finally, an initial aim of our research was to examine the possibility of developing a clinical prediction rule (CPR) based on the dataset. Therefore, a review of prognostic CPRs for low back pain concludes this review.

Prognostic Factors of Low Back Pain

Demographic

Age

Age has widely been used as an independent factor when it is considered as a possible prognostic factor in low back pain (LBP) studies.[1] A younger age has been shown to be a prognostic of a better outcome for LBP patients in many studies.[2,3,4] An older age has also been suggested as a negative prognostic factor for LBP patients.[5,6] One major common shortcoming of the existing literature is the lack of a standard definition of old and young age. Younger age has been defined as anywhere between 32 to 45 years, [5,7,8] and older age has been defined between 33 and 53 years.[5,6,7] There are also a number of studies that did not define older or younger age at all.[2,3] Campbell et al[7] used four age groups in their LBP prognostic study: ≤ 37, 38-45, 46-52, and ≥ 53 years old. In their statistical analysis, however, age did not predict outcome at either 6 months or 5 years.[7] The omission of age categorizations and overlapping of younger versus older age make this prognostic factor difficult to define. Moving forward, prognostic literature could benefit from defining a standardized and narrower age range when delineating old versus young age.

Gender

The gender of patients has been examined in LBP prognostic studies with the aim to determine if males or females have better outcomes. This factor can be easily dichotomized and is inherent in every study population. While Bekkering et al[2] did not find gender to be a significant prognostic factor, others have found the female gender predicted poor outcomes. [6,7,9] Based on the few studies that use gender as a prognostic factor, the female gender may be a stronger prognostic factor for poorer outcomes.

Physical Factors

Initial Oswestry Disability (ODI) score

The use of initial Oswestry Disability Index (ODI) score may be an important prognostic factor because it uses a baseline score on a measure validated for LBP patients. Cook et al[3] included initial ODI score as a prognostic factor of interest to examine generic predictors of outcome in LBP patients. They found that lower baseline ODI scores were individual prognostic variables within 2 of 4 of their statistical models.[3] Schwind et al[4] identified initial ODI score as a prognostic factor when using an MCID of 5 or 10 points on the ODI. A high baseline ODI has also been shown to predict a longer duration of days missed from work.[9]

Chronicity level

Chronicity level is used to describe the duration of LBP symptoms a patient has experienced. Duration of symptoms has been described in a variety of ways: time since the patient's last pain free month;[8] three different chronicity lengths (0–3 weeks/4–12 weeks/>12 weeks);[2] four different chronicity lengths (< 2 weeks, 2-3 weeks, 3-4 weeks, and 4-6 weeks);[10] and a long duration of chronicity as > 3 years since the onset of the current LBP episode.[8] Many have found a shorter duration of symptoms to be prognostic of better outcomes.[2,3,11,12] The majority of the studies reviewed agree that a low chronicity level may be a positive predictor of recovery from LBP.

Widespread pain

Widespread pain has been examined as a possible physical prognostic factor for LBP outcomes. The presence of widespread pain could be an indicator of centralization, which has been found to be a predictor of disability after 6 months of treatment.[13] Widespread pain has been defined as the presence of strong leg pain, distal leg pain, or upper body pain[8] or a drawing of

multiple areas on a pain diagram.[5] Many studies have demonstrated that widespread pain was a strong prognostic indicator for disability, both at 6 months[13] and 12 months.[8,14] One study by Abbott and Kingan[5] did not find widespread pain to be a significant prognostic factor in their analysis. The majority of the literature reviewed, however, shows strong correlation that widespread pain is a negative predictive factor in patients with LBP.

Radicular pain

Radicular pain or pain into the legs has been used in numerous prognostic studies for LBP. While George et al[13] found that men with leg pain at baseline was a prognostic factor for disability, a review of 9 LBP studies that measured pain as prognostic indicator identified radicular pain below the leg as a significant factor for long term disability.[15] Other studies simply mention that radicular pain was included as part of prognostic models or analyses.[10,1] The presence of radicular symptoms with LBP appears to be a significant negative indicator for poor outcomes in the long term.

Prescribed drugs

Traeger et al[10] used prescribed drug use as a prognostic factor in their study, delineating it as a yes/no question. They did not define a class or type of medication, but only asked if the patients were taking medication for their low back pain.

