CiSE
Clinics in Shoulder and Elbow

Reliability of the scapular dyskinesis test yes-no classification in asymptomatic individuals between students and expert physical therapists

Lawrence S. Ramiscal[1,#], Lori A. Bolgla[1], Chad E. Cook[2], John S. Magel[3], Stephen A. Parada[4], Raymond Chong[5]

[1]Department of Physical Therapy, Augusta University, Augusta, GA, USA
[2]Division of Physical Therapy, Duke University, Durham, NC, USA
[3]Department of Physical Therapy and Athletic Training, University of Utah, Salt Lake City, UT, USA
[4]Department of Orthopedic Surgery, Augusta University, Augusta, GA, USA
[5]Department of Interdisciplinary Health Sciences, Augusta University, Augusta, GA, USA

Background: Scapular dyskinesis is considered a risk factor for the shoulder pain that may warrant screening for prevention. Clinicians of all experience screen scapular dyskinesis using the scapular dyskinesis test yes-no classification (Y-N), yet its reliability in asymptomatic individuals is unknown. We aimed to establish Y-N's intra- and inter-reliability between students and expert physical therapists.
Methods: We utilized a cross-sectional design using consecutive asymptomatic subjects. Six students and two experts rated 100 subjects using the Y-N. Cohen's kappa (κ) and Krippendorff's alpha (K-α) were calculated to determine intra- and inter-rater reliability.
Results: Intra- and inter-rater values for experts were κ=0.92 (95% confidence interval [CI], 0.91–0.93) and 0.85 (95% CI, 0.84–0.87) respectively; students were κ=0.77 (95% CI, 0.75–0.78) and K-α=0.63 (95% CI, 0.58–0.67).
Conclusions: The Y-N is reliable in detecting scapular dyskinesis in asymptomatic individuals regardless of experience.

Keywords: Dyskinesias; Musculoskeletal system; Physical therapy specialty; Shoulder; Students

INTRODUCTION

Optimal shoulder function requires proper positioning and movement of the scapula on the thorax [1]. Abnormal scapular position or movement patterns during functional activities are defined as scapular dyskinesis [2,3]. Although it is typically associated with shoulder pain [4-6], dyskinesis also can be present in asymptomatic individuals [7-9]. More recent evidence suggests that scapular dyskinesis is a risk factor for shoulder pain [10] that may warrant screening as a preventative measure.

Physical therapists screen for scapular dyskinesis by visually comparing scapular movement asymmetries in overhead reach using the Scapular Dyskinesis Test [3]. The patient performs repeated shoulder elevation and lowering with weights on both

hands while the therapist observes scapular motion. The therapist identifies and labels scapular dyskinesis as type 1 when there is an excessive prominence of the inferior angle, as type 2 when there is excess prominence of the medial border or dysrhythmia, or as type 3 with excessive or premature movement of the scapula observed on a single plane of motion. The large numbers of possible abnormal movement patterns and combinations can make it difficult for therapists to agree on a final label. A variant of the test known as the Yes-No classification (Y-N) simply identifies the presence or absence of asymmetry between the shoulders and is more inclusive without need for the therapist to observe multiple separate planes, increasing the reliability [11]. The improved accuracy of the Y-N may be due to its simplicity and dichotomous decision [12]. Novice clinicians, such as physical therapy students, can quickly learn the Y-N as part of their training (e.g., clinical rotations). However, the Y-N involves subjectivity in that it relies heavily on clinician experience and is an observational method [13]. As novices, physical therapy students lack the experience needed for reliable and accurate measurement based on academic and clinical standards, especially in shoulder assessment tools [14,15]. Many studies have compared the reliability between novices and experienced clinicians using other assessment tools (primarily in balance) in physical therapy [16,17]. These studies also found evidence of rater discrepancy due to lack of experience. The Y-N has shown reliability among experienced clinicians [11,18,19]. However, its reliability across varied clinical experiences in the asymptomatic population is unknown.

Therefore, we aimed to determine the intra- and inter-rater reliability of the Y-N in detecting scapular dyskinesis in asymptomatic individuals between students and expert physical therapists. We hypothesized that the Y-N is a reliable tool in detecting scapular dyskinesis among asymptomatic individuals when used by experts but not by students due to lack of experience.

METHODS

Approval was obtained from the Institutional Review Board of Augusta University, and all subjects read and signed a consent form before participating in our study. Especially, the authors obtained consent from the participant whose body was exposed in the figure.

