VALIDITY AND RELIABILITY OF THE SHSQ-25 FOR ASSESSING SUBOPTIMAL HEALTH STATUS AMONG HANOI MEDICAL UNIVERSITY STUDENTS

Nguyen Thi Nguyen[1], Le Thi Thanh Truc[2], Pham Kim Ngoc[2], Nguyen Thi Anh Dong[2], Dang Quang Thanh[2], Luu Ngoc Minh[2], Phan Thanh Hai[2], Dinh Thai Son[2], Do Thi Thanh Toan[2*]

[1]University of Science and Technology of Hanoi

[2]Hanoi Medical University

Received 08 August 2024

Accepted 30 August 2024

Abstract: The Suboptimal Health Status Questionnaire (SHSQ-25) has been validated in different languages, such as Chinese and other Asian countries. This study, conducted in 2023, aims to examine the validity and reliability of the Vietnamese version of SHSQ-25 using data from 353 undergraduate students at Hanoi Medical University. Reliability was evaluated through test-retest reliability using the Intraclass Correlation Coefficient (ICC) and internal consistency using Cronbach's alpha. The subscales for fatigue, mental status, cardiovascular, and digestive symptoms demonstrated good test-retest reliability, with ICC values ranging from 0.416 to 0.703, and high internal consistency, with Cronbach's *alpha* ranging from 0.81 to 0.90. However, the immune system subscale exhibited lower reliability, with an ICC of 0.276 and a Cronbach's *alpha* of 0.53, indicating less consistency between measurements. Validity was assessed through Exploratory Factor Analysis (EFA), which was performed to determine the underlying structure of the Vietnamese version, and Confirmatory Factor Analysis (CFA), which was used to further assess its fit to the data. EFA results revealed a new 5-factor structure. CFA supported the 5-factor model, but findings indicated that further refinement of the SHSQ-25 may be warranted to optimize its factor structure and item composition.

Keywords: Suboptimal health status, questionnaire, reliability, validity.

1. Introduction

In 1946, the World Health Organization (WHO) defined health as a state of physical, mental, and social well-being

* Corresponding author
E-mail address: dothithanhtoan@hmu.edu.vn
https://doi.org/10.56086/jcvb.v4i3.176

combined, not just the absence of illness or disability [1]. This represents a significant shift from the traditional disease-focused perspective. This holistic conceptualization recognizes that health is not simply the absence of overt pathology but a multidimensional overall well-being across various domains [2].

Suboptimal health status refers to a condition that falls between full health and disease. Early recognition and management of suboptimal health symptoms are vital for preventing further development into chronic conditions and promoting overall well-being [3]. However, assessing suboptimal health remains challenging due to its multi-dimensional nature, encompassing physical and psychological domains [4,5]. Many existing measures focus on specific symptom clusters, failing to capture the interrelationships between various aspects of suboptimal health [6].

Developed and validated by researchers in China, the Suboptimal Health Status Questionnaire (SHSQ-25) provides a multidimensional evaluation of suboptimal health status (SHS), which refers to the presence of non-specific medical symptoms that do not meet the diagnostic criteria for overt disease, but indicate a state of elevated risk for the development of chronic illnesses. This scale has been validated for reliability and validity across various populations and geographical regions, demonstrating its wide applicability in evaluating suboptimal health status in research and clinical practice [6, 7].

Previous studies have validated and assessed for reliability the SHSQ-25's psychometric properties in general populations [6, 8, 9]. However, its performance in student populations, particularly those pursuing medical education, remains underexplored. In Vietnam, a recent study found that 13.8% of medical students at the University of Medicine and Pharmacy, Hue University, experienced suboptimal health status [10]. Medical students face unique challenges and risk factors that make them particularly susceptible to suboptimal health status. They must endure intensive study loads, demanding academic schedules, and high-stress environments, which can negatively impact their physical and emotional health. Given the high levels of stress and suboptimal health risks faced by medical students, a reliable and validated tool to identify and monitor their health status accurately is of great importance. Identifying a reliable and valid tool to evaluate SHS questionnaires in this population is crucial for enabling early detection, targeted interventions, and promoting overall well-being. The

findings of this comprehensive psychometric evaluation will provide valuable insights into the cross-cultural applicability and reliability of the SHSQ-25 within the context of medical education in Hanoi, Vietnam. This information can inform the development and implementation of tailored support services and health promotion strategies to address the unique needs of medical students and contribute to cultivating a healthier and more resilient future healthcare workforce in Vietnam and potentially in other educational settings.

