Jurnal Kedokteran dan Kesehatan: Publikasi Ilmiah Fakultas Kedokteran Universitas Sriwijaya
Volume 13, No 2. 2026/DOI: 10.32539/JKK.V13i2.728
p-ISSN 2406-7431; e-ISSN 2614-0411
Page: 183-191

VALIDITY, RELIABILITY, AND DIAGNOSTIC PERFORMANCE OF GERIATRIC-8 AS A SCREENING TOOL FOR ABNORMAL COMPREHENSIVE GERIATRIC ASSESSMENT

Ridzqie Dibyantari[1], Nina Kemala Sari [2], Edy Rizal Wahyudi[2], Muhadi[3], Aulia Rizka[2], Arif Mansjoer[4], Noorwati Sutandyo[5].

[1]Geriatric Medicine Division, Internal Medicine Department, Faculty of Medicine, Universitas Sriwijaya, Palembang
[2]Geriatric Medicine Division, Internal Medicine Department, Faculty of Medicine, Universitas Indonesia, Jakarta
[3]Clinical Epidemiology Unit, Internal Medicine Department, Faculty of Medicine, Universitas Indonesia, Jakarta
[4]Cardiology Division, Internal Medicine Department, Faculty of Medicine, Universitas Indonesia, Jakarta
[5]Haematologi and Medical Oncology Division, Internal Medicine Department, Faculty of Medicine, Universitas Indonesia, Jakarta

ARTICLE INFO

Corresponding author :
Ridzqie Dibyantari
Geriatric Medicine Division,
Faculty of Medicine,
Universitas Sriwijaya

Email:
ridzqie@fk.unsri.ac.id

Kata kunci:
Pengkajian paripurna pasien geriatri
G8
Uji kesahihan dan keandalan

Keywords:
Comprehensive geriatric assessment
Geriatric 8
Validity and reliability study

Original submisson:
August 27, 2025
Accepted:
March 27, 2026
Published:
April 25, 2026

ABSTRAK

Pengkajian Paripurna Pasien Geriatri (P3G) bermanfaat dalam tata laksana lansia, namun memerlukan waktu lama sehingga dikembangkan instrumen penapisan seperti Geriatric 8 (G8). Awalnya untuk pasien kanker, G8 juga terbukti valid pada lansia nonkanker, tetapi belum ada studi di Indonesia. Penelitian potong lintang ini (Agustus 2023–Maret 2024) di Poliklinik Geriatri RSUPN Cipto Mangunkusumo melibatkan 80 subjek yang menjalani P3G dan G8. P3G abnormal ditentukan bila ≥1 domain terganggu (ADL <20, IADL <8, MNA <24, MoCA-INA <26, atau GDS >4). Hasil menunjukkan reliabilitas kuat (kappa inter-rater 1; intra-rater 0,904; ICC 0,77–1; Cronbach α 0,697). Sensitivitas G8 70,27% dan spesifisitas 83,33% dengan AUC 0,846 (p<0,005). Temuan ini menunjukkan performa diagnostik yang baik dalam mengidentifikasi lansia yang memerlukan P3G lengkap. Disimpulkan bahwa G8 sahih dan andal sebagai alat penapisan lansia di populasi umum Indonesia dengan spesifisitas yang baik.

ABSTRACT

The Geriatric Patient Plenary Assessment (P3G) is useful in the management of the elderly, but it is time-consuming, so screening instruments such as the Geriatric 8 (G8) have been developed. Initially designed for cancer patients, the G8 has also been shown to be valid in non-cancer elderly patients, but there have been no studies in Indonesia. This cross-sectional study (August 2023–March 2024) at the Geriatrics Clinic of Cipto Mangunkusumo National Hospital involved 80 subjects who underwent P3G and G8. An abnormal P3G was defined as impaired in ≥1 domain (ADL <20, IADL <8, MNA <24, MoCA-INA <26, or GDS >4). The results showed strong reliability (inter-rater kappa 1; intra-rater 0.904; ICC 0.77–1; Cronbach's α 0.697). The G8 had a sensitivity of 70.27% and a specificity of 83.33%, with an AUC of 0.846 (p<0.005). These findings demonstrate good diagnostic performance in identifying elderly individuals requiring a complete P3G. It is concluded that the G8 is valid and reliable as a screening tool for elderly individuals in the general Indonesian population, with good specificity.