Total Visits

Cook et al[3] used total visits as an outcome measure in their study, placing the standard at 6 visits as they felt this number reflected thorough and efficient care in terms of cost.[3]

Additional Prognostic Factors

Other prognostic factors that were not available in this project's data have been used in LBP prognostic studies, including work status, eligibility for the clinical prediction rule (CPR) for spinal manipulation, and psychosocial factors. Employment factors such as absence from work and being unemployed have been shown to predict negative outcomes at 12 months.[8,14] It has also been demonstrated that a low initial ODI score can be predictive of worsening work status in the LBP population.[16] The influence of psychosocial factors as prognostic factors in LBP is a controversial subject. It has been suggested that psychosocial factors such as fear avoidance and depression have predicted negative outcomes in both the short and long term,[17] which runs counter to others who have found no influence of those factors on LBP.[7] Differing study populations or design may contribute to the differences in study findings. Perhaps subjects who are anxious about their symptoms and actively seek treatment may already have more severe chronicity or disability secondary to their LBP.

Others have suggested that negative beliefs regarding their own LBP, including fear of extended length of LBP symptoms or disability as a result of LBP, are negative predictors of disability in the short term of 6[18] to 12 months[19,20,21] and the long term of 5 years.[7] Finally, eligibility for the manipulation CPR includes meeting 4 of the 5 following criteria: pain > 16 days, no symptoms distal to the knee, FABQ score > 19, internal rotation of at least one hip > 35º, and hypomobility of at least one level of the lumbar spine. The CPR for spinal manipulation was found to be the most robust prognostic factor across 4 different predictive models in a study by Cook et al.[3] In this light, the manipulation CPR's role could serve as a valuable prognostic factor LBP patients.

Discussion of Psychometric Properties of the Oswestry Disability Index (ODI)

Validity

Validity is the ability of a tool to measure what it intends to measure.[22] The face and content validity of the Oswestry Disability Index (ODI) has been addressed in the literature. Patients experiencing low back pain (LBP) for the first time who were expected to improve over time was concurrently correlated with improvements in ODI scores.[23] Another analysis of 81 patients who improved over a 5-week period was shown to correlate with an expected improvement in disability as demonstrated by improved ODI scores.[24] A validation of the ODI that included relating patient behavior to the responses of the questionnaire demonstrated strong correlations with the sitting (tau = 0.41) and walking (tau = 0.46) sections of the questionnaire, and to a lesser extent, the lifting category (tau = 0.32).[25]

The construct validity of determining whether the wording used in the ODI reflects different aspects of disability related to LBP has support in the literature. The ODI has been correlated with a number of other outcome measures in terms of low back pain (LBP) that contributes to disability. The ODI has been moderately correlated with the Visual Analogue Scale (VAS) of pain ($r = 0.62$)[26] and ($\varrho = 0.69$).[27] A comparison of the intercorrelation between subject disability using the Pain Disability Index (PDI) and the ODI demonstrated a high correlation between the two outcome measures (ICC = 0.83).[26] The ODI was also moderately ($p < 0.05$) to strongly correlated ($p < 0.01$) with the PDI and the ability to perform three physical tests (repetitive sit ups ($r = 0.30$); repetitive arch ups ($r = 0.35$); and repetitive squat test ($r = 0.41$) used to determine overall disability of patients in daily activities in patients with LBP.[28] The ODI shows moderate correlation with another pain measure, the McGill Pain

Questionnaire.[29,30] Moderate correlations have been shown between the ODI and the 36-item Short Form Survey (SF-36) in demonstrating low back disability.[31]

Comparisons between the ODI and physical performance measures have supported the construct validity of the ODI. The ODI was a stronger predictor for return to work in LBP patients than an opto-electric device that measured trunk kinematics at 12 weeks (ODI, $r = 0.82$; device, $r = 0.70$), 24 weeks (ODI, $r = 0.81$; device, $r = 0.65$), and 52 weeks (ODI, $r = 0.85$; device, $r = 0.30$) .[32] The ODI has also successfully predicted isometric muscular endurance.[33] Improvements with ODI scores have been associated with patients who 'centralize' in the McKenzie evaluation system, suggesting that improved back symptoms are consonant with improved ODI scores.[34]

Reliability

The ODI has been examined by various authors to determine its test-retest reliability. One study of LBP in nurses separated the test-retest reliability into different subsections and found the following:[35]

- excellent test-retest reliability ($n = 33$) for ODI subsections of **walking** (ICC = 0.78, 95% CI [0.60, 0.88]), **sleep** (ICC = 0.82, 95% CI [0.67, 0.91]), and **total ODI score** (ICC = 0.88, 95% CI [0.77, 0.94]).
- adequate test-retest reliability ($n = 33$) for ODI subsections of **pain intensity** (ICC = 0.65, 95% CI [0.40, 0.81]), **lifting** (ICC = 0.74, 95% CI [0.53, 0.87]), **sitting** (ICC = 0.71, 95% CI [0.48, 0.85]), **standing** (ICC = 0.59, 95% CI [0.31, 0.78]), **social life** (ICC = 0.52, 95% CI [0.22, 0.73]), and **travel** (ICC = 0.51, 95% CI [0.21, 0.73]).

