

Test-retest reliability of the Nordic Musculoskeletal questionnaire (NMQ) in Vietnamese physical therapists

Le Thi Thach Thao

Hong Bang International University, Vietnam

ABSTRACT

Background: The Nordic Musculoskeletal Questionnaire (NMQ) is one of the screening tools to assess pain or discomfort of musculoskeletal disorders. It is widely used in various occupational populations. Many previous studies investigated the reliability of the NMQ, but there are no evidences in Vietnam. Objective: This study evaluated the test-retest reliability of the question about trouble with the locomotive organs in the NMQ in Vietnamese physical therapists (VPTs). Materials and method: It was a cross-sectional study in 10 VPTs by a convenient sampling technique. An online questionnaire was distributed to all participants twice times with a 7-day interval. They reported their musculoskeletal disorders during the last 12 months and the last 7 days and been prevented from doing normal work because of their troubles. The test-retest reliability of dichotomous data produced by the NMQ was accessed by using the kappa coefficient (k) and the level of significance was set at a p -value of less than 0.05. Results: Among 27 binary choice items, there were 5 items where k could not be computed as response were 100% negative on both testing occasions. Eight items showed not significant correlation poor to moderate agreement ($k \leq 0.40$, p -value > 0.05). Fourteen items showed substantial to excellent test-retest reliability ($k = 0.74$ -1.00, p -value < 0.05). Conclusion: The results suggest that the NMQ can be used with acceptable reliability for evaluating the musculoskeletal disorders and labour risks among VPTs.

Keywords: Nordic Musculoskeletal Questionnaire, musculoskeletal disorders, pain, discomfort, reliability

1. BACKGROUND

Nearly four decades ago, Kourinka et al. [1] developed the Nordic Musculoskeletal Questionnaire (NMQ) to assess the musculoskeletal symptoms in ergonomic context and occupational health. It was used widely as an indirect tool to identify the musculoskeletal problems in different occupational groups including physical therapists [2 - 5]. In Vietnam, the NMQ was also used as a screening tool to detect musculoskeletal disorders in different sample populations such as nurses or healthcare workers [6, 7]. Many previous studies showed that the NMQ had good psychometric properties to determine musculoskeletal disorders [8]. However, the test-retest reliability of the questionnaire about trouble with the locomotive organs in the NMQ is unknown in Vietnam. To make this questionnaire commonly used in Vietnam, one study should be done. Therefore, the purpose of this study was to evaluate the test-retest reliability of the NMQ in VPTs. The reliability of a ques-

tionnaire was considered as the consistency of the survey results by using its test-retest reliability in this study. Test-retest reliability is the degree to which test scores remain unchanged when measuring a stable individual characteristic from the same person on different occasions [9]. In this study, the evaluation of test-retest reliability of the self-reported questionnaire was conducted on two occasions. There is no evidence available to select the time interval between questionnaire administrations at occasion 1 and occasion 2. Authors recommended that a retest interval of one or two weeks is commonly used for the evaluation of test-retest reliability, so a 7-day interval was used in the present study [9 - 10].

2. MATERIALS AND METHOD

2.1. Participants

Potential participants included all licensed Vietnamese physical therapists (VPTs) at least 22

Corresponding Author: MSc. Le Thi Thach Thao

Email: thaolt2@hiu.vn

years old working at least 5 days per week and 6 hours per day at hospitals. PTs who were unable to work as physical therapist in the last 6 months and had any past MSDs history of recent within a time span of 1 year (e.g., experienced trauma or undergone orthopedic surgeries without associating work) were excluded.

2.2. Methodology

Study design: A cross-sectional study

Sample size: $n = 10$ VPTs who were working at 6 governmental hospitals ($n = 7$) and 3 private hospitals ($n = 3$) in Vietnam by convenient sampling technique.