INTRODUCTION

According to data from the National Centre of Statistics (BPS), the proportion of older adults in Indonesia has increased to 10.82% of the total population. The life expectancy of older adults also increased to 71.52 years.[1] Therefore, health services are increasingly needed for older adults.

The comprehensive geriatric assessment (CGA) is the gold standard for healthcare services for older adults.[2] CGA is beneficial for diagnosing and managing health problems in older adults.[2] CGA has been shown to be beneficial in emergency, inpatient, and outpatient settings.[3–5] Domains

reviewed in CGA include functional status and disability, cognitive status, nutritional status, mental status, and mobility.[2] However, the primary limitation of the CGA is the significant time.[6] Hence, extensive research was conducted to develop a short form of CGA as a screening tool for patients who needed the full CGA examination.[6–8]

Geriatric 8 (G8) is a screening tool proposed by Bellera et al. for older adults with cancer. Each question in the G8 was designed to correlate for its association with at least 2 CGA domains. G8 showed statistically significant associations with Activities of Daily Living (ADL), Instrumental Activities of Daily Living (IADL), Mini-Nutritional Assessment (MNA), Geriatric Depression Scale (GDS), and Mini-Mental State Exam (MMSE).[9] This instrument has good diagnostic performance as a screening tool for older adults with or without cancer.[10–12] To our knowledge, no research has been published regarding using G8 as a screening tool for older adults in Indonesia. Therefore, the aim of this study is to evaluate the validity and reliability of G8 as a screening tool for older adults requiring a full CGA.

METHODS

This cross-sectional, descriptive-analytic study aimed to determine the validity, reliability, and diagnostic performance of G8 as a screening tool for older adults in outpatient settings. This research was conducted in Cipto Mangunkusumo National Central General Hospital from August 2023 to March 2024. Consecutive sampling included patients who met the study's recruitment criteria. Inclusion criteria were age 60 or older and willingness to participate, as evidenced by signing an informed consent form. Exclusion criteria included deformity or paresis of the dominant hand, limiting the subject's ability to write and/or draw.

Data collected were baseline characteristics (age, sex, education, comorbidities, polypharmacy, caregivers, and residence), G8 score, CGA which is the gold standard for geriatric assessment consist of activity of daily living (ADL), instrumental activity of daily living (IADL), Montreal Cognitive Assessment Indonesia Version (MoCA-INA), Geriatric Depression Scale (GDS), Mini Nutritional Assessment (MNA), Cumulative Illness Rating Scale for Geriatrics (CIRS-G), SARC-F, SARCALF, FRAIL, and timed up and go. The instruments utilized represent standard tools in a full CGA to measure each domain of older adults. The cutoff values are: ADL <20, IADL <8, MNA <24, MoCA-INA <26, and GDS >4. Whereas in one domain below or above the cutoff value, the subject is considered to have abnormal CGA. Sample size was determined using the correlation sample size calculation and the diagnostic test sample size calculation. From each calculation, the sample size needed in this research to determine the correlation between G8 and CGA is 20 people, and the number of subjects needed for the diagnostic test is 80 people, respectively. Hence, the minimum sample size in this research was 80. To evaluate reliability, 21 subjects underwent duplicate G8 examinations by a general practitioner and an internist at two different times in order to calculate ICC.

This research has obtained ethical approval from the Medical Research Ethics Committee of the FKUI-RSCM: KET-1134/UN2.F1/ETIK/PPM. 00.02/2023.

RESULT

Eighty eligible subjects met the inclusion criteria. Among the subjects, 21 underwent a G8 examination by the second person on the same day. The second G8 score was obtained within 14 days of the first examination. Figure 1 explains the research flowchart.

