Validity and Reliability of the Thai Version of the Thyroid-Related Patient-Reported Outcome—A Thyroid-specific Quality of Life Questionnaire

Patorn Piromchai[1]   Supachat Chaiudomsom[1]   Pattaramon Wijakkanalan[1]   Torquil Watt[2]

[1] Department of Otorhinolaryngology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand
[2] Department of Medical Endocrinology, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark

Address for correspondence Patorn Piromchai, MD, MSc, PhD, FRCOT, FICS, Department of Otorhinolaryngology, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand (e-mail: patorn@gmail.com).

Abstract

Introduction The Thyroid-Related Patient-Reported Outcome (ThyPRO) is a new thyroid-specific quality of life patient-reported outcome measure for benign thyroid disorders.

Objective The objective of this study was to investigate the face validity, internal consistency, and test-retest reliability of the Thai version of the ThyPRO (ThyPROth).

Methods The translation of the ThyPRO questionnaire was performed using double forward translation, reconciliation, single backward translation, and cognitive debriefing, followed by a panel review. Five thyroid patients evaluated the face validity. The internal consistency and test-retest reliability were evaluated in 30 patients with thyroid diseases.

Results The overall validity score was 3.75 (range 0–4). The Cronbach α coefficient ranged from 0.76 to 0.95, with a total coefficient of 0.97 (95% CI 0.962–0.959), indicating excellent internal consistency. The test-retest reliability coefficient ranged from 0.70 to 0.97. All values were 0.70 and above. The total reliability coefficient was 0.86 (95% CI 0.724–0.932), indicating excellent reliability.

Conclusion The ThyPROth was found to be valid and to exhibit good internal consistency and test-retest reliability. The questionnaire is ready for implementation in the assessment of health-related quality of life in Thai patients with benign thyroid diseases.

Keywords
► thyroid
► quality of life
► questionnaire
► hyperthyroidism
► hypothyroidism
► thyroid neoplasms

Introduction with Objective

Thyroid disease is a common problem that physicians often encounter in general practice. The disease can be categorized according to the thyroid hormone level as either hyperthyroidism, euthyroid goiter, or hypothyroidism and also according to thyroid characteristics, such as either diffuse, single, or multi-nodular goiter.

In Thailand, the incidences of hypothyroidism and hyperthyroidism are 4.16 percent and 3.18 percent of the total population, respectively.[1] In iodine-sufficient countries, the prevalence of hyperthyroidism has been shown to range from 0.2 to 1.3%,[2,3] and that of hypothyroidism ranged from 1 to 2%.[4,5]

Thyroid tumors, such as thyroid nodules, have been found in 2 to 6% of patients via palpation; 19 to 35% with ultrasound; and 8 to 65% through autopsy.[6] The risk factors for

malignancy include increased age, female sex, exposure to radiation, and lack of iodine. The incidence of thyroid cancer has grown over the last two decades. The mean annual incidence of thyroid malignancies in Thailand is 1.5 and 5.1 per 100,000 of the male and female populations, respectively.[7] In some cases, either open or endoscopic thyroid surgery is necessary.[8–10]

Patients with thyroid diseases are affected by the fluctuations in thyroid hormone levels, resulting in abnormal symptoms (e.g., dry skin, hair loss, irregular menstruation, and bradycardia), mental disorders (e.g., depression, loss of memory, and insomnia), and decreased quality of life. According to a previous study conducted in Thailand, 22.1% of patients with a major depressive disorder had abnormal levels of thyroid hormones.[11]

Treatments for thyroid disease include medication, radioiodine treatment, and surgery. After treatment, these patients require life-long follow-up. Like other chronic diseases, the ultimate goal of treatment is a good quality of life, which is a multidimensional variable that measures a patient's physical, emotional, social, and mental status.

Questionnaires are commonly-used tools to assess a patient's quality of life. There are two types of quality-of-life questionnaire: generic and disease-specific. In Thailand, the 36-Item Short Form Health Survey (SF-36) has been retranslated using a forward-backward method and tested for construct validity and internal consistency reliability.[12] In terms of thyroid-specific quality of life questionnaires, one study developed and tested the validity and reliability of a Thai version of the Quality of life-Thyroid Version (QOL-TV).[13] However, this questionnaire is applicable only to patients with thyroid cancer. To our knowledge, there have been no Thai versions of questionnaires aimed at patients with other types of thyroid disease.

