

## **VALIDITY AND RELIABILITY OF THE NUMERIC RATING SCALE IN NON-MYOGENIC LOW BACK PAIN PATIENTS**

Faradila Risky Adha <sup>1</sup>, Dwi Rosella Komalasari <sup>2</sup>

<sup>1,2</sup>*Department of Physiotherapy, Faculty of Health Science, University Muhammadiyah Surakarta, Indonesia*

\*Corresponding author: Dwi Rosella Komalasari, [drks133@ums.ac.id](mailto:drks133@ums.ac.id)

### **Abstract**

**Introduction:** Pain is the main problem in LBP that can reduce quality of life. Several tools are used to determine the pain, including Numeric Rating Scale (NRS). NRS need validity and reliability to evaluate pain. **Objective:** To determine the validity and intra-rater and inter-rater reliability of the NRS in Non-Myogenic Low Back Pain patients. **Methods:** This research was an observational study with a methodological approach using a purposive sampling technique. The total sample were 55 people. **Results:** The validity of NRS showed well with  $p < 0.001$  and  $r$  calculated was higher than 0.9 (higher than  $r$  table = 0.260). SEM value: 0.20 and MDC: 0.55. While, intra-rater and inter-rater reliability of NRS was very high, with Cronbach Alpha and ICC being more than 0.9 and  $p < 0.05$ . **Conclusion:** The NRS was valid and reliable in terms of intra-rater and inter-rater for measuring the level of pain in non-myogenic low back pain.

**Keywords:** reliability, validity, low back pain, numeric rating scale

## **Introduction**

Low Back Pain (LBP) is a very common health problem and one of the main causes of disability worldwide. LBP is defined as pain that occurs between the rib margin and the crease of the buttocks which can be accompanied by radiation to the legs or numbness (1). This pain sensation can interfere with a person's work performance, including reducing a person's well-being (2). Based on the reasons, LBP is divided into two, namely myogenic LBP and non-myogenic LBP. Myogenic LBP is a form of spinal structural disorder that usually occurs due to trauma such as strains, sprains and lower back muscle spasms (3). Meanwhile, non-myogenic LBP is pain caused by structural problems in the spine or when pain radiates from other parts of the body (disc herniation, vertebral fractures, infections and tumors) (4).

Pain was the main complaint of LBP, it was found that around 70% of adult individuals had experienced episodes of lower back pain during their lifetime with a one-year prevalence of around 15% -45%, while the peak prevalence occurred between the ages of 35 to 55 years (5). In Central Java the prevalence of LBP in people aged over 65 years was around 18% in men and 14% in women (6). Non-myogenic LBP often occurs due to degenerative conditions and usually results in functional disorders that make the patient experience limitations in movement such as sitting for relatively long periods, bending, squatting, and twisting (7). The location of the pain felt by each individual is different, but usually occurs at L4-L5 to L5-S1, and can be accompanied by spreading to the feet (8).

There are many measuring tools used to evaluate the level of pain in non-myogenic LBP patients, one of which is the Numeric Rating Scale (NRS). NRS is a numbered version of the VAS where patients can choose one number that best reflects the intensity of pain, a score of 0 "no pain" to 10 "unbearable pain" (9). A study stated that the NRS is more popular with parents in measuring pain intensity, but is considered less objective because it provides inconsistent results in LBP patients (9). Based on a report from (10), the NRS has good reliability and validity in measuring pain intensity in LBP patients with a correlation coefficient result in test-retest reliability of 0.991. Reliability is the main requirement for a measuring instrument to be used as an evaluation tool.

Based on this description, the NRS can be a reference for measuring pain scales used in non-myogenic LBP patients. Moreover, in Indonesia, there have not been many reports regarding the reliability of the Numeric Rating Scale in measuring pain levels in non-myogenic LBP patients. So, this research aimed to determine the reliability and validity of the Numeric Rating Scale measuring instrument in terms of intra-rater and inter-rater in non-myogenic LBP patients.

## **Methods**

This research was carried out based on research permission from the Health Research

Ethics Committee of Dr Moewardi Regional General Hospital with number 111/ I/ HREC/ 2024. This type of research uses an observational study with a methodological research approach to determine the reliability of the Numeric Rating Scale as a tool for measuring pain intensity in non-myogenic LBP patients. This research was carried out at Pandan Arang Boyolali Hospital from February to March 2024. Sampling was carried out using a non-probability sampling technique, namely purposive sampling. An evaluation process first uses a sample screening form which contains questions about pain of the location of low back pain, radiating pain, as well as lumbar flexion and extension pain, medical diagnosis, and history of complaints. After that, the respondent fills out (informed consent) which is a consent form to become a respondent.