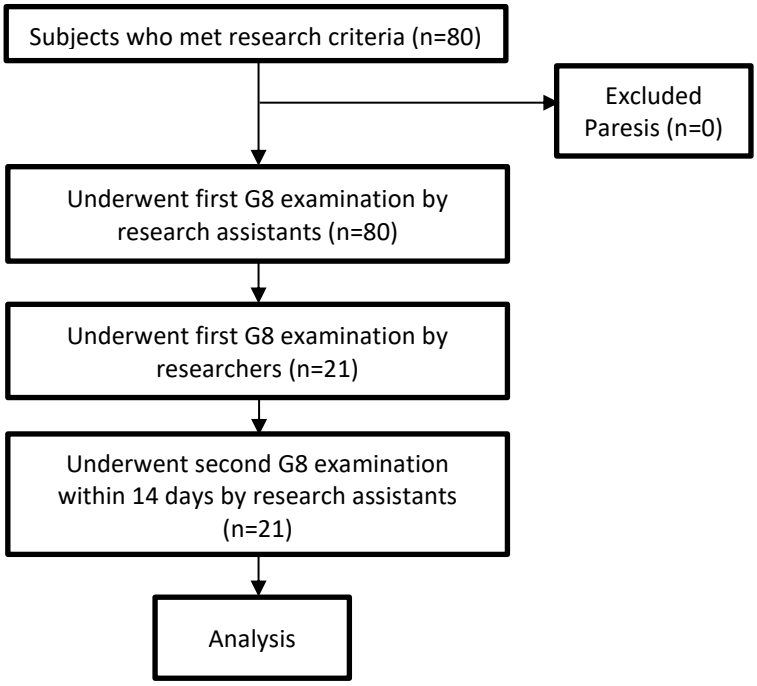

Figure 1. Subject recruitment flowchart

The mean age was 73.68 (SD 6.36) years. Most subjects were female (66.3%) and had advanced levels of education (academy to doctorate; 56.3%). The majority of the subjects were independent (82.5%), without disability (67.5%), and had no depression (86.3%). Cognitive impairments were found in 47.5% of the research subjects. Mobility disorder was found in 48.8% of the subjects. Nutritional problems, i.e., being at risk of malnutrition and malnutrition, were found in 32.5% of the subjects. Most subjects did not have sarcopenia (87.5%), while a fit/robust proportion was found in 46.3%. Most subjects (92.5%) had a disorder in at least one CGA domain.

Table 1. Baseline characteristics

Variables	N=80
Age, mean (SD)	73.68 (6.36)
Sex, n (%)	
Male	27 (33.8)
Female	53 (66.3)
Education, n (%)	
None	1 (1.3)
Basic education	34 (42.5)
Advance education	45 (56.3)
Functional status, n (%)	
Independent	66 (82.5)
Dependence	14 (17.5)
Disability, n (%)	
With disability	26 (32.5)

Without disability	54 (67.5)
Cognitive status, n (%)	
None	42 (52.5)
Mild cognitive impairment	38 (47.5)
Mobility	
Normal	41 (51.2)
Abnormal	39 (48.8)
Nutritional status	
Malnutrition/risk of malnutrition	26 (32.5)
Normal	54 (67.5)
Mental status, n (%)	
Not depressed	69 (86.3)
Probably depressed	11 (13.8)
Polypharmacy, n (%)	
Medicine < 5	22 (27.5)
Medicine ≥ 5	58 (72.5)
CIRS-G, median (IQR)	3,55 (2.80 – 3.98)
Sarcopenia, n (%)	
None	70 (87.5)
Probable	10 (12.5)
Frailty, n (%)	
Prefrail/frail	43 (53.8)
Fit	37 (46.3)
CGA, n (%)	
Normal	6 (7.5)
Problem	74 (92.5)

Internal validity was assessed by testing the correlation of the G8 score with the score of each CGA domain. A normality test was performed before the correlation test.