The Thyroid-Related Patient-Reported Outcome (ThyPRO) is a new thyroid-specific quality of life patient-reported outcome measure for benign thyroid disorders.[14,15] A recent systematic review of the quality of thyroid-specific health-related quality-of-life instruments recommended using ThyPRO for patients with benign thyroid diseases.[16] The objective of this study was to investigate face validity, internal consistency, and test-retest reliability of the Thai version of the ThyPRO (ThyPROth).

Method

Characteristics of the ThyPRO Questionnaire

The ThyPRO questionnaire consists of 85 items to evaluate various aspects of quality of life in patients with thyroid disease, such as goiter symptoms, hyperthyroid symptoms, hypothyroid symptoms, eye symptoms, tiredness, cognitive impairment, anxiety, depressivity, emotional susceptibility, impaired social life, impaired daily life, impaired sex life, cosmetic complaints, and overall quality of life. The questionnaire requires around 15 to 20 minutes to complete. Each item is rated using a 5-point Likert scale (0–not at all, 1–a little, 2–some, 3–quite a bit, and 4–very much/completely),[17] and the final scores are translated into a scale from 0 to 100. Higher scores indicate more symptoms and worse quality of life.

Translation Validity

Approval was sought from the questionnaire's developer. This study used the standard method for translation of patient-reported outcomes.[18,19] The questionnaire was translated from the original English to Thai by two native Thai professional translators from Khon Kaen University's Department of Humanities, who were fluent in both Thai and English and familiar with both Thai and Western culture.

The two translators were independent and blinded to each other. After the translations were complete, any inconsistencies were solved by a mutual agreement. We called this version of the questionnaire the "Reconciled Forward Translation of the ThyPROth."

The reconciled forward translation questionnaire was then translated back to English by an American professional translator (D. S.) who is fluent in both Thai and English. The backward translation version was compared with the original version of the questionnaire by 1. a native English speaker and professional translator (D. F.), 2. the developer of the ThyPRO (T. W.), 3. thyroid specialists (P. P., P. W.), and 4. a language coordinator (S. C.). Any issue that arose during the process was documented, and if revisions were made, the forward-backward translation was repeated, until a final version was developed.

The prefinal version of the ThyPROth was tested in 5 patients with thyroid diseases using a cognitive interview technique. The questions were as follows:

1. Did you have any difficulty understanding this item?
2. What does this item mean to you?
3. Would you rephrase this item?
4. Is this item relevant to your situation?
5. If the response options are not clear and consistent with this item, how would you rephrase them?

The cognitive debriefing results were then reviewed, and the translation was finalized.

Face Validity

The 5 patients mentioned above were also asked to assess the validity of the ThyPROth questionnaire by rating each of the following items on a 5-point Likert scale (0–not at all, 1–a little, 2–some, 3–quite a bit, and 4–very much/completely):

1. The questions were generally clear, easy to understand, and easy to answer.
2. Overall, the assessment of the symptoms matched your situation.
3. The instructions on the questionnaire were clear and easy to understand.
4. This questionnaire is practicable.

Internal Consistency

The internal consistency was measured based on the correlations among various items in the following 14 areas (85 items):

1. Goiter symptoms (11 items)
2. Hyperthyroid symptoms (8 items)
3. Hypothyroid symptoms (4 items)
4. Eye symptoms (8 items)

5. Tiredness (7 items)
6. Cognitive impairment (6 items)
7. Anxiety (6 items)
8. Depressivity (7 items)
9. Emotional susceptibility (9 items)
10. Impaired social life (4 items)
11. Impaired daily life (6 items)
12. Impaired sex life (2 items)
13. Cosmetic complaints (6 items)
14. Overall quality of life (1 item)

Test-retest Reliability

Thirty patients with thyroid disorders were enrolled to determine whether the results of the ThyPROth were reproducible over time, provided that all other variables remained the same. All patients completed the baseline ThyPROth questionnaire and were retested with the same questionnaire 2 weeks after their first response.

Participants and Setting

This study was conducted at Khon Kaen University's Srinagarind Hospital, in Thailand, from September 2017 to March 2019. Patients over 18 years old with thyroid diseases (including hyperthyroidism, hypothyroidism, and thyroid nodule) were recruited. We excluded patients who had cancer or other chronic diseases (such as heart disease, diabetes mellitus, or kidney disease), those who were pregnant, and those who had cognitive impairment or psychological disorders. All participants completed the questionnaire by themselves.