Instrument: An online questionnaire about trouble with the locomotive organs in the NMQ consists of two parts. The first part is questions about age, gender, weight and height. The second part is 27 binary choice items for 3 questions about trouble (ache, pain, discomfort) in 9 body parts (i.e., neck, shoulders, elbows, wrist/hands, upper back, low back, hips/thighs, knees and ankles/feet). Instruction describes how to answer the questionnaire and picture shows the approximate position of the nine body parts referred to in the questionnaire. To answer the questions, the participants are asked to rate “yes” or “no” for each area to the following question: “Have you at any time during the last 12 months had trouble (ache, pain, discomfort, numbness)” in - followed by a list of the nine body parts that means an annual prevalence question. If they mark “yes”, then they continue to rate “yes” or “no” for the two questions “Have you at any time during the last 12 months been prevented from doing your normal work (at home or away from home) because of the trouble?” as an annual prevention question and “Have you any trouble at

any time during the last 7 days?” refers to a weekly prevalence question.

Procedure: Test-retest reliability was assessed with the questionnaire completed twice by self-administration at a 7-day interval. An online questionnaire was created on Google Form and then was distributed to 10 participants who were self-administered within a 24-hour and returned after completing. All participants were self-reported all over again after one week.

Statistical Analysis: The Statistical Package for the Social Sciences (SPSS) version 23.0 for Windows was used to analyzed data. Descriptive techniques including number, frequency or mean, percentage (%) or standard deviation (SD). The test-retest reliability of dichotomous data produced by the NMQ was accessed using the *kappa coefficient* (k) and the level of significance was set at a *p-value* of less than 0.05 [11]. The proportion of observed agreement (P_o) for dichotomous data was calculated with 100% representing perfect agreement. *Kappa* (k) is a ratio expressed as:

$$kappa (k) = \frac{P_o - P_e}{1 - P_e}$$

In which, P_o : proportion of observed agreement;
 P_e : Proportion of chance agreement

Range of k is -1.00 to +1.00 wherein k is 0 if agreement equals chance, positive if agreement is better than chance, equals 1.00 if perfect agreement, and is negative if agreement is worse than expected due to chance. When interpreting k strength of agreement, values 0.80 - 1.00 are considered excellent agreement, above 0.60 substantial agreement, 0.40 - 0.60 moderate agreement, and less than 0.4 poor to fair agreement [11].

3. RESULTS

Table 1. Sample Characteristics ($n = 10$)

Variables	Frequency or mean	Percentage or SD
Age (years)	27.9	2.81
Gender		
Male	4	40%
Female	6	60%
Weight (kg)	54.8	6.94
Height (m)	158.2	4.21

The personal characteristics of the sample were presented in Table 1. The mean age of the respondents was 27.9 ± 2.81 years, with range of 23 to 34 years,

with nearly equal distribution regarding their gender. The mean weight and height of the respondents were 54.8 ± 6.94 kg and 158.2 ± 4.21 m respectively.

Table 2. Data of test and retest the NMQ for each question applied to nine body part (n = 10)

Variables		Annual Prevalence		Weekly Prevalence		Annual prevention	
		1 st assessment	2 nd assessment	1 st assessment	2 nd assessment	1 st assessment	2 nd assessment
		Frequency (%)	Frequency (%)	Frequency (%)	Frequency (%)	Frequency (%)	Frequency (%)
Neck	Yes	7 (70)	7 (70)	6 (60)	5 (50)	3 (30)	2 (20)
	No	3 (30)	3 (30)	4 (40)	5 (50)	9 (90)	8 (80)
Shoulders	Yes	5 (50)	2 (20)	2 (20)	2 (20)	1 (10)	1 (10)
	No	5 (50)	8 (80)	8 (80)	8 (80)	9 (90)	9 (90)
Upper Back	Yes	2 (20)	3 (30)	1 (10)	1 (10)	0 (0)	0 (0)
	No	8 (80)	7 (70)	9 (90)	9 (90)	10 (100) *	10 (100) *
Elbows	Yes	1 (10)	1 (10)	0 (0)	0 (0)	0 (0)	0 (0)
	No	9 (90)	9 (90)	10 (100) *	10 (100) *	10 (100) *	10 (100) *
Wrists/Hand	Yes	2 (20)	3 (30)	2 (20)	2 (20)	1 (10)	1 (10)
	No	8 (80)	7 (70)	8 (80)	8 (80)	9 (90)	9 (90)
Lower Back	Yes	7 (70)	7 (70)	1 (10)	4 (40)	3 (30)	3 (30)
	No	3 (30)	3 (30)	9 (90)	6 (60)	7 (70)	7 (70)
Hips/Thighs	Yes	1 (10)	1 (10)	0 (0)	0 (0)	0 (0)	0 (0)
	No	9 (90)	9 (90)	10 (100) *	10 (100) *	10 (100) *	10 (100) *
Knees	Yes	3 (30)	3 (30)	1 (10)	0 (0)	0 (0)	1 (10)
	No	7 (70)	7 (70)	9 (90)	10 (100)	10 (100)	9 (90)
Ankles/Feet	Yes	2 (20)	3 (30)	1 (10)	0 (0)	1 (10)	1 (10)
	No	8 (80)	7 (70)	9 (90)	10 (100)	9 (90)	9 (90)