Table 2. G8 score correlation with CGA.

CGA	*r*	*p-value**
ADL	0.421	<0.0001
IADL	0.420	<0.0001
MoCA-INA	0.611	<0.0001
GDS	-0.278	0.013
MNA	0.721	<0.0001
CIRS-G	-0.083	0.465

*Spearman correlation test

G8 has a strong correlation with MNA and MoCA-INA, a moderate correlation with ADL and IADL, and a weak correlation with GDS. Each G8 question was also tested for correlation with at least two CGA domains to determine the validity. Table 3 shows correlation test results between G8 questions and CGA domains.

Table 3. Correlation of G8 questions with CGA domains.

G8 Questions	CGA	*r*	*p-value*
A (food intake)	ADL	0.465	<0.0001
	IADL	0.232	0.038
	MNA	0.602	<0.0001
B (weight loss)	ADL	0.244	0.029
	IADL	0.217	0.053
	MNA	0.380	0.001
C (mobility)	ADL	0.394	<0.0001
	IADL	0.216	0.054
	TUG	-	0.047*
D (neuropsychology problem)	MoCA-INA	0.444	<0.0001
	GDS	-0.057	0.616
E (BMI)	ADL	0.191	0.090
	IADL	0.14	0.210
	MNA	0.402	<0.0001
F (consume > 3 medicines)	CIRS-G	-0.290	0.009
G (health perspective)	ADL	0.240	0.032
	IADL	0.289	0.009
	GDS	-0.236	0.035
	CIRS-G	-0.146	0.196

Spearman correlation test; *Mann Whitney-Test

In this study, we employed test-retest reliability, inter-rater agreement, and internal consistency to assess the reliability of G8 as a screening instrument. The test-retest analysis was performed by administering the G8 twice, by different examiners at two separate time points (within 14 days of the initial assessment) of the first examination. The kappa for different times was 0.904 ($p<0.001$), and for the different persons was 1 ($p<0.001$).

Interrater reliability was assessed by calculating the interclass correlation coefficient (ICC) for the G8 scores obtained by two examiners. Table 4 shows the ICC. Cronbach was used to evaluate internal consistency. This study found a Cronbach α of 0,697, above the required significance, i.e., 0.6.[13]

Table 4. ICC G8 between two examiners

G8 Questions	ICC (95%CI)
A	0.941 (0.847 – 0.977)
B	1.000 (1.000 – 1.000)
C	1.000 (1.000 – 1.000)
D	1.000 (1.000 – 1.000)
E	1.000 (1.000 – 1.000)
F	1.000 (1.000 – 1.000)
G	0.774 (0.412 – 0.913)
H	1.000 (1.000 – 1.000)

The diagnostic performance of G8 as a geriatric syndrome screening tool is displayed in Table 5. G8 has good sensitivity and specificity to CGA. The sensitivity was 70.27 (95%CI 58.82 – 80.34), and the specificity was 83.33 (95%CI 35.88 – 99.58).

Table 5. Diagnostic Performance of G8 as A Screening Tool for Geriatric Syndrome

	Sensitivity	Specificity	PPV	NPV
CGA	70.27	83.33	98.11	18.52
	(58.82 – 80.34)	(35.88 – 99.58)	(89.62 – 99.68)	(12.11 – 27.27)

The diagnostic accuracy of G8 compared to CGA was displayed as ROC. We determined the disorder of one domain as disordered or abnormal CGA. The *Area Under Curve* (AUC) was 0.846, p <0.005 (95%CI 0.677-1.000) at a cut-off score of 14.