Sample Size

During the translation process, five patients were recruited according to the standard methodology of original ThyPRO questionnaire.[14,15] To evaluate internal consistency and test-retest reliability, 30 patients were recruited based on the acceptable sample size for the application of Cronbach α coefficient.[20]

Statistical Analysis

The statistical analysis was performed using Stata version 15 (StataCorp LLC, College Station, TX, USA). For patient characteristics and outcomes, continuous variables were presented as mean ± standard error (SE), and categorized variables were presented as frequencies or percentages. The internal consistency was evaluated using the Cronbach α test.[21] Test-retest reliability was evaluated using the intraclass correlation coefficient (ICC) between the two measurements (two-way mixed effect models).[22,23] In the case of non-normality of data, the Spearman rank-order correlation coefficient was used. A 95% confidence interval (95%CI) was calculated for all tests.

The Cronbach α coefficient can be interpreted as followed: excellent ($0.9 \leq \alpha$), good ($0.8 \leq \alpha < 0.9$), acceptable ($0.7 \leq \alpha < 0.8$), questionable ($0.6 \leq \alpha < 0.7$), poor ($0.5 \leq \alpha < 0.6$), and unacceptable ($\alpha < 0.5$).[24]

The ICC can be interpreted as followed: poor ($ICC < 0.5$), moderate ($0.5 \leq ICC < 0.75$), good ($0.75 \leq ICC < 0.9$), excellent ($ICC \geq 0.9$).[25]

The Spearman rank-order correlation coefficient assesses the relationship between the two variables. It has a value between +1 and −1, in which 1 is total positive correlation, 0 is no linear correlation, and −1 is total negative correlation.[26]

Ethical Considerations

This study was approved by the Khon Kaen University Ethics Committee in Human Research (HE601276). Written informed consent to participate was provided by all patients enrolled.

Results

Translation of the ThyPRO

In the forward translation step, the word/phrasing choices of the two independent translators were slightly different, but the meaning of the content was the same. There were 8 items upon which the translators had to reach an agreement.

The backward translation of ThyPROth was compared with the original questionnaire. There were 8 items about which issues were raised by the native English-speaking professional translator and developer of the ThyPRO. A consensus among the translator, developer of the ThyPRO, thyroid specialists, and language coordinator was eventually reached. Forward and backward translations were repeated until a final version was completed.

Two issues were raised during the cognitive interviews with the five patients mentioned above. The authors decided to change the wording of these two items to be more consistent with the local language and culture. However, we did so in a way that ensured these items still corresponded to the variables to be measured and that their original content and meaning remained unchanged.

Face Validity

All five patients who were enrolled to participate during the translation process were female (100%). The mean age of the patients was 39.5 years. Three had nontoxic goiters (60%), 1 had Graves' disease, and 1 had hypothyroidism. All participants condition was euthyroid when administered the questionnaire.

The overall validity score was 3.75 (range 0–4). The 1st item ("The questions were generally clear, easy to understand, and easy to answer") received a mean score of 3.6; the 2nd item ("Overall, the assessment of the symptoms matched your situation") received a mean score of 3.6; the 3rd item ("The instructions on the questionnaire were clear and easy to understand") received a mean score of 3.8, and the 4th item ("This questionnaire is practicable") received a mean score of 4.

Internal Consistency

Thirty patients were enrolled in the internal consistency and test-retest reliability study, most of whom were female (80%). The most common diagnosis was nontoxic goiter (50%). The average amount of time taken to complete the questionnaire was 15.36 minutes. (►**Table 1**)

Table 1 Patient characteristics

Characteristics	N (%)
Sex	
• Male	6 (20%)
• Female	24 (80%)
Age (mean, years)	48.5 ± 2.36
Diagnosis	
• Nontoxic goiter	15 (50%)
• Toxic nodular goiter	6 (20%)
• Graves' disease	5 (16.67%)
• Hypothyroid	4 (13.33%)
Time used to complete the questionnaire (mean, min)	15.36 ± 0.24

Table 2 Internal consistency

Areas measured	Cronbach α coefficients (95% CI)
1. Goiter symptoms (11 items)	0.76 (0.641–0.882)
2. Hyperthyroid symptoms (8 items)	0.92 (0.839–0.999)
3. Hypothyroid symptoms (4 items)	0.78 (0.615–0.953)
4. Eye symptoms (8 items)	0.89 (0.692–0.999)
5. Tiredness (7 items)	0.84 (0.745–0.929)
6. Cognitive impairment (6 items)	0.95 (0.896–0.996)
7. Anxiety (6 items)	0.91 (0.861–0.968)
8. Depressivity (7 items)	0.89 (0.758–0.999)
9. Emotional susceptibility (9 items)	0.87 (0.771–0.975)
10. Impaired social life (4 items)	0.83 (0.671–0.987)
11. Impaired daily life (6 items)	0.84 (0.736–0.953)
12. Impaired sex life (2 items)	0.93 (0.853–0.999)
13. Cosmetic complaints (6 items)	0.88 (0.673–0.999)
14. Overall quality of life (1 item)	–
Total internal consistency (85 items)	0.97 (0.962–0.959)