- poor test retest reliability (n = 33) for ODI subsection of **sex** (ICC = 0.25, 95% CI [-0.11, 0.55]).

In terms of time period, the ODI has shown excellent test-retest reliability in patients with at least 2 months of back pain when the test-retest interval is 1 week (ICC = 0.83).[26] In a comparison of 5 different LBP disability questionnaires, the ODI was shown to be one of the most reliable in demonstrating an ability to detect improvement or worsening.[36] The ODI has been shown to have a test-retest reliability over 24 hours (r = 0.99)[23] and over 1-14 days (ICC = 0.94).[37]

Responsiveness

The ODI has been deemed a useful functional status questionnaire for measuring change in patients with LBP.[38] While some authors have had difficulty determining a minimal clinically important difference (MCID) for the ODI,[38,39] others have proposed a variety of values to describe an MCID. Some have opted to describe an MCID of in terms of a percentage, namely a 50% improvement.[3,40] Others have described the MCID in terms of an improvement in the raw score. These values are mainly in the range of 4 to 11 points,[24,41,42,43,4,44,45,46,47] with the most consistent value set at 10 points.[44,45,48,47] The lower range of 4-6 points has been suggested to be a useful clinical cut-off score to determine if patients' disability had improved or not after a 6 week period.[24] This lower MCID threshold may be appropriate for patients with acute back pain because of its ability to establish a meaningful change after 6 weeks. Hagg et al[45] found that an MCID of 10 was the lowest number they could identify within a 95% confidence interval. Ostelo

et al[46] proposed that acute sufferers of LBP may have higher ODI scores than those suffering from chronic LBP and suggested that 10 was an acceptable MCID value on the ODI, based on previous research. This recommendation was bolstered by the fact that Lauridsen et al[49] found an average of 11 for their MCID across a stratification of patients with differing baseline ODI values and symptoms. Because a range of values from 4 to 11 have been reported in the literature, it would stand to reason that a range of MCIDs for the ODI would be appropriate for use in the logistic regression analysis used for this project. The trichotimization of the MCID ranges for stastical analysis is discussed in the "Discussion of Logistic Regression Model in Prognostic Studies" section below.

Discussion of Logistic Regression Model in Prognostic Studies

Multivariable analysis is a statistical tool used for understanding the contributions of different factors to a single event.[50] Logistic regression modeling is a form of multivariate analysis and is used with dichotomous discrete categorical outcomes. Multinomial regression is a form of categorical regression modeling in which trichotomous discrete outcomes are possible and one of the outcomes used functions as the referent variable.[50] In our research, the outcomes are trichotomized by patients reaching a predetermined MCID of >10, 4-10, or < 4 points improvement on the ODI representing a high, medium, or low change on the ODI. The lowest MCID "did not meet" was the referent variable in the multinomial regression analysis.

The multivariable analysis helps to determine which prognostic factors lead to improvements in disability, as measured by the ODI. Multivariate analysis allows many factors to be examined independently, adjusting for potential confounding variables.[50] This is performed by estimating the conditional probability (holding all other variables constant), and the odds-ratio

describes what happens to the probability of the outcome if a variable is increased by 1 unit. Predictive models typically require more than "goodness-of-fit" statistics to account for how well a predictive model explains the outcome. Often, an area under the receiver-operating curve (ROC) displays the assessment of the predictive value of a logistic regression model over various cut points of the probability of the outcome.[51]

The validity of a logistic regression model is affected by the sample size. For simple univariate multinomial or logistic regression, Hosmer and Lemeshow[52] have recommended a minimum observation-to-variable ratio of 10:1, but cautioned that a number this low will likely overfit a model. Their preferred observation-to-variable ratio is 20:1 for the multivariate modeling.[52] These sample size requirements for logistic regression are stated as outcomes per variable rather than person per variable.[50]

An advantage of a logistic regression model includes the fact that few assumptions are made, including the distribution of the outcome. Furthermore, the results can be interpreted through the regression coefficients expressed as odds-ratios.[53] The selection of prognostic variables for use in a logistic regression should include quantitative and qualitative information from published studies and clinical knowledge,[53] as has been discussed earlier in the review of the prognostic literature.