Therefore, this study aims to evaluate the reliability, and validity of the SHSQ-25 using data collected from medical students in Hanoi, addressing the topic: "Validity and Reliability of the SHSQ-25 for Assessing Suboptimal Health Status Among Hanoi Medical University Students."

2. Methods

2.1. ***Data collection***

Data were collected at Hanoi Medical University. Students were selected if (1) they were studying at Hanoi Medical University at the time of data collection and (2) they voluntarily participated and responded to the interview. A total of 353 individuals participated. The questionnaire included demographic information and 25 questions from the SHSQ-25 questionnaire. Data were collected online via REDCap in the form of a self-completion questionnaire. People who had mental health problems or a history of diagnosed physical illnesses were excluded.

2.2. ***SHSQ-25 questionnaire***

SHSQ-25 is a multidimensional subclinical state that contains 25 questions across five health domains: fatigue (9 items), cardiovascular system (3 items), digestive tract (3 items), immune system (3 items), and mental status (7 items) [5, 6].

Participants answered questions about their health status in the preceding three months using a five-point Likert-type scale (1 = never or almost never, 2 = occasionally, 3 = often, 4 = very often, and 5 = always) [6].

2.3. ***Translation***

The questionnaire was translated from English into Vietnamese. In order to preserve the original meaning of the questionnaire, a reverse translation back to English by other team members was held. We piloted the questionnaire with 20 students in Hanoi Medical University.

2.4. ***Sample size and sampling.***

The study employed a subject-to-item ratio of 1:10. Given that the SHSQ-25 questionnaire contains 25 questions, the

minimum required sample size was 250 subjects [11, 12].

The actual sample size used in this study was 353 subjects, which exceeds the minimum requirement.

2.5. Data analysis

All analyses were executed in **Rstudio version 4.0**. Descriptive statistics of items were calculated using the `psych` package, while CFA was carried out using the `lavaan` package. Scale reliability analysis employed the Intraclass Correlation Coefficient (ICC) to assess test-retest reliability and Cronbach's **α** with item analyses to evaluate internal consistency. An initial EFA utilized the Maximum Likelihood method and Promax rotation to determine key factors for dimensionality assessment. Subsequently, to validate the results, we used CFA to evaluate the fit of a 5-factor solution using multiple model fit indices. Factor loading patterns, total variance explained, and inter-factor correlations were inspected to confirm the measurement model's validity. Scale reliability, EFA, and CFA collectively established the evidence for scale validation.

2.6. Ethical considerations

Information provided by participants is kept confidential and used for research purposes only. Participants were informed of the research purpose, guaranteed to understand it, and completely voluntarily participated

3. Results

3.1. Characteristics of participants

Table 1 presents the characteristics of 353 students participating in the study. Sex, BMI, year of study, majors, current residence, and academic performance were reported.

Table 1. Characteristics of participants (n = 353)

Characteristics	n	%
Sex		
Male	147	41.6
Female	206	58.4
BMI		
<18.5	61	17.3
18.5-24.9	268	75.9
>25	24	6.8
Year of study		

Year 1	122	34.6
Year 2	35	9.9
Year 3	37	10.5
Year 4	80	22.6
Year 5	69	19.5
Year 6	10	2.8
Major		
Medicine	131	37.1
Dentistry	16	4.5
Traditional Medicine	11	3.1
Public Health	195	55.3
Current Residence		
Dormitory	111	31.4
Stay with family	83	23.5
Staying alone	29	8.3
Staying with friends	100	28.3
Staying with relatives	24	6.8
Other (specify)	6	1.7
Most recent academic performance		
Excellent (9.0 - 10)	13	3.7
Outstanding (8.0 - 8.9)	84	23.8
Good (7.0 - 7.9)	121	34.3
Above Average (6.0 - 6.9)	83	23.5
Average (5.0 - 5.9)	39	11.0
Poor (4.0 - 4.9)	39	11.0

3.2 Reliability

3.2.1 Test-retest reliability

Table 2 shows that the Intraclass Correlation Coefficient (ICC) of the four subscales demonstrates good reliability, with values ranging from 0.416 to 0.703. However, the “Immune system” subscale exhibits lower agreement (0.276), indicating less consistency between measurements.