Study Design
A cross-sectional intra- and inter-rater reliability design was utilized.

Subjects
Participants were conveniently sampled from students on the Health Sciences campus of Augusta University. Asymptomatic adults 18–35 years old were recruited using word of mouth and referrals. Table 1 summarizes the exclusion criteria. A screening tool for eligibility included existing medical problems, medications, and pain ratings. The first consecutive 100 healthy asymptomatic subjects that met the criteria were included in the study and underwent evaluation via the Y-N (see Procedures and Instrumentation). Table 2 summarizes the demographic characteristics of the subjects.

Raters
There were eight raters: two experts and six students. The expert raters were licensed and certified orthopedic physical therapy specialists, one with 25 years of clinical experience, considered

Table 2. Demographic characteristics of the subjects

Variable	Value (n = 100)
Age (yr)	24 ± 3
Women	63 (63)
Handedness (right)	89 (89)
History of repeated overhead movement	71 (71)

Values are presented as mean ± standard deviation or number (%).

Table 1. Exclusion criteria

Any of the following
Shoulder pain with activity of 2/10 or greater on the numeric pain rating scale
History of shoulder pain within the past year
Adhesive capsulitis, defined as loss of greater than 50% of passive shoulder range of motion in shoulder external rotation and one other plane of motion
Previous shoulder surgery within the past year
History of shoulder fracture
Systemic musculoskeletal disease (rheumatoid arthritis, fibromyalgia, etc.)
Shoulder pain that was reproduced with active/passive cervical spine motion

the expert gold standard, and the other with 21 years of clinical experience. The student raters were second-year PT students. All raters were blinded to other's data during the study period. Table 3 summarizes the demographic characteristics of the raters.

Procedures and Instrumentation

Scapular dyskinesis test yes-no classification
The Y-N was performed on the 100 subjects and video recorded

Table 3. Demographic characteristics of reliability study raters

Variable	Expert (n = 2)	Student (n = 6)
Year of experience	23 ± 3	0
PT education	DPT	2nd year DPT
OCS	2 (100)	0

Values are presented as mean ± standard deviation or number (%).
PT: physical therapy, DPT: doctor of physical therapy, OCS: licensed and certified orthopedic physical therapy specialist.

for later evaluation of presence or absence of scapular dyskinesis (Fig. 1). Male participants were asked to remove their shirts, while women wore sports bras to expose both scapulae. Using a metronome at a rate of 60 beats per minute, participants performed five consecutive non-stop repetitions of bilateral, active, and weighted 120° shoulder flexion using dumbbells based on body weight: 1.4kg (3lb) for those weighing < 68.1 kg (150 lb) and 2.3 kg (5 lb) for those > 68.1 kg (150 lb) according to the scapular dyskinesis test protocol by McClure at al. [18].

An eight-foot PVC pipe on a wooden base was placed in front of the subjects (two feet from their toes) to standardize shoulder flexion and assure accuracy among the five repetitions. A spring clamp with handles wrapped with bright neon orange tape was clamped to the pole for easy visibility. Subjects' shoulders were passively elevated to align with a goniometer (fixed at 120°) and were held in that position. The clamp was moved roughly at the level of the subjects' middle fingers or a level they would remember to raise their arms during the test. To establish reliability be-

Fig. 1. Scapular dyskinesis test yes-no classification and video recording set-up.

tween repetitions, after determining the clamp's ideal height on the pole, subjects were asked to put their arms to their sides, raise them again to the clamp level, and hold. The fixed goniometer was placed at the shoulders one at a time to verify alignment. This process was repeated until elevation of both arms aligned with the goniometer.

To record the movement, a high-definition digital camera on a tripod equipped with lighting was set up one meter behind the participant at the level of the seventh thoracic spinous process (between the inferior angles of the scapulae). Each video was saved in an MP4 format and labeled with an unidentified subject number assigned during the consent process. All videos were stored in a secure Box folder (server) provided by the Institutional Review Board. After watching the videos independently, raters used the Y-N to label the presence or absence of scapular dyskinesis for each subject they evaluated.

Definitions of operational terms

Yes: Scapular dyskinesis is present (asymmetrical shoulders). Either or both of the following motion abnormalities may be present on either shoulder: (1) dysrhythmia: the scapula demonstrates premature or excessive elevation or protraction, nonsmooth or stuttering motion during arm elevation or lowering, or rapid downward rotation during arm lowering or (2) winging: the medial border or inferior angle of the scapula is posteriorly displaced from the posterior thorax.