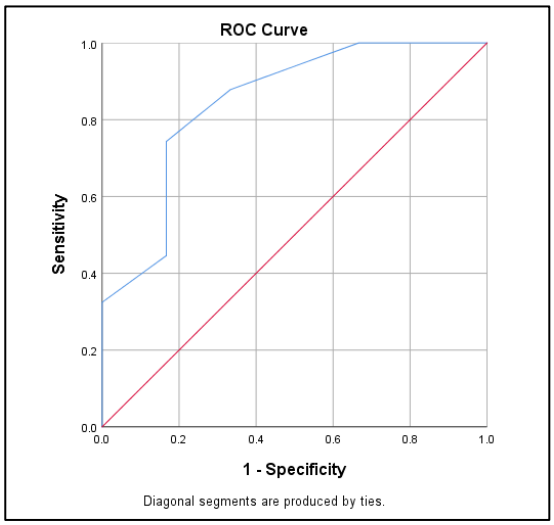

Figure 2. Receiver Operating Characteristic (ROC) G8 – CGA

DISCUSSION

Generally, The baseline characteristics of these study subjects are consistent with national data in previous research. However, our subjects have higher educational levels than the nationwide data concerning the educational attainment of older adults in Indonesia. The nationwide census reported that the mean level of formal education for the older adult population was 5 years, whereas our study found that 56.3% of subjects held degrees ranging from vocational diplomas to doctorates.[1] Another study conducted in the same Geriatric Integrated Outpatient Clinic found a result similar to ours, with an advanced education level proportion of 69.5%.[14]

Validity was evaluated to determine the degree to which the instrument measures its intended construct. Construct validity was assessed in five steps: assessment of the questionnaire's representativeness, feedback from respondents, the questionnaire's internal structure, comparison with other instruments, and evaluation.[15] The G8 was originally designed as a screening instrument for older adults with cancer who need a complete CGA examination. This instrument consists of 8 questions on food intake, weight loss, mobility, neuropsychological problems, body mass index, daily medication intake, self-perceived health, and age. The maximum score is 17, and the lowest is 0.[9]

The next aspect of construct validity is internal structure assessment, which determines the

correlation coefficient between questions and the G8 total score. In this research, correlation coefficients ranged from very weak to very strong (-0.083 to 0.721). Cavusoglu found a moderate negative correlation between GDS and G8 total score ($r = -0.541$, $p<0.001$).[12] The weak correlation with the G8 score suggests low internal consistency for the G8.

The G8 instrument is compared to other known instruments to assess geriatric syndrome. In this research, the instruments used to compare were ADL, IADL, MNA, TUG, MoCA-INA, GDS, and CIRS-G. Correlation coefficients varied from very weak to strongly correlated. GDS score was weakly correlated ($r=-0.057$) with the neuropsychiatry question ($p=0.616$). This finding differs from previous research, which showed a correlation between the GDS score and the neuropsychiatric question of G8 ($p<0.001$).[9]

A reliability test was performed using test-retest, interrater, and intra-rater coefficients, and an internal consistency test.[15] This research showed a high correlation coefficient in both inter-rater and intra-rater. The kappa value in this research was higher than that from Cavusoglu et al. research, i.e., 0.886 and 0.875 ($p<0.001$) for interrater and intra-rater, respectively.

Inter-rater agreement was determined by counting ICC. Subjects were reassessed by a second examiner. A very strong correlation was found on questions A, B, C, D, E, F, and H. Question G showed a strong correlation.[15] This result showed that examiners were adequately trained to use G8, and subjects' ability to understand G8 questions was also good. An internal consistency test was performed by calculating Cronbach α for G8 with an acceptable value of> 0,6.13. This study found Cronbach α of 0,697, which means G8 has satisfactory internal consistency.

Diagnostic performance analysis G8 was conducted against CGA and showed a sensitivity of 70.27% (58.82-80.34) and a specificity of 83.33% (35.88-99.58) with a cut-off value of 14.25. Cavusoglu et al. also found that the cut-off value 14.25 has a sensitivity of 80.92 (73.1-87.3) and a specificity of 89.86 (80.2-95.8).[12] Research conducted by Smets et al. found G8 sensitivity was 78 (74-82) and specificity was 68 (63-72) on cut-off value 14.[10] G8 and other CGA short forms (i.e., VES-13, TRST, GFI, and aCGA) were tested to determine their diagnostic performance as screening tools for older adults in both the cancer and noncancer groups.[11] These studies demonstrated that the G8 outperforms other abbreviated CGA regarding its sensitivity, specificity, and ability to predict mortality and morbidity in cancer patients. G8 has been tested for internal consistency by assessing the correlation between the CGA score and each CGA domain question.[9]