The inclusion criteria are as follows (1) age 40-75 years, (2) male and female, (3) pain > 3 or more when moving lumbar flexion and extension using NRS, (4) diagnosis of low back pain > 3 months (chronic) (5) Diagnosed with HNP, spondylitis, spondylosis, spondylolisthesis;. Meanwhile, the exclusion criteria are; (1) unable to communicate well (2) patient has neurological diseases such as stroke, Parkinson's, cerebral ataxia, or impaired coordination (3) post-surgery (4) the patient has spinal abnormalities such as tumors, and infection (spinal TB) (5) ) patients who cancelled their availability became study respondents.

The sample size calculation is based on the (11) formula with the Sample Size Calculation formula as follows ;

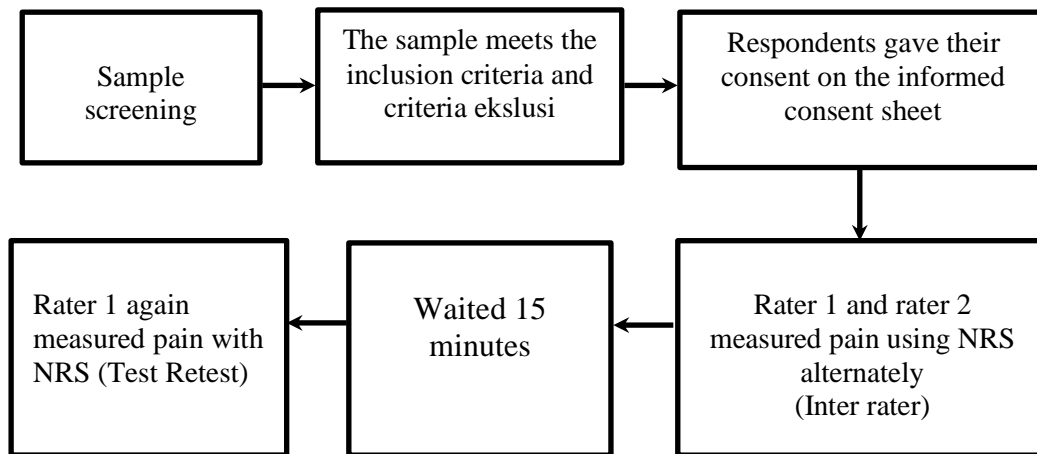
**Tabel 1.** Sampel Size Calculation formula

<i>Intraclass Correlation Coefficient (ICC) - Hypothesis Testing</i>	
Minimum Expectable Reliability (ICC) ( $\rho_0$ )	0.60
Expectable Reliability ( $\rho_1$ )	0.8
Significance Level ( $\alpha$ )	0.05 (Two-tailed)
Power (1 - $\beta$ )	80 %
Number Of Rater/repetitions per subject (k)	2
Expected dropout rate	10 %
Sample Size, n	49
Sampel Size (with 10% drop out), $n_{drop}$	55

So the number of samples in this study that met the inclusion and exclusion criteria was 55 people (11).

The research instrument consists of (1) Numeric Rating Scale, which is a measuring tool that can be used as an evaluation tool to measure the intensity of pain during lumbar flexion and extension movements. How to determine the score, patients can choose one number that best reflects the intensity of pain, if 0 is "no pain" and 10 is "unbearable pain" (9). (2) Research examination form. This form contains questions that discuss pain.

There are two types of reliability tests, namely test-retest or intra-rater reliability and inter-rater reliability. Inter-rater measurements were carried out first by two raters simultaneously (rater 1 and rater 2). The patient is briefly explained how the NRS works in describing the perceived pain value. Patients can choose a number that best reflects the intensity of the pain, where 0 is “no pain” to 10 “unbearable pain. The patient is asked to perform lumbar flexion and extension movements, then the patient is asked to choose one number according to the pain felt during the flexion and extension movements. Then after 15 minutes, rater 1 took intra-rater measurements using the same testing procedure. The research flow is explained in Figure 1.