Table 3 Total internal consistency

Areas measured	Total Cronbach α coefficients (95% CI)
1. Without goiter symptoms (11 items)	0.97 (0.957–0.985)
2. Without hyperthyroid symptoms (8 items)	0.97 (0.950–0.983)
3. Without hypothyroid symptoms (4 items)	0.97 (0.955–0.984)
4. Without eye symptoms (8 items)	0.97 (0.953–0.984)
5. Without tiredness (7 items)	0.97 (0.957–0.985)
6. Without cognitive impairment (6 items)	0.97 (0.951–0.983)
7. Without anxiety (6 items)	0.97 (0.952–0.983)
8. Without depressivity (7 items)	0.97 (0.957–0.985)
9. Without emotional susceptibility (9 items)	0.97 (0.956–0.985)
10. Without impaired social life (4 items)	0.97 (0.956–0.985)
11. Without impaired daily life (6 items)	0.97 (0.954–0.984)
12. Without impaired sex life (2 items)	0.97 (0.957–0.985)
13. Without cosmetic complaints (6 items)	0.97 (0.954–0.984)
14. Without overall quality of life (1 item)	0.97 (0.957–0.985)

The Cronbach α for the total score without one domain at a time was calculated. The total coefficient without domain 1 to 14 was 0.97 for all analyses, indicating the uniformity of total internal consistency across the questionnaire. (►**Table 3**)

Test-retest Reliability

The Spearman rank-order correlation coefficient ranged from 0.70 to 0.97. Most of the areas measured were 0.70 and above, indicating excellent reliability.[28] The total coefficient was 0.86 (95% CI 0.724–0.932). (►**Table 4**).

To evaluate the floor and ceiling effect, the percentage frequency of lowest or highest possible score achieved by respondents was assessed. In the first test, 0% of the patients achieved the best score (100) and 0% achieved the worst score (0). In the second test, the results remained the same.

Discussion

As there was no disease-specific questionnaire for benign thyroid diseases available in Thai, there was a need to translate a standard questionnaire that is used for this purpose. We decided to translate the ThyPRO questionnaire because a systematic review of the quality of thyroid-specific

The internal consistency was calculated for each of the measured areas and summarized in ►**Table 2**. The Cronbach α coefficients ranged from 0.76 to 0.95, with a total coefficient of 0.97 (95% CI 0.962–0.959), indicating excellent internal consistency.[27] (►**Table 2**)

Table 4 Test-retest reliability

Areas measured	Spearman rank-order correlation coefficient (95% CI)
1. Goiter symptoms (11 items)	0.76 (0.550–0.879)
2. Hyperthyroid symptoms (8 items)	0.75 (0.534–0.874)
3. Hypothyroid symptoms (4 items)	0.78 (0.584–0.890)
4. Eye symptoms (8 items)	0.73 (0.502–0.863)
5. Tiredness (7 items)	0.81 (0.635–0.906)
6. Cognitive impairment (6 items)	0.84 (0.688–0.921)
7. Anxiety (6 items)	0.90 (0.799–0.952)
8. Depressivity (7 items)	0.70 (0.454–0.847)
9. Emotional susceptibility (9 items)	0.73 (0.502–0.863)
10. Impaired social life (4 items)	0.93 (0.857–0.966)
11. Impaired daily life (6 items)	0.86 (0.724–0.932)
12. Impaired sex life (2 items)	0.97 (0.937–0.986)
13. Cosmetic complaints (6 items)	0.89 (0.780–0.947)
14. Overall quality of life (1 item)	0.85 (0.706–0.927)
Total test-retest reliability (85 items)	0.86 (0.724–0.932)

health-related quality-of-life instruments recommended using it to evaluate patients with benign thyroid diseases.[16] The purpose of this study was to translate the ThyPRO questionnaire into Thai and to establish the face validity, internal consistency, and test-retest reliability of the translated version.