A hierarchical multinomial model can be used when the outcome of a response variable can have multiple outcomes, or "polytomous [(≥ 3)] responses."[54] These responses are often qualitative such as describing the severity of a disease. Polytomous responses may have classifications that are not independent of each other, rather the response occurs in a consecutive way, or "one category is nested in the previous one."[54] For example, if the response is the number of alcoholic drinks a person drinks in a day, the first level is whether the person is a

drinker or not. The next level could be defined as 1-3 drinks per day, and the following level as 4-6 drinks per day. The risk group at each level changes accordingly.[54]

Discussion of Prognostic Clinical Prediction Rules (CPR) for the Low Back Pain

A prognostic Clinical Prediction Rule (CPR) aims to identify specific clinical findings that have shown significant predictability in the determination of a patient outcome when given a specific treatment.[55] Prognostic CPRs are generally developed using a three-step procedure with the purpose to predict success or failure with an intervention.[55] First, CPRs are generally derived using multivariate statistical methods to determine the predictive ability of groups of specific clinical or descriptive variables.[55] Second, a randomized control trial (RCT) is used to validate the CPR and to reduce the possibility that the originally chosen variables for the CPR were not chosen by chance in the derivation stage.[56] Finally, an impact analysis of the CPR is performed to determine if the CPR can improve care, reduce costs, and accurately identify the originally stated objective.[56] A discussion of selected examples of prognostic CPRs for low back pain (LBP) follows.

George et al[57] examined the development of two CPRs for LBP: (1) to predict pain 6 months after receiving specific exercise for LBP; and (2) to predict disability 6 months after receiving specific exercises for LBP. For pain, the researchers found that a lack of centralization phenomenon at baseline predicted a higher pain intensity 6 months later.[57] For disability, the researchers reported patients who had a lack of centralization phenomenon and increased fear-avoidance beliefs about work predicted higher disability after 6 months.[57] These CPRs remain in the derivation stage of CPRs.

Hancock et al[58] sought to develop a CPR to predict the time to recover from acute LBP. Variables included in the rule were baseline pain ≤ 7/10, duration of episode ≤ 5 days, and ≤ 1 previous episodes. When all 3 variables were present, the median days to recovery was 6 days (95% CI [4, 8]).[58] If none of the variables were present, then the median days to recovery was 22 days (95% CI [11, 33]).[58] This CPR also remains in the derivation stages of CPRs.

To date, only three CPRs for LBP have undergone validation status. The first is the "Cassandra rule," a CPR used to determine patients' long-term risk for low back pain related disability. The "Cassandra rule" was formulated from a population of LBP patients in a primary care setting.[59] The prediction variables were determined using the Symptoms Checklist 90 Revised (SL-90-R) questionnaire. Measurements of depression and of somatization from the SL-90-R were used to classify patients considered at risk for having a score on the Roland-Morris Disability questionnaire (RMDQ) indicating 50% or greater disability at 2 years.[60] The rule was validated in a prospective cohort study of 860 patients absent from work secondary to LBP[61] and in a prospective cohort study of 1,262 patients presenting to the emergency room with nonspecific LBP.[62]

The other two CPRs that have undergone validation are the five-item and two-item Flynn manipulation CPRs, each of which are comprised of specific guidelines for the appropriateness of lumbopelvic manipulation.[60] The CPRs purport that patients deemed appropriate for the manipulation criteria are predicted to have positive outcomes in terms of back disability.[60] The five-item CPR for lumbopelvic manipulation includes the following five variables: symptom duration (< 16 days), fear-avoidance beliefs (work subscale < 19), lumbar hypomobility (≥ 1 hypomobile segment), hip internal rotation range of motion (≥ 1 hip > 35º IR), and no symptoms distal to the knee.[63] Patients who demonstrate at least 4 of the 5 variables are considered

appropriate for lumbopelvic manipulation, with a successful outcome defined as an improvement of at least 50% on the ODI.[63] At least nine validation studies have investigated this CPR in terms of generalizability and whether the CPR is a treatment effect modifier.[60] A treatment effect modifier is generally a patient attribute at baseline that can be used to classify subgroups of patients who experience different outcomes from a specified intervention.[58] A predictive CPR that is comprised of individual patient attributes, such as the Flynn manipulation CPR,[64] can itself be considered a treatment effect modifier.[60] Positive baseline status of the Flynn manipulation CPR, generally considered as ≥ 4 of the above conditions met, has been shown to predict reduced disability in numerous studies,[60] but most specifically when the definition of success was a 50% improvement on the ODI.[4]