**Figure 1.** Research Flow

This research uses data analysis techniques in the form of (1) Univariate tests are used to see the frequency distribution of one variable which is described in the form of frequency (n), average (mean), minimum (min), maximum (max), and standard deviation (SD) (12), (2) Reliability tests consist of intra rater and inter-rater who present data in the form of Cronbach Alpha results which are used to determine the consistency of repeated testing and measurements which are interpreted as follows; <0.5 is unacceptable, 0.5–0.6 is poor, 0.6–0.7 is doubtful, 0.7–0.8 is acceptable, 0.8–0.9 is good, and more than 0.9 is very good (13). Intra Class Correlation (ICC) is used to determine agreement between 2 or more raters which is interpreted as <0.5 low reliability, 0.5 – 0.75 medium reliability, 0.75 – 0.9 high reliability, >0.9 high reliability (14), (3) Validity test using the Pearson Product Moment test. In this study, we looked at the validity of the NRS scale factors. NRS is valid if the Pearson Product Moment test is  $p < 0.05$  and the r-count is greater than the r-table ( $r = 0.260$ ) for a sample size of 55 people.

## Results

**Table 2.** Characteristic data of non-myogenic LBP patients

Variable	Min	Max	Mean $\pm$ SD	N (%)
Age	45	75	62.02 $\pm$ 7.04	-
Gender				
Female	-	-	-	28 (50.9%)
Men	-	-	-	27 (49.1%)
Job				
Work	-	-	-	25 (45.5%)
Not Working	-	-	-	30 (54.5 %)
Diagnosis				
HNP	-	-	-	54 (98.2%)
Spondylolisthesis	-	-	-	1 (1.8%)
Pain (NRS)	3	9	5.4 $\pm$ 1.4	-

Table 2 shows that the average respondent is 62 years old, with the number of female respondents being 28 compared to 27 male respondents, almost female and male respondents balanced. Non-working status is higher than working status. The cause of LBP condition is dominated by HNP (98.2%) with an average NRS pain scale is 5.4.

**Table 3.** Test reliability and validity of visual analogue scale

Test Variable	<i>Cronbach's Alpha</i>	ICC	95%CI	<i>p-value</i>	<i>r-value</i>
Intra Rater (T1-T2)	0.970	0.970	0.949-0.982	<0.001	-
Inter Rater	0.968	0.968	0.945-0.981	<0.001	-
Intra rater validity	-	-	-	<0.001	0.942
Inter-rater validity	-	-	-	<0.001	0.939

Table 3 shows the intra-rater and inter-rater reliability tests of the NRS in non-myogenic LBP patients. With Cronbach's Alpha and ICC values above 0.9 which means the NRS scale has very high reliability (very good). Intra rater is carried out at intervals of 15 minutes in 2 repeated times. Meanwhile, inter-rater is carried out by 2 raters at a time which at the same time. The Pearson Product Moment test shows positive results significantly ( $p < 0.001$ ) for intra-rater and inter-rater. R-calculated value explains the validity of the NRS, with an *r*-calculated value above 0.9 for both intra-rater and inter-rater, which is greater than the *r*-table ( $r = 0.260$ ) for the number 55 patients. So the NRS is valid as a pain measurement tool for people non-myogenic LBP patients.

**Table 4.** SEM and MDC<sub>95</sub> values

	<b>Score</b>
SEM	0.20
MDC95	0.55

The SEM search yielded a value of 0.20, indicating a very low probability of error. Moreover, the MDC95 search resulted in a 0.55 value, which means that the probability of a change in this study is small.

## **Discussion**

The study examined the characteristics of respondents, who had an average age of 62 years, with a minimum age of 45 years and a maximum age of 75 years. At the age of 60 years, muscle strength begins to decline, which triggers other complaints such as the shifting of the vertebral bodies, causing pain that radiates to the lower back (15). These age-related changes highlight the importance of assessing pain and its underlying causes in older adults. This occurs due to degeneration caused by tissue damage, tissue change into scar tissue, and reduced body fluids. This results in reduced stability in the muscles, resulting in pain (16). Based on gender, women and men are almost the same, with only one difference, namely 28 women and 27 men. Because women's muscle capacity tends to be lower than men's muscle and the level of the hormone estrogen decreases during menopause which has an impact on decreasing bone density (17).

In terms of employment, more people are not working 30 (54.5%) than those who are working 25 (45.5%). So it can be concluded that the pain felt can interfere with a person's work performance, including reducing an individual's well-being (2). HNP is a common cause of low back pain which is the main reason someone undergoes spinal surgery, especially in working-age individuals (18). The majority of respondents suffered from HNP 54 (98.2%) and only 1 person (1.8%) suffered from spondylolisthesis, with an average age of respondents of 60 years. This occurs due to unergonomic body posture when doing work (19). In this case, pain is the main complaint that affects daily activities with the average NRS pain scale felt by respondents being 5.4 with the interpretation of moderate pain a score of 3 for minimal pain and a score of 9 for maximum pain. A pain score of more than 3 means that in this condition a person begins to feel that the pain has hampered activities and disrupted the person's quality of life with the interpretation of the pain being moderate to severe and included in the chronic category or lasting more than 3 months (2).