During the translation process, we found some differences in the grammatical structures of the English and Thai version, such as the use of the temporal expressions and reliance on context.[29]

After testing the final version of ThyPROth in five patients, they rated the overall validity of the test at 3.75 (range 0–4). All patients rated the practicability of this questionnaire as excellent.

Internal consistency was evaluated using the standard Cronbach α approach. A previous study conducted by Watt et al in 904 patients had found the Cronbach α coefficients of this questionnaire to be acceptable (> 0.70).[15] In our study, we found the Cronbach α coefficients ranged from 0.76 to 0.95, with a total coefficient of 0.97 (95% CI 0.962–0.959). This finding was comparable with other reliability studies. One study using a Romanian version of the ThyPRO[30] found

that the Cronbach α coefficients ranged from 0.78 to 0.93, and another using a Serbian version found that they ranged from 0.83 to 0.95.[31]

There was no standard cut-point on the acceptable value of Cronbach α coefficients. However, the acceptable values in most reports were ranging from 0.70 to 0.95.[32,33] A low value of α could be due to a low number of questions, poor inter-relatedness between items or heterogeneous constructs. If α is too high, it may suggest that some items are redundant, as they are testing the same question but in a different guise.[34]

In our study, the Cronbach α coefficient in each area ranged from 0.76 to 0.95, which is within the acceptable range. However, the total internal consistency was 0.97, which is slightly above the acceptable range. This can be explained by the length of the test, as an 85-item questionnaire is slightly lengthy and can increase the total α coefficient.[35]

The test-retest reliability was analyzed using Spearman rank-order correlation coefficient. Most of the areas analyzed were above 0.70. The total coefficient was 0.86 (0.724–0.932), indicating excellent reliability. This finding was comparable with the original study by Watt et al, in which the ICCs ranged from 0.77 to 0.89.[14]

The limitation of this study was a small sample size. However, the power calculation was performed to assess whether the sample size of 30 was enough to reject the null hypothesis when the alternative hypothesis is true. The ThyPRO questionnaire has 85 items, we found the Cronbach α coefficients ranged from 0.76 to 0.95. Based on an α value fixed at 0.05, the power of this study was higher than 90%.[36]

14 Furthermore, it can be used to guide the surgeon to make the surgical decision on benign thyroid tumor. If the benign tumor was severely affecting the quality of life, the surgeon may advise the patient to undergo thyroid surgery. (►**Appendix 1 - Available online**)We found that the Thy-PROth had adequate internal consistency and test-retest reliability. We are currently planning to conduct a study that focuses on the construct validity of the ThyPROth by comparing the responses of benign thyroid disease patients with those of a normal population.

Conclusion

The ThyPROth was found to be valid and to exhibit good internal consistency and test-retest reliability. The questionnaire is ready to be implemented in the assessment of health-related quality of life in Thai patients with benign thyroid diseases.

References

1 Sriphrapradang C, Pavarangkoon S, Jongjaroenprasert W, Chailurkit LO, Ongphiphadhanakul B, Aekplakorn W. Reference ranges of serum TSH, FT4 and thyroid autoantibodies in the Thai population: the national health examination survey. Clin Endocrinol (Oxf) 2014;80(05):751–756. Doi: 10.1111/cen.12371

2 Garmendia Madariaga A, Santos Palacios S, Guillén-Grima F, Galofré JC. The incidence and prevalence of thyroid dysfunction in Europe: a meta-analysis. J Clin Endocrinol Metab 2014;99(03): 923–931. Doi: 10.1210/jc.2013-2409

3 Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). J Clin Endocrinol Metab 2002;87(02):489–499. Doi: 10.1210/jcem.87.2.8182

4 Vanderpump MP. The epidemiology of thyroid disease. Br Med Bull 2011;99:39–51. Doi: 10.1093/bmb/ldr030

5 Parle JV, Franklyn JA, Cross KW, Jones SC, Sheppard MC. Prevalence and follow-up of abnormal thyrotrophin (TSH) concentrations in the elderly in the United Kingdom. Clin Endocrinol (Oxf) 1991;34(01):77–83

6 Dean DS, Gharib H. Epidemiology of thyroid nodules. Best Pract Res Clin Endocrinol Metab 2008;22(06):901–911

7 Tangjaturonrasme N, Vatanasapt P, Bychkov A. Epidemiology of head and neck cancer in Thailand. Asia Pac J Clin Oncol 2018;14 (01):16–22. Doi: 10.1111/ajco.12757