*100% negative on both testing occasions

Table 2 showed data of the NMQ at the 1st and 2nd assessment for three questions applied to nine body parts. For the first question related to the annual prevalence, there were 4 body parts (i.e., neck, elbows, lower back and knees) that were reported the similarity on both testing occasions, while 5 remaining body parts were reported the difference between test and retest. For the second question according to the weekly prevalence, respondents

reported the same answers in 5 body parts (i.e., shoulders, upper back, elbows, wrists/hands and hips/thighs) in which elbows and hips/thighs were 100% negative on both testing occasions, while their answers differed in 4 remaining body parts. For the last question related to the annual prevention, 7 in 9 body parts were reported the same answers in which upper back, elbows and hips/thighs were 100% negative on both testing occasions.

Table 3. Test–retest reliability of the NMQ in physical therapists (n = 10)

Location of the pain	Annual Prevalence		Weekly prevalence		Annual Prevention	
	kappa	p-value	kappa	p-value	kappa	p-value
Necks	1.00	0.002	0.40	0.197	0.21	0.490
Shoulders	0.40	0.114	1.00	0.002	1.00	0.002
Upper Back	0.74	0.016	1.00	0.002	*	*
Elbows	1.00	0.002	*	*	*	*
Wrists/Hands	0.74	0.016	1.00	0.002	1.00	0.002
Lower Back	1.00	0.002	0.29	0.197	1.00	0.002
Hips/Thighs	1.00	0.002	*	*	*	*
Knees	1.00	0.002	0.00	*	0.00	*
Ankles/Feet	0.21	0.490	0.00	*	1.00	0.002

*A kappa coefficient and p-value could not be calculated
Level of significance p-value<0.05

Reliability statistics for the NMQ was presented in Table 3. Among 27 binary choice items for 3 questions, there were 5 items where k could not be computed as response were 100% negative on both testing occasions. Six items had k low ($k < 0.40$ and $p\text{-value} > 0.05$, not significant correlation poor to fair agreement), including 1 item in the annual prevalence question for ankles/feet, 3 items in the weekly prevalence questions (i.e., lower back, knees and ankles/feet), and 2 others in the annual prevention questions (i.e., neck and knees) but high proportion observed agreement as k is affected by prevalence and uneven data distributions. Two items had not significant correlation moderate agreement between the 1st and 2nd times in the assessment including one item in the annual prevalence question for shoulder ($k = 0.40$, $p\text{-value} = 0.114$) and other one in the weekly prevalence questions for neck ($k = 0.40$, $p\text{-value} = 0.197$). Moreover, the findings showed the significant correlation substantial agreement between two assessment times of 2 items in the annual prevalence question for upper back ($k = 0.74$, $p\text{-value} = 0.016$) and wrists/hands ($k = 0.74$, $p\text{-value} = 0.016$). The significant correlation excellent agreement between the 1st and 2nd assessment in 12 items including 5 items in the annual prevalence question (i.e., neck, elbows, lower back and hips/thighs and knees), 3 items in the weekly prevalence questions (i.e., shoulders, upper back and wrists/hands) and 5 items in the annual prevention questions (i.e., shoulders, wrists/hand, lower back and ankles/feet), ($k = 1.00$, $p\text{-value} = 0.002$).

4. DISCUSSION

Question about annual prevalence and annual prevention had only 2 body parts that showed low reliability, whereas there were 4 body parts in weekly prevalence question had the lower reliability statistics. To explain this, the test-retest reliability with 7-day interval might affect the results because it caused to increase the chance of weekly prevalence differences.