The two-item CPR for lumbopelvic manipulation has been proposed that uses the variables of duration (< 16 days) and distribution of symptoms (none distal to the knee).[65] The two-item version has been found to correctly classify patients as the same as the five-item version 84% of the time.[60] Patients treated with both thrust and non-thrust manipulation that met the variables of the two-item version have been shown to experience greater clinical improvements than those who did not receive manipulation.[66] Furthermore, the patients receiving thrust manipulation gained similar clinical outcomes more efficiently than those who received non-thrust manipulation.[66] This favorable effect of thrust manipulation when meeting the criteria for the two-item manipulation CPR has been corroborated elsewhere in the literature.[67] Both validation studies of the two-item CPR excluded patients who were negative for the criteria, which prevents the evaluation of the rule's predictive accuracy in identifying patients with alternate likelihoods of improving.[60] The eligibility for the five-item manipulation CPR has been

found to be a prognostic indicator of multiple outcomes including improved ODI score, improved Numeric Pain Rating Scale change scores, total visits, and reported rate of recovery.[68]

A 2015 systematic review identified 30 prognostic or predictive CPRs, with only the above 3 mentioned CPRs reaching the validation stage.[60] Because the others have not been validated, they cannot be recommended for clinical use at this time. However, clinicians may wish to use their knowledge of identified individual predictors that make up portions of these rules to carefully inform their prognostic clinical judgements.[60] The use of a checklist has been suggested to create consensus-based guidelines to improve the standards of developing CPRs in further research.[69] The novel addition of our current capstone project to the possible development of a prognostic CPR for LBP includes costs, which is rarely if ever captured in the existing literature.

References

1. Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results-guidance for future prognosis reviews. *J Clin Epidemiol*. 2009;62(8):781-796.e1. doi:10.1016/j.jclinepi.2008.09.004.
2. Bekkering GE, Hendriks HJM, van Tulder MW, et al. Prognostic factors for low back pain in patients referred for physiotherapy: comparing outcomes and varying modeling techniques. *Spine (Phila Pa 1976)*. 2005;30(16):1881-1886. doi:10.1097/01.brs.0000173901.64181.db.
3. Cook CE, Learman KE, O'Halloran BJ, et al. Which prognostic factors for low back pain are generic predictors of outcome across a range of recovery domains? *Phys Ther*. 2013;93(1):32-40. doi:10.2522/ptj.20120216.
4. Schwind J, Learman K, O'Halloran B, Showalter C, Cook C. Different minimally important clinical difference (MCID) scores lead to different clinical prediction rules for the Oswestry disability index for the same sample of patients. *J Man Manip Ther*. 2013;21(2):71-78. doi:10.1179/2042618613Y.0000000028.
5. Haxby Abbott J, Kingan E-M. Accuracy of physical therapists' prognosis of low back pain from the clinical examination: a prospective cohort study. *J Man Manip Ther*. 2014;22(3):154-161. doi:10.1179/2042618613Y.0000000045.
6. Lukas C, Dougados M, Combe B. Factors associated with a bad functional prognosis in early inflammatory back pain: results from the DESIR cohort. *RMD open*. 2016;2(1):e000204. doi:10.1136/rmdopen-2015-000204.
7. Campbell P, Foster NE, Thomas E, Dunn KM. Prognostic indicators of low back pain in primary care: Five-Year Prospective Study. *J Pain*. 2013;14(8):873-883. doi:10.1016/j.jpain.2013.03.013.
8. Dunn KM, Jordan KP, Croft PR. Contributions of prognostic factors for poor outcome in primary care low back pain patients. *Eur J Pain*. 2011;15(3):313-319. doi:10.1016/j.ejpain.2010.07.008.
9. Steenstra IA, Verbeek JH, Heymans MW, Bongers PM. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. *Occup Environ Med*. 2005;62(12):851-860. doi:10.1136/oem.2004.015842.
10. Traeger AC, Henschke N, Hübscher M, et al. Estimating the Risk of Chronic Pain: Development and Validation of a Prognostic Model (PICKUP) for Patients with Acute Low Back Pain. *PLOS Med*. 2016;13(5):e1002019. doi:10.1371/journal.pmed.1002019.
11. Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH. Can rate of recovery be predicted in patients with acute low back pain? Development of a clinical prediction rule. *Eur J Pain*. 2009;13(1):51-55. doi:10.1016/j.ejpain.2008.03.007.
12. Steffens D, Hancock MJ, Maher CG, et al. Prognosis of chronic low back pain in patients presenting to a private community-based group exercise program. *Eur Spine J*. 2014;23(1):113-119. doi:10.1007/s00586-013-2846-x.
13. George S, Bialosky J, Donald D. The centralization phenomenon and fear-avoidance beliefs as prognostic factors for acute low back pain. *J Orthop Sports Phys Ther*. 2005;35(12):844-845-847.
14. Grotle M, Foster NE, Dunn KM, Croft P. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *Pain*.