No: Scapular dyskinesis is not present (symmetrical shoulders). Both scapulae are stable with minimal motion during the initial 30º to 60º of shoulder elevation. Smooth and continuous scapular rotation upward during elevation and downward during humeral lowering. No evidence of winging.

Student training

Students underwent a two-part standardized training provided by the expert gold standard (Fig. 2). The first part was a didactic format to educate the students on use of the Y-N. The second part was a practical application format where all student raters independently rated sample videos of subjects performing the Y-N to achieve a baseline minimum of substantial agreement (Krippendorff's alpha or K-α = 0.61–0.80) [20] before the study proper.

Rating process

After reaching the required baseline level of agreement (substantial) among the six student raters, the 100 study videos were released to all raters at a rate of 10 per week over the next 10 weeks
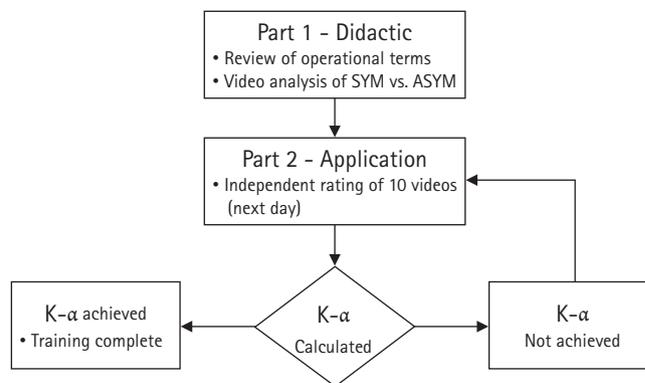

Fig. 2. Student rater training. SYM: symmetrical, ASYM: asymmetrical, K-α: Krippendorff's alpha.

for independent rating. The ratings in this part were used to calculate inter-rater reliability. Access to the videos was closed and the ratings were due at the end of the week. At the end of the 10th week, videos from the first week were re-released for the second round of ratings. Ratings in this part were used to calculate intra-rater reliability.

Sample Size Estimation

A priori power analysis using Real Statistics Resource Pack software, release 7.2, was used to establish reliability. Based on the previously determined inter-reliability Cohen's kappa (κ) value of 0.64 [21] with a significance level of 0.05 and power of 90%, the minimum sample size required to test the null hypothesis $κ = 0.3$ versus the alternative hypothesis $κ = 0.6$ was 72.

Statistical Methods

To determine the intra-rater reliability in student and expert raters, κ [22] and its 95% confidence interval (CI) for each rater were calculated between the first and second ratings of the videos from the first week (10 weeks apart) and then averaged. To determine the inter-rater reliability between student raters only, K-α [23] with its 95% CI was calculated. To determine the inter-rater reliability between expert raters only, the κ was calculated. Bootstrapping using the nonparametric (resampling) method, with a sample size of 1,000 that yielded 1,500 pairs, was performed to improve the accuracy of distribution of the alphas and Kappas [20,24]. Without bootstrapping, the CIs were wider (Table 4). The suggested interpretation of both K-α and κ is as follows: <0.0, poor agreement; 0.0–0.2, slight; 0.21–0.4, fair; 0.41–0.6, moderate; 0.61–0.8, substantial; and 0.81–1, near-perfect [22]. Statistical significance was set at $α = 0.05$. Statistical tests were performed with IBM SPSS ver. 27 (IBM Corp., Armonk, NY, USA).

Table 4. Summary of rater reliability

Variable	Intra-rater		Inter-rater	
Expert	κ	95% CI	κ	95% CI
	0.92	0.91–0.93	0.85	0.84–0.87
		0.85–0.99*		0.75–0.96*
Student	κ	95% CI	K-α	95% CI
	0.77	0.75–0.78	0.63	0.58–0.67
		0.59–0.95*		0.47–0.79*

κ: Cohen's kappa, CI: confidence interval, K-α: Krippendorff's alpha.
*CIs were calculated without bootstrapping.

RESULTS

Experts and students were reliable in using Y-N to detect scapular dyskinesis in asymptomatic individuals. Table 4 summarizes the reliability results of experts and students. The intra-rater reliability of the experts was near perfect (κ = 0.92), while that of students was substantial (κ = 0.77). The inter-rater reliability of the experts also was nearly perfect (κ = 0.85), and that of the students remained substantial (K-α = 0.63). The prevalence rate of scapular dyskinesis in our sample of 100 subjects as identified by the experts was 59%.