G8 tool reliability further demonstrated by its ability to bridge gap between clinical questionnaires, physiological markers, and long-term prognostic outcomes across diverse geographic regions. Such as by data in oncology studies. In these cohorts, G8 has shown a clinically significant relationship with bioimpedentiometry (BIA) parameters, specifically the Phase angle (PhA) ($p = 0.013$).[17] This clinical utility is echoed in specialized oncology settings, where the Turkish version of the tool demonstrated "good" measurement power (AUC 0.824) and concurrent validity when measured against the Edmonton Frail Scale. While these studies highlight G8's ability to reflect functional and frailty status, its reliability is also confirmed by objective physiological markers.[18] Recent longitudinal data from older Indian patients indicates that G8 is a stronger independent predictor of 2-year mortality than more complex tools like the Onco-MPI.[19] This versatility is further confirmed by a study in a multi-ethnic Asian population in Singapore, which found that G8 performs well (AUC 0.73) in identifying patients who would benefit from a Comprehensive Geriatric Assessment (CGA). [20]

To our knowledge, this is the first study to evaluate the validity, reliability, and diagnostic performance of G8 as a screening tool for geriatric syndromes among the general older adult population in Indonesia. The study involved trained general practitioners (GPs), supporting the potential implementation of the G8 as a reliable screening tool.

This research has several limitations. Notably, mental status (as measured by the GDS) demonstrated a weak correlation with the G8 score. Hence it may be necessary to adapt the G8 questionnaire to improve its sensitivity for older adults with neuropsychiatric problems.

CONCLUSION

According to the results, we conclude that G8 is valid and reliable for use as a screening tool for geriatric syndromes in outpatient settings. The G8 also demonstrated favorable diagnostic performance, characterized by high sensitivity

We recommend further research to refine or modify the neuropsychiatric components for G8 screening and a multicenter study to validate the diagnostic performance of the G8 at a national level. We also propose training in the use of G8 for general practitioners and paramedics.

REFERENCES

1. Badan Pusat Statistik Republik Indonesia. Statistik Penduduk Lanjut Usia 2022. Jakarta: Badan Pusat Statistik; 2022.
2. Soejono CH. Pengkajian paripurna pada pasien geriatri. In: Sudoyo AW, Setiati S, Simadibrata M, Setiyohadi B, Alwi I, Syam AF, editors. Buku Ajar Ilmu Penyakit Dalam. Edisi 6. Jakarta: Interna Publishing; 2016. p. 3705-13.
3. Briggs R, McDonough A, Ellis G, Bennett K, O'Neill D, Robinson D. Comprehensive geriatric assessment for community-dwelling, high-risk, frail, older people. Cochrane Database Syst Rev. 2022;(5):CD012705. doi:10.1002/14651858.CD012705.pub2 [cited 2023 Jan 17]. Available from: https://doi.org/10.1002/14651858.CD012705.pub2
4. Mazya AL, Garvin P, Ekdahl AW. Outpatient comprehensive geriatric assessment: effects on frailty and mortality in old people with multimorbidity and high health care utilization. Aging Clin Exp Res. 2019;31(4):519-25. doi:10.1007/s40520-018-1002-x
5. Wei YJ, Hsieh CF, Huang YT, Huang MS, Fang TJ. The influence of integrated geriatric outpatient clinics on the health care utilization of older people. BMC Geriatr. 2020;20(1):379. doi:10.1186/s12877-020-01777-8
6. Overcash JA, Beckstead J, Extermann M, Cobb S. The abbreviated comprehensive geriatric assessment (aCGA): a retrospective analysis. Crit Rev Oncol Hematol. 2005;54(2):129-36. doi:10.1016/j.critrevonc.2004.11.003
7. Beauche O, Chabot J, Fung S, Launay CP. Update of the 6-Item Brief Geriatric Assessment screening tool of older inpatients at risk for long length of hospital stay: results from a prospective and observational cohort study. J Am Med Dir Assoc. 2018;19(8):720-1. doi:10.1016/j.jamda.2018.03.018
8. Tai CJ, Yang YH, Huang CY, Pan SC, Hsiao YH, Tseng TG, et al. Development of the brief geriatric assessment for the general practitioner. J Nutr Health Aging. 2021;25(1):134-40. doi:10.1007/s12603-020-1491-2
9. Bellera CA, Rainfray M, Mathoulin-Pélissier S, Mertens C, Delva F, Fonck M, et al. Screening older cancer patients: first evaluation of the G-8 geriatric screening tool. Ann Oncol. 2012;23(8):2166-72. doi:10.1093/annonc/mdr491