2010;151(3):790-797. doi:10.1016/j.pain.2010.09.014.
15. Shaw WS, Pransky G, Fitzgerald TE. Early prognosis for low back disability: intervention strategies for health care providers. *Disabil Rehabil*. 2001;23(18):815-828. doi:10.1080/09638280110066280.
16. Harris SA, Rampersaud YR. The importance of identifying and modifying unemployment predictor variables in the evolution of a novel model of care for low back pain in the general population. *Spine J*. 2016;16(1):16-22. doi:10.1016/j.spinee.2015.09.034.
17. Burton AK, McClune TD, Clarke RD, Main CJ. Long-term follow-up of patients with low back pain attending for manipulative care: Outcomes and predictors. *Man Ther*. 2004;9(1):30-35. doi:10.1016/S1356-689X(03)00052-3.
18. Foster NE, Bishop A, Thomas E, et al. Illness perceptions of low back pain patients in primary care: what are they, do they change and are they associated with outcome? *Pain*. 2008;136(1-2):177-187. doi:10.1016/j.pain.2007.12.007.
19. Schultz IZ, Crook J, Meloche GR, et al. Psychosocial factors predictive of occupational low back disability: Towards development of a return-to-work model. *Pain*. 2004;107(1-2):77-85. doi:10.1016/j.pain.2003.09.019.
20. Henschke N, Maher CG, Refshauge KM, et al. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. *BMJ*. 2008;337:a171. http://www.ncbi.nlm.nih.gov/pubmed/18614473. Accessed January 15, 2017.
21. Ramond A, Bouton C, Richard I, et al. Psychosocial risk factors for chronic low back pain in primary care-a systematic review. *Fam Pract*. 2011;28(1):12-21. doi:10.1093/fampra/cmq072.
22. Finch, E. Brooks, D. Stratford, P. Mayo N. *Physical Rehabilitation Outcome Measures : A Guide to Enhanced Clinical Decision Making*. 2nd ed. Baltimore, MD: Lippincot Williams and WIlkins; 2002.
23. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy*. 1980;66(8):271-273. http://www.ncbi.nlm.nih.gov/pubmed/6450426. Accessed March 2, 2017.
24. Beurskens AJ, de Vet HC, Köke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*. 1996;65(1):71-76. http://www.ncbi.nlm.nih.gov/pubmed/8826492. Accessed March 2, 2017.
25. Fisher K, Johnston M. Validation of the Oswestry Low Back Pain Disability Questionnaire, its sensitivity as a measure of change following treatment and its relationship with other aspects of the chronic pain experience. *Physiother Theory Pract*. 1997;13(1):67-80. http://cirrie.buffalo.edu/database/4598/. Accessed March 2, 2017.
26. Grönblad M, Hupli M, Wennerstrand P, et al. Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin J Pain*. 1993;9(3):189-195. http://www.ncbi.nlm.nih.gov/pubmed/8219519. Accessed March 2, 2017.
27. DeVine J, Norvell DC, Ecker E, et al. Evaluating the Correlation and Responsiveness of Patient-Reported Pain With Function and Quality-of-Life Outcomes After Spine Surgery. *Spine (Phila Pa 1976)*. 2011;36(21 Suppl):S69-S74. doi:10.1097/BRS.0b013e31822ef6de.
28. Grönblad M, Järvinen E, Hurri H, Hupli M, Karaharju EO. Relationship of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) with three dynamic physical tests in a group of patients with chronic low-back and leg pain. *Clin J Pain*. 1994;10(3):197-203. http://www.ncbi.nlm.nih.gov/pubmed/7833577. Accessed