Table 2. Subscale Characteristics

Subscale	No. of Items	Mean ± SD	Cronbach's α	IIC	ICC* (95% CI)
Fatigue	9	2.4 ± 0.69	0.90	0.91	0.494 (0.452-0.539)
Cardiovascular system	3	1.8 ± 0.78	0.88	0.83	0.703 (0.659-0.745)
Digestive tract	3	1.9 ± 0.76	0.81	0.75	0.591 (0.536-0.644)
Immune system	3	2.4 ± 0.69	0.53	0.49	0.276 (0.209-0.345)
Mental status	7	2.5 ± 0.74	0.90	0.90	0.558 (0.514-0.602)
Total	25	2.3 ± 0.61	0.95	0.96	0.416 (0.379-0.456)

* ICC from the 2-way mixed model

3.2.2 Internal consistency

The SHSQ-25 demonstrates strong internal consistency, with Cronbach's alpha values ranging from 0.81 to 0.90 and IIC values between 0.75 and 0.91, except for the Immune system subscale, which shows lower internal consistency (Cronbach's α = 0.53, IIC = 0.49), indicating potential areas for improvement in this particular dimension of the questionnaire.

3.3 Validity

3.3.1 Exploratory factor analysis

Bartlett's test of sphericity (p-value < 0.001) and the KMO value of 0.94 suggest that the variables share common variance and are suitable for factor analysis.

Table 3 compares the key characteristics and findings from the method used in the current study with other commonly employed EFA approaches [6, 8, 9]. While both Maximum Likelihood extract five inter-correlated factors, the use of varimax rotation resulted in slightly improved model fit indices (CFI = 0.890, TLI = 0.875, RMSEA = 0.082) compared to the promax rotation used in the current study. However, the loading of one item (c17) is poor (≤0.30) onto the factors, which is undesirable given the goal of preserving the full 25-item scale. Parallel analysis suggested a 2-factor solution, resulting in a notably poorer fit (CFI = 0.802, TLI = 0.783, RMSEA = 0.106). As such, the Promax solution represents the most viable factor structure for preserving the integrity of the original SHSQ-25 items and subscales based on these analyses.

Table 3. EFA Methods and Model Fit Indices

EFA Method	No. of Factors	Model fit indices		
		CFI	TLI	RMSEA
Maximum Likelihood, Promax	5	0.878	0.862	0.085
Maximum Likelihood, Varimax	5	0.890	0.875	0.082
Parallel Analysis	2	0.802	0.783	0.106

Table 4 shows the factor loadings using the Maximum-Likelihood Method and rotation Promax, with the factor interpreted when loadings were greatest in each variable. Based on the structure revealed by the analysis, factor 1 shares items 11 to 16; factor 2 shares items 20 to 24; factor 3 shares items 3 to 10 and 25; factor 4 shares items 1 and 2; factor 5 shares items 17 to 19. They are different from the original subscales. Depending on the characteristics of each item in the newly defined factors, these factors were renamed to better reflect the domains being assessed by this questionnaire. The five factors of the Vietnamese version in this study include 1) Cardiovascular/Digestive issues, 2) Mental status, 3) General somatic symptoms, 4) Exhaustion, and 5) Sleep disturbances.