8 Piromchai P, Wijakkanalan P, Teeramatwanich W, Kasemsiri P, Laohasiriwong S, Ratanaanekchai T. Postauricular-submental approach endoscopic thyroidectomy. Clin Otolaryngol 2018;43(02): 767–769. Doi: 10.1111/coa.12791

9 Piromchai P, Vatanasapt P, Reechaipichitkul W, Phuttharak W, Thanaviratananich S. Is the routine pressure dressing after thyroidectomy necessary? A prospective randomized controlled study. BMC Ear Nose Throat Disord 2008;8:1. Doi: 10.1186/1472-6815-8-1

10 Piromchai P. Early Outcomes of the Postauricular-Submental Approach Endoscopic Thyroidectomy. J Med Assoc Thai 2017;100:202

11 Suttajit S, Pilakanta S. Prevalence of and predictors for depression in individuals with schizophrenia. Eur Neuropsychopharmacol 2016;2:S497–S498

12 Leurmarnkul W, Meetam P. Properties testing of the retranslated SF-36 (Thai version). Thaiphesatchasan 2005;29:69–90

13 Vasupongaiya J, Lerkiatbundit S, Premprabha T. Testing of the Validity and Reliability of the Quality of Life-Thyroid Version (Thai). Thai J Pharm Pract 2013;5:43–62

14 Watt T, Hegedüs L, Groenvold M, et al. Validity and reliability of the novel thyroid-specific quality of life questionnaire, ThyPRO. Eur J Endocrinol 2010;162(01):161–167. Doi: 10.1530/EJE-09-0521

15 Watt T, Bjorner JB, Groenvold M, et al. Establishing construct validity for the thyroid-specific patient reported outcome measure (ThyPRO): an initial examination. Qual Life Res 2009;18(04): 483–496. Doi: 10.1007/s11136-009-9460-8

16 Wong CK, Lang BH, Lam CL. A systematic review of quality of thyroid-specific health-related quality-of-life instruments recommends ThyPRO for patients with benign thyroid diseases. J Clin Epidemiol 2016;78:63–72. Doi: 10.1016/j.jclinepi.2016.03.006

17 Albaum G. The Likert scale revisited. Market Res Soc J 1997; 39:1–21

18 Wild D, Eremenco S, Mear I, et al. Multinational trials-recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report. Value Health 2009;12(04):430–440

19 Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res 2013;22(08):1889–1905

20 Yurdugül H. Minimum sample size for Cronbach's coefficient alpha: a Monte-Carlo study. Hacettepe Üniv Eğitim Fakültesi Dergisi 2008;35:1–9

21 Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297–334

22 Pearson K. VII. Note on regression and inheritance in the case of two parents. Proc R Soc Lond 1895;58:240–242

23 Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 2005;19 (01):231–240

24 Loewenthal K, Lewis CA. An introduction to psychological tests and scales. Psychology press; 2018

25 Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016;15(02):155–163. Doi: 10.1016/j.jcm.2016.02.012

26 Lehman A. JMP for basic univariate and multivariate statistics: a step-by-step guide. SAS Institute; 2005

27 Mallery P, George D. SPSS for Windows step by step: a simple guide and reference. Allyn, BaconBoston2003

28 Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284

29 Thammawan S. A study of the differences of some grammatical structures of English and Thai that cause English speaking problems for Thai people. Language Institute, Thammasat University; 2008

30 Zahan AE, Watt T, Pascanu I, et al. The romanian version of the thyroid-related patient-reported outcomes thypro and thypro-39. translation and assessment of reliability and cross-cultural validity. Acta Endocrinol (Bucur) 2018;14(02):192–200

31 Bukvić B, Živaljević V, Šipetić S, Diklić A, Taušanović K, Paunović I. Validation and cross-cultural adaptation of the questionnaire ThyPRO in thyroid patients in Serbia. Vojnosanit Pregl 2015;72 (07):583–588

32 Bland JM, Altman DG. Cronbach's alpha. BMJ 1997;314 (7080):572. Doi: 10.1136/bmj.314.7080.572

33 DeVellis RF. Scale development: Theory and applications. Sage publications; 2016

34 Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011;2:53–55. Doi: 10.5116/ijme.4dfb.8dfd

35 Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J Pers Assess 2003;80(01): 99–103. Doi: 10.1207/S15327752JPA8001_18

36 Bujang MA, Omar ED, Baharum NA. A Review on Sample Size Determination for Cronbach's Alpha Test: A Simple Guide for Researchers. Malays J Med Sci 2018;25(06):85–99. Doi: 10.21315/mjms2018.25.6.9