March 2, 2017.
29. Haas M, Nyiendo J. Diagnostic utility of the McGill Pain Questionnaire and the Oswestry Disability Questionnaire for classification of low back pain syndromes. *J Manipulative Physiol Ther*. 1992;15(2):90-98. http://www.ncbi.nlm.nih.gov/pubmed/1532978. Accessed March 2, 2017.
30. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)*. 2000;25(22):2940-52; discussion 2952. doi:10.1097/00007632-200011150-00017.
31. Grevitt M, Khazim R, Webb J, Mulholland R, Shepperd J. The short form-36 health survey questionnaire in spine surgery. *J Bone Joint Surg Br*. 1997;79(1):48-52. http://www.ncbi.nlm.nih.gov/pubmed/9020444. Accessed March 2, 2017.
32. Loisel P, Md, Poitras S, et al. Is Work Status of Low Back Pain Patients Best Described by an Automated Device or by a Questionnaire? *Spine (Phila Pa 1976)*. 1998;23(15):1588-1594. http://www.asipp.org/reference/41Loisel.pdf. Accessed March 2, 2017.
33. Kuukkanen T, Mälkiä E. Muscular performance after a 3 month progressive physical exercise program and 9 month follow-up in subjects with low back pain. A controlled study. *Scand J Med Sci Sports*. 1996;6(2):112-121. http://www.ncbi.nlm.nih.gov/pubmed/8809928. Accessed March 2, 2017.
34. Sufka A, Hauger B, Trenary M, et al. Centralization of Low Back Pain and Perceived Functional Outcome. *J Orthop Sport Phys Ther*. 1998;27(3):205-212. doi:10.2519/jospt.1998.27.3.205.
35. Dawson AP, Steele EJ, Hodges PW, Stewart S. Utility of the Oswestry Disability Index for studies of back pain related disability in nurses: Evaluation of psychometric and measurement properties. *Int J Nurs Stud*. 2010;47(5):604-607. doi:10.1016/j.ijnurstu.2009.10.013.
36. Davidson, Megan. Keating J. A Comparison of Five Low Back Disability Questionnaires: Reliability and Responsiveness. *Phys Ther*. 2002;82(1):8-24. http://vb3lk7eb4t.search.serialssolutions.com.libproxy.lib.unc.edu/?sid=Entrez:PubMed&id=pmid:11784274. Accessed March 2, 2017.
37. Kopec JA, Esdaile JM, Abrahamowicz M, et al. The Quebec Back Pain Disability Scale. Measurement properties. *Spine (Phila Pa 1976)*. 1995;20(3):341-352. http://www.ncbi.nlm.nih.gov/pubmed/7732471. Accessed March 2, 2017.
38. Beattie P, Maher C. The role of functional status questionnaires for low back pain. *Aust J Physiother*. 1997;43(1):29-38. doi:10.1016/S0004-9514(14)60400-1.
39. Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference. Low back pain: outcome measures. *J Rheumatol*. 2001;28(2):431-438. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=11246692&retmode=ref&cmd=prlinks.
40. Fritz JM, Hebert J, Koppenhaver S, Parent E. Beyond Minimally Important Change. *Spine (Phila Pa 1976)*. 2009;34(25):2803-2809. doi:10.1097/BRS.0b013e3181ae2bd4.
41. Maughan EF, Lewis JS. Outcome measures in chronic low back pain. *Eur Spine J*. 2010;19(9):1484-1494. doi:10.1007/s00586-010-1353-6.
42. Parker SL, Adogwa O, Mendenhall SK, et al. Determination of minimum clinically important difference (MCID) in pain, disability, and quality of life after revision fusion for symptomatic pseudoarthrosis. *Spine J*. 2012;12(12):1122-1128. doi:10.1016/j.spinee.2012.10.006.

43. Resnik L, Dobrzykowski E. Guide to outcomes measurement for patients with low back pain syndromes. *J Orthop Sports Phys Ther*. 2003;33(6):307-316-318. doi:10.2519/jospt.2003.33.6.307.
44. Eleswarapu A, Divi S, Dirschl D, Mok J. How Effective is Physical Therapy for Common Low Back Pain Diagnoses? A Multivariate Analysis of 4597 Patients. *Spine (Phila Pa 1976)*. 2016;41(16):1325-1329. doi:10.1097/BRS.0000000000001506.
45. Hägg O, Fritzell P, Nordwall a. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J*. 2003;12(1):12-20. doi:10.1007/s00586-002-0464-0.
46. Ostelo RWJG, de Vet HCW. Clinically important outcomes in low back pain. *Best Pract Res Clin Rheumatol*. 2005;19(4):593-607. doi:10.1016/j.berh.2005.03.003.
47. Khan M, Bashir M, Kumar R, Enam S. Enduring improvement in Oswestry Disability Index outcomes following lumbar microscopic interlaminar decompression: An appraisal of prospectively collected patient outcomes. *J Craniovertebr Junction Spine*. 2015;6(4):195. doi:10.4103/0974-8237.167881.
48. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting Change Scores for Pain and Functional Status in Low Back Pain. *Spine (Phila Pa 1976)*. 2008;33(1):90-94. doi:10.1097/BRS.0b013e31815e3a10.
49. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskelet Disord*. 2006;7:82. doi:10.1186/1471-2474-7-82.
50. Katz MH. Multivariable Analysis: A Primer for Readers of Medical Research. *Ann Intern Med*. 2003;138(8):644-650. doi:10.7326/0003-4819-138-8-200304150-00012.
51. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36. doi:10.1148/radiology.143.1.7063747.
52. Hosmer, DW. Lemeshow S. *Applied Logistic Regression*. (2, ed.). New York: Wiley; 2000.
53. Steyerberg EW, Eijkemans MJ, Harrell Jr. FE, Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Mak*. 2001;21(1):45-56. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11206946.
54. Hierarchical Multinomial Models. MathWorks. https://www.mathworks.com/help/stats/hierarchical-multinomial-models.html. Published 2017. Accessed March 20, 2017.
55. Cook C. Potential Pitfalls of Clinical Prediction Rules. *J Man Manip Ther*. 2008;16(2):69-71. doi:10.1179/106698108790818477.
56. Childs JD, Cleland JA. Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Phys Ther*. 2006;86(1):122-131. http://www.ncbi.nlm.nih.gov/pubmed/16386067. Accessed March 3, 2017.
57. George SZ, Bialosky JE, Donald DA. The centralization phenomenon and fear-avoidance beliefs as prognostic factors for acute low back pain: a preliminary investigation involving patients classified for specific exercise. *J Orthop Sports Phys Ther*. 2005;35(9):580-588. doi:10.2519/jospt.2005.35.9.580.