For the questions about trouble with the locomotive organs in the NMQ overall, excluding 5 items where k could not be calculated because participants rated 'No' on both occasions, nearly 64% of the remaining 22 questions had high strength of agreement ($k > 0.60$ and $p\text{-value} < 0.05$, significant correlation substantial to excellent agreement on both testing occasions).

Consistency with this, there were the evidence to stated that the NMQ had good psychometric properties for evaluating the musculoskeletal disorders and possible associated psychosocial and labor risks, and the researchers used the NMQ [8]. One study was done in a group of 50 participants aged 20 to 75 years in Greece. The results showed very good test-retest reliability for almost all items ($k > 0.81$ and $p\text{-value} > 0.001$) and rather good for 2 items addressing neck and elbow disorders in the weekly prevalence question ($k = 0.64$ and $p\text{-value} > 0.001$) [12]. Kahraman T. et al. [13] assessed the test-retest reliability of the NMQ and the findings showed all items had moderate to almost perfect reliability ($k = 0.57 - 0.90$ and $p\text{-value} < 0.001$). Legault E.P. et al. [14] found that range of Kappa value for 27 dichotomous questions was 0.57 to 1.00. Palmer et al. [15] reported that most of the items in the questionnaire showed a good level of agreement while some showed excellent agreement.

Moreover, the popularity of the NMQ was reported with over 259 publications in 42 different countries that applied in three main sectors including human health, manufacturing industries and agriculture [8]. Many previous studies used the NMQ to investigate the prevalence and risk factors of work-related musculoskeletal disorders among physical therapists in different countries such as America, Australia, Malaysia, Nigeria, the UK and Turkey [2-5]. Therefore, this study was feasible as conducted in a sample population of VPTs.

This study might have some limitations. Firstly, the self-administered questionnaire could have some recall bias, so participants might forget to mention all incidents of musculoskeletal pain during last 12 months. Secondly, binary choice questions might have some response bias, so participants can respond untruthfully leading to provide inaccurate answer arises. However, participants in this study were careful instructed the purpose and questions of survey to ensure they can understand to answer all questions. They had 24-hour to rate and resubmit the questionnaire, so we believe that they had enough time to participate in this study. Furthermore, they were informed that their personal information would not be divulged to encourage honest responses.

5. CONCLUSION

The results suggest that the questionnaire about trouble

with the locomotive organs in the Nordic Musculoskeletal Questionnaire can be used with acceptable reliability for evaluating the musculoskeletal disorders among

Vietnamese physical therapists. However, further studies should be done to confirm and expand these findings in a variety of occupational groups.