Table 4. Standardized Loadings (Pattern Matrix)

Item no	Abbreviated item-label	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Com
1	Exhaustion	0.03	0.01	0.04	**0.79**	−0.05	1.0
2	Chronic fatigue	0.14	0.05	0.01	**0.70**	0.05	1.1
3	Lethargy when working	−0.22	0.28	**0.37**	0.36	0.05	3.6
4	Headache	−0.05	−0.21	**0.70**	0.15	0.27	1.6
5	Dizziness	0.16	−0.10	**0.57**	0.17	0.07	1.4
6	Aching or tired eyes	−0.11	0.09	**0.77**	−0.02	−0.03	1.1
7	Sore throat	0.16	0.03	**0.56**	0.04	−0.15	1.3
8	Muscle or joint stiffness	0.44	−0.09	**0.57**	−0.02	−0.11	2.0
9	Ache in shoulder/neck/waist	0.03	0.03	**0.64**	−0.03	0.05	1.0
10	Heavy feeling in legs	0.38	−0.01	**0.44**	0.09	−0.09	2.2
11	Breathlessness	**0.87**	−0.03	−0.10	0.07	0.04	1.0
12	Chest congestion	**0.89**	−0.01	−0.07	0.06	−0.02	1.0
13	Heart palpitations	**0.76**	0.10	−0.09	0.09	−0.06	1.1
14	Poor appetite	**0.62**	−0.06	0.03	0.00	0.20	1.2
15	Upset stomach	**0.72**	−0.01	0.05	−0.07	0.03	1.0
16	Indigestion	**0.63**	0.00	0.25	−0.13	−0.01	1.4
17	Cold intolerance	−0.16	0.04	0.25	−0.08	**0.32**	2.6
18	Difficulty falling asleep	0.05	0.13	0.00	0.06	**0.59**	1.1
19	Waking up during the night	0.37	−0.04	−0.18	0.05	**0.71**	1.7
20	Impaired short-term memory	0.06	**0.62**	−0.02	0.08	0.09	1.1
21	Inability to respond quickly	0.11	**0.65**	0.08	0.09	−0.06	1.1
22	Difficulty concentrating	−0.07	**0.99**	−0.01	−0.10	0.02	1.0
23	Distracted for no reason	−0.04	**0.98**	−0.07	0.03	−0.03	1.0
24	Nervous or jittery	0.30	**0.35**	0.13	−0.02	0.18	2.8
25	Frequently catch colds	0.20	0.14	**0.38**	−0.16	0.05	2.3

3.3.2 Confirmatory factor analysis

CFA results supported a reasonable fit (χ^2 = 937; p < 0.001; RMSEA = 0.085 (95% CI); CFI = 0.878; TLI = 0.862) for this 5-factor model, providing evidence for the underlying factor structure identified through the exploratory analysis.