Intra-rater reliability is the repetition of pain measurements using NRS carried out by raters on non-myogenic LBP patients at 15-minute intervals. In this study, Cronbach's Alpha value was 0.970, which means there is very high consistency (13). The ICC value obtained a result of 0.970

which shows high reliability (14). This research is in line with previous research that intra-rater reliability of the NRS has good reliability in measuring pain intensity in LBP patients with an ICC result of 0.991, meaning that the NRS has very high consistency if repeated by the same rater (10). Inter-rater is a measurement of pain intensity carried out by two raters alternately without any time lag. This test aims to see whether there is agreement or approval of the measurements that have been carried out (20). In this study, the Cronbach's alpha value was 0.968 with an Intraclass Correlation Coefficient (ICC) of more than 0.9, which means that rater 1 and rater 2 have a very high (excellent) agreement between test 1 and test 2 (14).

Factor validity is the Pearson product-moment test or calculated r-value on intra-rater and inter-rater reliability. This test is used to see the validity of a measuring instrument, where the correlation between tests 1 and 2 obtained a result of  $r=0.942$  which is greater than the r-table (0.260) for a sample size of 55. This shows that the NRS is valid even though it is carried out by the same rater. Meanwhile, the relationship between rater 1 and rater 2 with a value of  $r=0.939$  shows very strong validity whether carried out by different raters (21).

Several things need to be considered apart from reliability and validity, namely Standard Error Measurement (SEM) and Minimum Detectable Change (MDC). SEM is an estimate of measurement error as proven by the validity value provided it does not exceed the standard deviation. The smaller the standard error measurement, the better the measurement results (22). Minimum Detectable Change (MDC) is an estimate of the smallest change that can be detected in a measurement, with the formula  $SEM \times \sqrt{2} \times 1.96$  (22). In this study, an SEM value of 0.20 was obtained, this is in line with previous research which produced an SEM value of 0.48, meaning that the possibility of error is small (23). Meanwhile, the MDC got a value of 0.55, which is under previous research that the MDC value was 1.33, meaning there is a slight possibility of error (23). This study illustrates that NRS measurements in non-myogenic LBP patients are said to be consistent and reliable and the possibility of errors and changes is very small.

## **Conclusion**

The conclusion of this study is the reliability and validity of the numeric rating scale in non-myogenic LBP patients at the Pandan Arang Regional General Hospital, it can be concluded that the numeric rating scale is reliable and valid in intra-rater and inter-rater cases to measure pain in non-myogenic LBP patients.