58. Hancock M, Herbert RD, Maher CG. A guide to interpretation of studies investigating subgroups of responders to physical therapy interventions. *Phys Ther*. 2009;89(7):698-704. doi:10.2522/ptj.20080351.
59. Dionne CE, Koepsell TD, Von Korff M, Deyo RA, Barlow WE, Checkoway H. Predicting long-term functional limitations among back pain patients in primary care settings. *J Clin Epidemiol*. 1997;50(1):31-43. http://www.ncbi.nlm.nih.gov/pubmed/9048688. Accessed April 4, 2017.
60. Haskins R, Osmotherly PG, Rivett DA. Validation and impact analysis of prognostic clinical prediction rules for low back pain is needed: a systematic review. *J Clin Epidemiol*. 2015;68(7):821-832. doi:10.1016/j.jclinepi.2015.02.003.
61. Dionne C. Psychological distress confirmed as predictor of long-term back-related functional limitations in primary care settings. *J Clin Epidemiol*. 2005;58:714-718.
62. Dionne CE, Le Sage N, Franche R-L, Dorval M, Bombardier C, Deyo RA. Five questions predicted long-term, severe, back-related functional limitations: evidence from three large prospective studies. *J Clin Epidemiol*. 2011;64(1):54-66. doi:10.1016/j.jclinepi.2010.02.004.
63. Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine (Phila Pa 1976)*. 2002;27(24):2835-2843. doi:10.1097/01.BRS.0000035681.33747.8D.
64. Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med*. 2004;141(12):920-928. http://www.ncbi.nlm.nih.gov/pubmed/15611489. Accessed April 3, 2017.
65. Fritz JM, Childs JD, Flynn TW. Pragmatic application of a clinical prediction rule in primary care to identify patients with low back pain with a good prognosis following a brief spinal manipulation intervention. *BMC Fam Pract*. 2005;6(1):29. doi:10.1186/1471-2296-6-29.
66. Fritz JM, Brennan GP, Leaman H, Fogarty WT, Kirk DH, Adams A. Does the evidence for spinal manipulation translate into better outcomes in routine clinical care for patients with occupational low back pain? A case-control study. *Spine J*. 2004;6(3):289-295. doi:10.1016/j.spinee.2005.11.002.
67. Hallegraeff JM, Hallegraeff HJM, de Greef M, Winters JC, Lucas C. Manipulative therapy and clinical prediction criteria in treatment of acute nonspecific low back pain. *Percept Mot Skills*. 2009;108(1):196-208. doi:10.2466/PMS.108.1.196-208.
68. Cook CE, Learman KE, O'Halloran BJ, et al. Which prognostic factors for low back pain are generic predictors of outcome across a range of recovery domains? *Phys Ther*. 2013;93(1):32-40. doi:10.2522/ptj.20120216.
69. Cook C, Brismée J-M, Pietrobon R, Sizer P, Hegedus E, Riddle DL. Development of a Quality Checklist Using Delphi Methods for Prescriptive Clinical Prediction Rules: the QUADCPR. *J Manipulative Physiol Ther*. 2010;33(1):29-41. doi:10.1016/j.jmpt.2009.11.010.