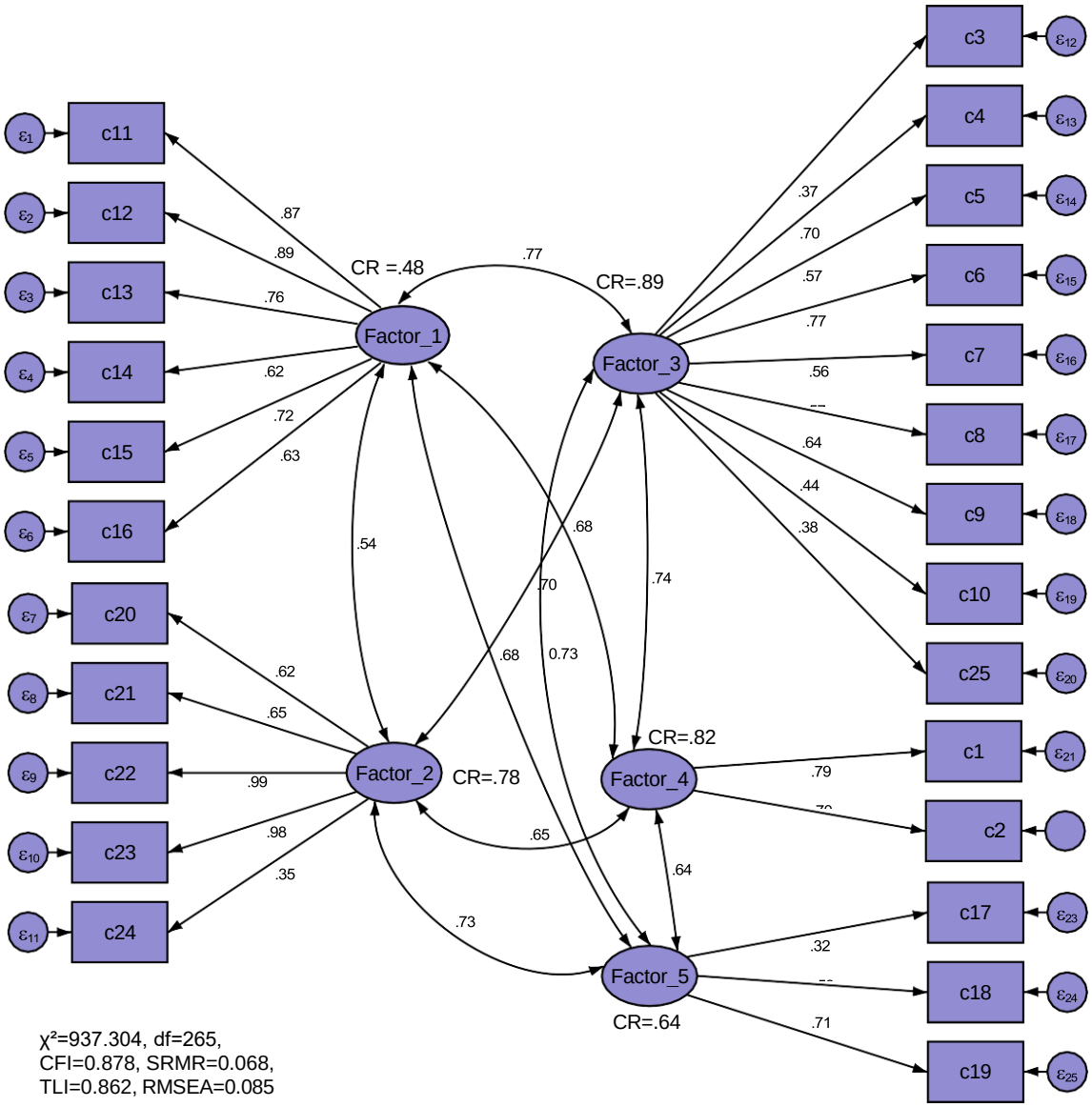

Figure 1. SEM model

Table 5 shows that Factors 2, 3, and 4 have high Cronbach's **α** (ranging from 0.78 to 0.89) and acceptable to good reliability based on ICCs (0.471 to 0.701) while Factors 1 and 5 exhibit lower Cronbach's **α** of 0.48 and 0.64, respectively, and poor reliability based on ICCs (0.315 and 0.369).

Table 5. Description of Factors reliability

Subscale	No. of Items	Mean ± SD	Cronbach's α	IIC	ICC* (95% CI)
Factor 1	6	2.5 ± 0.76	0.48	0.32	0.315(0.218 − 0.406)
Factor 2	5	2.3 ± 0.79	0.78	0.73	0.541(0.482 − 0.598)
Factor 3	9	2.4 ± 0.66	0.89	0.89	0.471(0.428 − 0.516)
Factor 4	2	2.3 ± 0.89	0.82	0.70	0.701(0.644 − 0.701)

Factor 5	3	2.4 ± 0.76	0.64	0.58	0.369(0.303 − 0.435)

* ICC from the 2-way mixed model

4. Discussion

The Vietnamese version of SHSQ-25 appeared valid and reliable among a sample of Hanoi Medical University students. Our results demonstrate acceptable reliability and validity for most subscales while highlighting areas that require further refinement.

Reliability:

The subscales measuring fatigue, mental status, cardiovascular system, and digestive tract symptoms exhibited high internal consistency (0.81-0.90). Additionally, these subscales demonstrated good reliability according to intraclass correlation coefficients (0.416-0.703). These findings suggest that the items within these subscales effectively measured the intended constructs related to suboptimal health status. The high internal consistency and reliability provide confidence in the subscales' ability to consistently capture the targeted aspects of fatigue, mental health issues, cardiovascular symptoms, and digestive problems. However, the immune system subscale performed relatively poorly, with a low Cronbach's **α** of 0.53 and an ICC of only 0.276, indicating poor internal consistency and reliability. The low agreement between items in this subscale suggests that it may not adequately capture the intended aspects of immune system dysfunction or suboptimal immune health. Revisions to the items within this subscale may be warranted to improve its psychometric properties.