## **Acknowledgements**

The author would like to thank the non-myogenic LBP patients at the Pandan Arang Hospital, Boyolali, Central Java who have joined in this research.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Fujii T, Oka H, Takano K, Asada F, Nomura T, Kawamata K, et al. Association between high fear-avoidance beliefs about physical activity and chronic disabling low back pain in nurses in Japan. *BMC Musculoskelet Disord*. 2019;20(1):1–10.
2. Rahmawati A. Risk factor of low back pain. *Jmh*. 2021;3(1):402–6.
3. Zuhri S, Rustanti M. Beda Efektivitas Elektroakupunktur Dengan Tens Terhadap Nyeri Dan Fleksibilitas Lumbal Pada Pasien Nyeri Punggung Bawah Myogenik. *J Keterapian Fis*. 2017;2(2):93–102.
4. World Health Organization. Low back pain. 2023. Available from: <https://www.who.int/news-room/fact-sheets/detail/low-back-pain>
5. Ariyanah NS, Ersila W, J KE, Pekalongan T Di. The Correlation between Working Period and Long Sitting Driving with Low Back Pain Incidence on J & T Expeditionary Couriers in Pekalongan Hubungan Masa Bekerja Dan Lama Duduk Berkendara Dengan Kejadian Low Back Pain Pada. *Pros 16th Univ Res Colloquium*. 2022;43:1113–9.
6. Kurniawidjaja LM, Purnomo E, Maretti N, Pujiriani I. Pengendalian Risiko Ergonomi Kasus Low Back Pain pada Perawat di Rumah Sakit. *Maj Kedokt Bandung*. 2014;46(4):225–33.
7. Widiyarsari K., Ahmad A, Budiman F. Hubungan Faktor Individu Dan Faktor Risiko Ergonomi Dengan Keluhan Low Back Pain (LBP) Pada Penjahit Sektor Usaha Informal CV. Wahyu Langgeng Jakarta Tahun 2014. *J Inohim [Internet]*. 2014;2(2):90–9. Available from: <https://inohim.esaunggul.ac.id/index.php/INO/article/view/107>
8. Rusmayanti MY, Kurniawan SN. Hnp Lumbalis. *JPHV (Journal Pain, Vertigo Headache)*. 2023;4(1):7–11.
9. Chiarotto A, Maxwell LJ, Ostelo RW, Boers M, Tugwell P, Terwee CB. Measurement Properties of Visual Analogue Scale, Numeric Rating Scale, and Pain Severity Subscale of the Brief Pain Inventory in Patients With Low Back Pain: A Systematic Review. *J Pain [Internet]*. 2019;20(3):245–63. Available from: <https://doi.org/10.1016/j.jpain.2018.07.009>
10. Yao M, Xu B, Li Z, Zhu S, Tian Z, Li D, et al. Perbandingan antara skala nyeri punggung bawah untuk pasien dengan herniasi lumbal: validitas , reliabilitas , dan daya tanggap. 2020;2:1–12.
11. Arifin WN. Sample size calculator (web). 2023. Available from: <http://wnarifin.github.io>



12. Notoatmodjo S. *Metodologi Penelitian Kesehatan*. Rineka Cipta Kerja; 2018.
13. Singgaravi P, Yamata Ahmad H. The Effectiveness of Gamification to Improve Writing Simple S-V-O Sentence among Year 2 Pupils. *Int J Acad Res Bus Soc Sci*. 2022;12(2):1–15.
14. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med [Internet]*. 2016;15(2):155–63. Available from: <http://dx.doi.org/10.1016/j.jcm.2016.02.012>
15. Tenny S, Gillis, Christopher C. *Spondylolisthesis*. StatPearls; 2023. Available from: [https://www.ncbi.nlm.nih.gov/translate/goog/books/NBK430767/?\\_x\\_tr\\_sl=en&\\_x\\_tr\\_tl=id&\\_x\\_tr\\_hl=id&\\_x\\_tr\\_pto=tc](https://www.ncbi.nlm.nih.gov/translate/goog/books/NBK430767/?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc)
16. Andini F, Fauzia Andini | Risk Factors of Low Back Pain in. *Work J Major |*. 2015;4:12.
17. Wijayanti F, Oktafany, Ramadhian RM, Saftarina F, Cania E. Kejadian Low Back Pain ( LBP ) pada Penjahit Konveksi di Kelurahan Way Halim Kota Bandar Lampung The Incidence of Low Back Pain at Tailor Convection in Housing Way Halim Bandar Lampung. *J Kedokt*. 2019;8:82–8.
18. Ecklund JM, Babington PW. Lumbar Disc Herniation and Military Rank. *World Neurosurg*. 2014;82(1–2):e157–8. Available from: <http://dx.doi.org/10.1016/j.wneu.2013.07.101>
19. Sitepu D, Sinaga M, Lubis H. Faktor-Faktor Yang Berhubungan Dengan Keluhan Low Back Pain Pada Petani Jeruk Di Desa Dokan Kecamatan Merek Kabupaten Karo. 2015;4:1–9.
20. Denteneer L, Van Daele U, Truijzen S, De Hertogh W, Meirte J, Stassijns G. Reliability of physical functioning tests in patients with low back pain: a systematic review. *Spine J [Internet]*. 2018;18(1):190–207. Available from: <https://doi.org/10.1016/j.spinee.2017.08.257>
21. Sanny B, Dewi R. Pengaruh Net Interest Margin ( NIM ) Terhadap Return on Asset (ROA) Pada. *J E-Bis (Ekonomi Bisnis)*. 2020;4(1):78–87.
22. Dontje ML, Dall PM, Skelton DA, Gill JMR, Chastin SFM. Reliability, minimal detectable change and responsiveness to change: Indicators to select the best method to measure sedentary behaviour in older adults in different study designs. *PLoS One*. 2018;13(4):1–16.
23. Alghadir AH, Anwer S, Iqbal A, Iqbal ZA. Test-retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. *J Pain Res*. 2018;11:851–6.