DISCUSSION

The results showed that the Y-N was reliable when used by students or experts in subjects without shoulder pain. Although student reliability was substantial, there was a 20-point difference from experts with near-perfect reliability. This was consistent with similar studies that investigated student reliability compared to that of experts using other clinical tests [16,17,25]. This finding was not surprising as experience may be the most obvious explanation for such a discrepancy. All authors of these studies concluded that experience was the most significant factor that explained the difference.

Our study found that reliability among students was consistently substantial when the Y-N was applied to asymptomatic subjects. This was consistent with the findings of a similar study by Møller, with student κ scores in the range of 0.70–0.90 [12]. Although their research also used PT students as raters, their reliability scores were higher than those of our study. This could be because they used PT students in their final year instead of PT students in their second year. This difference emphasizes the importance of clinical experience.

Our study found that expert reliability was consistently near perfect when the Y-N was applied to asymptomatic subjects. In a

previous study by Uhl et al. [11] utilized the Y-N for measuring reliability, the kappa score was only moderate between experts (κ = 0.41). Interestingly, the definition of "expert" in the Uhl et al.'s study [11] was limited to "experienced clinicians." In contrast, we defined experts as those board certified in orthopedic physical therapy and with at least two decades of clinical experience. This indicates that experience remains the most significant defining factor for higher reliability, even among experts. This was the same as the conclusion of Lluch et al. [26] in their comparison of inter-rater reliability among licensed physical therapists with different levels of experience.

Our study prevalence rate of scapular dyskinesis among asymptomatic individuals was 59%. It has been reported that about 60%–70% of individuals suffering shoulder pain have scapular dyskinesis [7-9]. However, many of those studies reported a similar proportion of patients with scapular dyskinesis even among healthy asymptomatic individuals reflective of our study's prevalence result.

The Y-N is very subjective, and there is possibility of an expectation bias because of an expected outcome. This may have influenced the scapular dyskinesis labeling because raters "see what they want to see;" in this case, the presence of scapular dyskinesis.

Most of the experiments took place during the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic [27]. The rating period stretched over 10 weeks at the pandemic height, which may have introduced history and timing biases from subject recruitment to rater performance.

Use of convenience sampling and its associated sampling bias may contribute to the weak generalizability of the results. It is possible that the sample was not representative of the general population due to the nature of volunteer subject enrollment and its associated response bias.

In conclusion, the Y-N is reliable in detecting scapular dyskinesis regardless of experience level when used in an asymptomatic population for screening.

ORCID

Lawrence S. Ramiscal   https://orcid.org/0000-0002-0230-277X
Lori A. Bolgla   https://orcid.org/0000-0002-4298-4779
Chad E. Cook   https://orcid.org/0000-0001-8622-8361
John S. Magel   https://orcid.org/0000-0001-9087-4753
Stephen A. Parada   https://orcid.org/0000-0002-6304-3038
Raymond Chong   https://orcid.org/0000-0002-2706-1300

REFERENCES

1. Neumann DA. Kinesiology of the musculoskeletal system: foundations for rehabilitation. St. Louis, MO: Mosby; 2016.

2. Burkhart SS, Morgan CD, Kibler WB. The disabled throwing shoulder: spectrum of pathology Part III: the SICK scapula, scapular dyskinesis, the kinetic chain, and rehabilitation. Arthroscopy 2003;19:641-61.

3. Kibler WB, Ludewig PM, McClure PW, Michener LA, Bak K, Sciascia AD. Clinical implications of scapular dyskinesis in shoulder injury: the 2013 consensus statement from the 'Scapular Summit'. Br J Sports Med 2013;47:877-85.

4. Roche SJ, Funk L, Sciascia A, Kibler WB. Scapular dyskinesis: the surgeon's perspective. Shoulder Elbow 2015;7:289-97.

5. Timmons MK, Thigpen CA, Seitz AL, Karduna AR, Arnold BL, Michener LA. Scapular kinematics and subacromial-impingement syndrome: a meta-analysis. J Sport Rehabil 2012;21:354-70.

6. Struyf F, Cagnie B, Cools A, et al. Scapulothoracic muscle activity and recruitment timing in patients with shoulder impingement symptoms and glenohumeral instability. J Electromyogr Kinesiol 2014;24:277-84.

7. Burn MB, McCulloch PC, Lintner DM, Liberman SR, Harris JD. Prevalence of scapular dyskinesis in overhead and nonoverhead athletes: a systematic review. Orthop J Sports Med 2016;4:2325967115627608.