10. Smets IH, Kempen GI, Janssen-Heijnen ML, Deckx L, Buntinx FJ, van den Akker M. Four screening instruments for frailty in older patients with and without cancer: a diagnostic study. BMC Geriatr. 2014;14(1):26. doi:10.1186/1471-2318-14-26
11. Kenig J, Szabat K, Mituś J, Mituś-Kenig M, Krzeszowiak J. Usefulness of eight screening tools for predicting frailty and postoperative short- and long-term outcomes among older patients with cancer who qualify for abdominal surgery. Eur J Surg Oncol. 2020;46(11):2091-8. doi:10.1016/j.ejso.2020.05.016
12. Cavusoglu C, Deniz O, Dogrul RT, Ileri I, Yildirim F, Caliskan H, et al. Validity and reliability of the G8 screening test in older non-cancer patients. Eur Geriatr Med. 2021;12(2):397-404. doi:10.1007/s41999-020-00415-7
13. Cronbach LJ, Shavelson RJ. My current thoughts on coefficient alpha and successor procedures. Educ Psychol Meas. 2004;64(3):391-418. doi:10.1177/0013164404266386
14. Heppy F, Rooshoeroe AG, Dwimartutie N, Muhadi M. Performa penapisan instrumen MSRA dan instrumen MSRA-LB terhadap diagnosis sarkopenia pada pasien lanjut usia. Tesis. Jakarta: Fakultas Kedokteran, Universitas Indonesia; 2023.
15. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med. 2006;119(2):166.e7-166.e16. doi:10.1016/j.amjmed.2005.10.036
16. Cavusoglu C, Tahtaci G, Dogrul RT, Ileri I, Yildirim F, Candemir B, et al. Predictive ability of the G8 screening test to determine probable sarcopenia and abnormal comprehensive geriatric assessment in older patients with solid malignancies. BMC Geriatr. 2021;21(1):269. doi:10.1186/s12877-021-02544-9
17. Leo S, Ferrari M, Montini M, et al. Bioimpedentiometry parameters used as indicators of frailty and malnutrition: association between G8 score and phase angle (PHA) in elderly cancer patients. Aging Clin Exp Res. 2023;35(10):2219-25. doi:10.1007/s40520 023 02512 w.
18. Atakul E. Geriatric assessment in oncology: G8 screening tool validation and reliability study. Turk J Oncol [Preprint]. 2019. doi:10.5505/tjo.2019.2010.
19. Sable S, Shukla S, Kapoor S, et al. Validation of onco MPI and geriatric 8 for predicting 2 year mortality in older Indian patients with breast cancer: a prospective observational study. ecancermedicalscience. 2026;20:e2099. doi:10.3332/ecancer.2026.2099.
20. Chen MZ, Tan Y, Tan DS, et al. Evaluation of the geriatric 8 (G8) screening tool in a multi ethnic Asian population of older adults with cancer: results from the Geriatric Oncology Longitudinal End to end (GOLDEN) study. J Clin Oncol. 2022;40(16_suppl):e24029. doi:10.1200/JCO.2022.40.16_suppl.e24029.