Validity:

EFA revealed a different factor structure compared to the original SHSQ-25. In the current study, the EFA used the Maximum Likelihood method and Promax rotation. The confirmatory factor analysis indicated a satisfactory fit of the data to the factor structure (CFI=0.878; TLI=0.862).

Five intercorrelated factors emerged, labeled Cardiovascular/Digestive issues, Mental status, General somatic symptoms, Exhaustion, and Sleep disturbances in the results. The strong psychometric properties of Factors 2, Factor 3, and Factor 4 are encouraging. The high internal consistency, strong item-total correlations, and acceptable to good reliability indicate these factors are well-defined and measured reliably within the overall model. This suggests the constructs represented by these factors are being captured effectively by the assessment instrument. These results

support the validity and reliability of these specific factors in the five-factor model. However, given the poorer performance of Factor 1 and Factor 5, additional research may be needed to further optimize the factor structure and item composition of the SHSQ-25. These findings suggest that while the SHSQ-25 effectively captures certain aspects of suboptimal health status, some factors may require further refinement or revision. Correlational analyses indicated strong positive correlations among fatigue, cardiovascular symptoms, immune symptoms, digestive issues, and mental status. These associations align with existing literature highlighting the interrelationships between various suboptimal health symptoms and domains [7].

In previous research on the SHSQ-25, Korean and Chinese employed EFA using Maximum Likelihood extraction and Promax rotation, similar to the approach taken in the current investigation [6,9]. A study of Korean and Chinese found a 5-factor structure close to the original conceptualization of the SHSQ-25 subscales with Cronbach's α values from 0.77 to 0.94 (Korean) and 0.70 to 0.86 (Chinese). Meanwhile, in the current research, the EFA revealed a slightly different 5-factor structure. Additionally, Cronbach's α values range from 0.48 to 0.89, with two factors having poor internal consistency. Korean and Chinese validation studies reported excellent fit indices for a 5-factor model, with RMSEA values of 0.069 (Korean) and 0.044 (Chinese), GFI values of 0.929 (Korean) and 0.914 (Chinese). However, the CFA result in the current study only provided a reasonable RMSEA of 0.085 and an excellent GFI of 0.988.

A study in Ghana also used the SEM model to evaluate the construct validity of the SHSQ-25 [8]. Using parallel analysis, the researcher in Ghana explored a 3-factor model, which demonstrated good internal consistency, with Cronbach's α values ranging from 0.821 to 0.861. Moreover, the CFA in Ghana has excellent fit indices when RMSEA = 0.049 $<$ 0.08 (90%CI: 0.041, 0.056), CFI = 0.903, GFI = 0.880, TLI = 0.907, SRMR = 0.055.

Although the current study uses the same method as Korean and Chinese and extracts new factor models as Ghana, it suggests potential cultural or contextual differences, which may have affected the factor structure and psychometric properties of SHSQ-25 in this setting.

5. Conclusion

This research evaluated the validity and reliability of the SHSQ-25 among Hanoi Medical University students. The

findings demonstrated that the Vietnamese version of the SHSQ-25 exhibited strong reliability, with internal consistency measures (Cronbach's α) meeting recommended thresholds for all five subscales. Additionally, the questionnaire showed good construct validity, as Confirmatory Factor Analysis (CFA) results supported the proposed five-factor structure, demonstrating a good fit with the data.

References

[1] WHO. (1946). Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference. Official Records of the World Health Organization. 2:100

[2] Ruggeri K, Garcia-Garzon E, Maguire Á, et al. (2020). Well-being is more than happiness and life satisfaction: a multidimensional analysis of 21 countries. Health Qual Life Outcomes 18, 192.