REFERENCES

- [1] I. Kuorinka, B. Jonsson, A. Kilbom, H. Vinterberg and F. Biering-Sørensen, G. Andersson, et al, "Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms", *Applied ergonomics*, Vol. 18, No. 3, pp. 233-237, 1987.
- [2] J.E. Cromie, V.J. Robertson and M.O Best, "Work-related musculoskeletal disorders in physical therapists: prevalence, severity, risks, and responses", *Physical Therapy*, Vol. 80, No. 4, pp. 336-351, 2000.
- [3] D.J. West and D. Gardner, "Occupational injuries of physiotherapists in North and Central Queensland", *Australian of Journal Physiotherapy*, Vol. 47, No. 3, pp. 179-186, 2001
- [4] Y. Salik and A. Ozcan, "Work-related musculoskeletal disorders: a survey of physical therapists in Izmir-Turkey", *BMC Musculoskeletal Disorders*, Vol. 5, pp. 27, 2004.
- [5] NAM Nordin, J.H. Leonard and N.C. Thye, "Work-related injuries among physiotherapists in public hospitals: a Southeast Asian picture", *Clinics (Sao Paulo)*, Vol. 66, No. 3, pp. 373-378, 2011.
- [6] T.V. Nong, S. Kesornthong and N. Homkham, "Work-related Musculoskeletal Disorders Among Healthcare Workers in a General Provincial Hospital in Vietnam", *International Journal of Ergonomics*, Vol. 10, No. 1, pp. 18-26, 2020.
- [7] L.D. Luan, N.T. Hai, P.T. Xanh, H.T. Giang, P.V. Thuc P, N.M. Hong and et al, "Musculoskeletal Disorders: Prevalence and Associated Factors among District Hospital Nurses in Haiphong, Vietnam", *BioMed Research International*, Vol. 2018, pp. 9, 2018.
- [8] L. López-Aragón, "Applications of the Standardized Nordic Questionnaire: A Review", *Sustainability*, Vol. 9, pp. 9, 2017.
- [9] G. Vilagut, "Test-Retest Reliability", *Encyclopedia of Quality of Life and Well-Being Research*, pp. 6622-6625, 2014.
- [10] R.G. Marx, A. Menezes, L. Horovitz, E.C. Jones and R.F. Warren, "A comparison of two-time intervals for test-retest reliability of health status instruments", *Journal of Clinical Epidemiology*, Vol. 56, No. 8, pp. 730-735, 2003.
- [11] L.G. Portney and M.P. Watkins, *Foundations of clinical research: application to practice*, 3rd ed. Inc., USA: Pearson Education, 2009.
- [12] M. Antonopoulou, C. Ekdahl, M. Sgantzos, N. Antonakis and C. Lionis, "Translation and standardisation into Greek of the standardised general Nordic questionnaire for the musculoskeletal symptoms", *The European Journal of General Practice*, Vol. 10, pp. 33-34, 2004.
- [13] T. Kahraman, A. Genc and E. Goz, "The Nordic Musculoskeletal Questionnaire: Cross-cultural adaptation into Turkish assessing its psychometric properties", *Disability and Rehabilitation*, Vol. 38, pp. 2153-2160, 2016.
- [14] E.P. Legault, V. Cantin and M. Descarreaux, "Assessment of musculoskeletal symptoms and their impacts in the adolescent population: Adaptation and validation of a questionnaire", *BMC Pediatrics*, Vol. 14, pp. 173, 2014.
- [15] K. Palmer, G. Smith, S. Kellingray and C. Cooper, "Repeatability and validity of an upper limb and neck discomfort questionnaire: the utility of the standardized Nordic questionnaire", *Occupational Medicine*, Vol. 49, No. 3, pp. 171-175, 1999.

Đánh giá độ lặp lại của bảng câu hỏi cơ-xương-khớp Bắc Âu trên các chuyên viên vật lý trị liệu Việt Nam

Lê Thị Thạch Thảo

TÓM TẮT

Đặt vấn đề: Bảng câu hỏi cơ-xương-khớp Bắc Âu là một trong những công cụ sàng lọc được sử dụng rộng rãi để đánh giá các rối loạn cơ-xương-khớp. Nhiều nghiên cứu đã đánh giá độ tin cậy của bảng câu hỏi này nhưng chưa thấy các bằng chứng tại Việt Nam. **Mục đích nghiên cứu:** Đánh giá độ tin cậy bằng phương

pháp lặp lại của bảng câu hỏi cơ-xương-khớp Bắc Âu trên các chuyên viên Vật lý trị liệu Việt Nam. Đối tượng và phương pháp nghiên cứu: Là một nghiên cứu cắt ngang trên 10 chuyên viên Vật lý trị liệu bằng kỹ thuật lấy mẫu thuận tiện. Một bảng câu hỏi trực tuyến được gửi cho tất cả người tham gia hai lần cách nhau 7 ngày. Hệ số kappa (k) được dùng để tính độ lặp lại của bảng câu hỏi này và mức ý nghĩa được đặt ở $p < 0.05$. Kết quả: Trong số 27 mục, có 5 mục không thể tính được hệ số k , 8 mục cho thấy mối tương quan không đáng kể ($k \leq 0.40$, $p > 0.05$) và 14 mục cho thấy độ lặp lại là tốt đến hoàn hảo ($k = 0.74 - 1.00$, $p < 0.05$). Kết luận: Bảng câu hỏi cơ-xương-khớp Bắc Âu có thể được sử dụng với độ tin cậy chấp nhận được để đánh giá các rối loạn cơ-xương-khớp trên các chuyên viên Vật lý trị liệu Việt Nam.

Từ khóa: Bảng câu hỏi cơ-xương-khớp Bắc Âu, đau, độ tin cậy, khó chịu, rối loạn cơ xương

Received: 05/04/2023

Revised: 06/06/2023

Accepted for publication: 08/06/2023