8. Plummer HA, Sum JC, Pozzi F, Varghese R, Michener LA. Observational scapular dyskinesis: known-groups validity in patients with and without shoulder pain. J Orthop Sports Phys Ther 2017;47:530-7.

9. Ramiscal L, Bolgla L, Chong R. Scapular muscle activity and pectoralis minor muscle length of asymptomatic scapular dyskinesis: a pilot study [abstract]. J Orthop Sports Phys Ther 2019;50:CSM81-177. Abstract no. OPO184.

10. Hickey D, Solvig V, Cavalheri V, Harrold M, Mckenna L. Scapular dyskinesis increases the risk of future shoulder pain by 43% in asymptomatic athletes: a systematic review and meta-analysis. Br J Sports Med 2018;52:102-10.

11. Uhl TL, Kibler WB, Gecewich B, Tripp BL. Evaluation of clinical assessment methods for scapular dyskinesis. Arthroscopy 2009;25:1240-8.

12. Møller M, Attermann J, Myklebust G, et al. The inter- and intra-rater reliability and agreement for field-based assessment of scapular control, shoulder range of motion, and shoulder isometric strength in elite adolescent athletes. Phys Ther Sport 2018;32:212-20.

13. Lange T, Struyf F, Schmitt J, Lützner J, Kopkow C. The reliability of physical examination tests for the clinical assessment of scapular dyskinesis in subjects with shoulder complaints: a systematic review. Phys Ther Sport 2017;26:64-89.

14. Christiansen DH, Møller AD, Vestergaard JM, Mose S, Maribo T. The scapular dyskinesis test: reliability, agreement, and predictive value in patients with subacromial impingement syndrome. J Hand Ther 2017;30:208-13.

15. Rajasekar S, Bangera RK, Sekaran P. Inter-rater and intra-rater reliability of a movement control test in shoulder. J Bodyw Mov Ther 2017;21:739-42.

16. Maqueda CE, Patel R. Novice versus experienced rater reliability of the Mini Balance Evaluation Systems Test (Mini-BESTest) in patients with acquired brain injury (ABI). J Stud Phys Ther Res 2015;8:92-109.

17. Gulgin H, Hoogenboom B. The functional movement screening (fms)™: an inter-rater reliability study between raters of varied experience. Int J Sports Phys Ther 2014;9:14-20.

18. McClure P, Tate AR, Kareha S, Irwin D, Zlupko E. A clinical method for identifying scapular dyskinesis, part 1: reliability. J Athl Train 2009;44:160-4.

19. Rossi DM, Pedroni CR, Martins J, de Oliveira AS. Intrarater and interrater reliability of three classifications for scapular dyskinesis in athletes. PLoS One 2017;12:e0181518.

20. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Commun Methods Mea 2007;1:77-89.

21. Ramiscal L. Reliability of the scapular dyskinesis test between student and expert physical therapists in asymptomatic scapular dyskinesis: a pilot study [abstract]. J Orthop Sports Phys Ther 2020;51:CSM57-165. Abstract no. CSM166.

22. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. London: Routledge; 2013.

23. Krippendorff K. Computing Krippendorff's alpha-reliability [Internet]. Philadelphia, PA: University of Pennsylvania, Annenberg School of Communication; 2011 [cited 2022 Nov 7]. Available from: https://repository.upenn.edu/asc_papers/43.

24. Krippendorff K. Bootstrapping distributions for Krippendorff's alpha [Internet]. Philadelphia, PA: University of Pennsylvania, Annenberg School of Communication; 2016 [cited 2022 Nov 7]. Available from: https://www.asc.upenn.edu/sites/default/files/2021-03/Algorithm%20for%20Bootstrapping%20a%20Distribution%20of%20Alpha.pdf.

25. Kuo KT, Hunter BC, Obayashi M, et al. Novice vs expert inter-rater reliability of the balance error scoring system in children between the ages of 5 and 14. Gait Posture 2021;86:13-6.

26. Lluch E, Benítez J, Dueñas L, et al. The shoulder medial rotation test: an intertester and intratester reliability study in overhead

athletes with chronic shoulder pain. J Manipulative Physiol Ther 2014;37:198-205.

27. World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that causes it [Internet]. Geneva: World Health Organization; 2020 [cited 2021 Mar 15]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it#:~:text=Official%20names%20have%20been%20announced,%2DCoV%2D2).