[3] Wang W, Yan Y, Guo Z, et al. (2021). All around suboptimal health — a joint position paper of the Suboptimal Health Study Consortium and European Association for Predictive, Preventive and Personalised Medicine. EPMA Journal 12, 403–433. https://doi.org/10.1007/s13167-021-00253-2

[4] Zhou LM, Shi Y. (2019). Association of Suboptimal Health Status with Psychosocial Stress, Plasma Cortisol and mRNA Expression of Glucocorticoid Receptor α/β in Lymphocytes. Stress, 22(6), 697-706. doi: 10.1080/10253890.2019.1628145

[5] Hu F, Wang W. (2023). Suboptimal Health Innovation: From Bench to Bedside. In: Pod-Bielska, H., Kapalla, M. (eds) Predictive, Preventive, and Personalised Medicine: From Bench to Bedside. Advances in Predictive, Preventive and Personalised Medicine, vol 17. Springer, Cham. https://doi.org/10.1007/978-3-031-34884-6_2

[6] Yan YX, Liu YQ, Li M, et al. (2009). Development and Evaluation of a Questionnaire for Measuring Suboptimal Health Status in Urban Chinese. J Epidemiol. 19(6):333–41.

[7] Wang W, Yan Y. (2012). Suboptimal health: a new health dimension for translational medicine. Clinical and Translational Medicine. doi: 10.1186/2001-1326-1-28

[8] Adua E, Afrifa-Yamoah E, Frimpong K, et al. (2021). Construct validity of the Suboptimal Health Status Questionnaire-25 in a Ghanaian

population. Health Qual Life Outcomes. 19(1):180. doi: 10.1186/s12955-021-01810-z.

[9] Guo Z, Meng R, Zheng Y, et al. (2022). Translation and cross-cultural validation of a precision health tool, the Suboptimal Health Status Questionnaire-25, in Korean. J Glob Health. 12:04077. doi: 10.7189/jogh.12.04077.

[10] Duong, Thi & Thien, Phuoc & Huong, Lam. (2022). Characteristics according to traditional medicine in suboptimal health status and some related factors among medical students in University of Medicine and Pharmacy, Hue University. Journal of Medicine and Pharmacy. 59-66.

[11] Anthoine E, Moret L, Regnault A, et al. (2014). Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes. 12(1):2. doi:10.1186/s12955-014-0176-2

[12] Gunawan J, Marzilli C, Aungsuroch Y. (2021). Establishing appropriate sample size for developing and validating a questionnaire in nursing research. Belitung Nurs J. 7(5):356-360. doi:10.33546/bnj.1927

APPENDIX. The following questions inquire about health events during the last 3 months. Answer every question by marking the appropriate box with an 'x'. You may choose from one of the following answers:

1. never or almost never
2. occasionally
3. often
4. very often
5. always

Table 6. Suboptimal Health Status Questionnaire (SHSQ-25)

No.	Questions	1	2	3	4	5
1	Were exhausted without greatly increasing your physical activity.	□	□	□	□	□
2	Fatigue could not be substantially alleviated by rest.	□	□	□	□	□
3	Were lethargic when working.	□	□	□	□	□
4	Suffered from headaches.	□	□	□	□	□

5	Suffered from dizziness.	□	□	□	□	□
6	Eyes ached or were tired.	□	□	□	□	□
7	Suffered from a sore throat.	□	□	□	□	□
8	Muscles or joints felt stiff.	□	□	□	□	□
9	Have pain in your shoulder/neck/waist.	□	□	□	□	□
10	Have a heavy feeling in your legs when walking.	□	□	□	□	□
11	Feel out of breath while sitting still.	□	□	□	□	□
12	Suffered from chest congestion.	□	□	□	□	□
13	Were bothered by heart palpitations.	□	□	□	□	□
14	Appetite is poor.	□	□	□	□	□
15	Suffered from heartburn.	□	□	□	□	□
16	Suffered from nausea.	□	□	□	□	□
17	Could not tolerate the cold.	□	□	□	□	□
18	Had difficulty falling asleep.	□	□	□	□	□
19	Had trouble waking up during the night.	□	□	□	□	□
20	Had trouble with your short-term memory.	□	□	□	□	□
21	Could not respond quickly.	□	□	□	□	□
22	Had difficulty concentrating.	□	□	□	□	□
23	Were distracted for no reason.	□	□	□	□	□
24	Felt nervous or jittery.	□	□	□	□	□
25	Caught a cold in the past 3 months.	□	□	□	□	□