



RESEARCH ARTICLE

Diagnostic Performance of the Japanese Version of the GAD-7 in Primary Care and Comparing its Accuracy with Other Language Versions

Kumiko Muramatsu, MD, PhD^{1,6,14}, Shunta Maeda, PhD², Hitoshi Miyaoka, MD, PhD³, Kunitoshi Kamijima, MD, PhD⁴, Yoshiyuki Muramatsu, MD^{5,6}, Yutaka Tanaka, DDS, PhD⁷, Michio Hosaka, MD, PhD^{4,8}, Yusuke Miwa, MD, PhD⁹, Katsuya Fuse, MD¹⁰, Fumitoshi Yoshimine, MD¹¹, Ichiro Mashima, MD^{5,6}, Toshiaki Kikuchi, MD, PhD⁶, Yo Seino, MD, PhD¹², Natsue Shimizu, MD, PhD¹³, Megumi Saito, MA¹⁴, Daisuke Kobayashi, PhD^{14,15}, Saeko Kamoshida, PhD¹⁶, Kohei Koiba, PhD¹⁷, Hiroto Ito, PhD¹⁸, Noriko Numata, PhD¹⁹, Eiji Shimizu, MD, PhD¹⁹.

¹. Niigata Seiryō University Health Service Center, Niigata Seiryō University, Niigata, Japan.

². Graduate School of Education, Tohoku University, Miyagi, Japan.

³. Kitasato University, Kanagawa, Japan.

⁴. Showa University, Tokyo, Japan.

⁵. Kohdo Hospital, Niigata, Japan.

⁶. Department of Respiratory Medicine and Infectious Diseases, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan.

⁷. Niigata University Graduate School of Medicine and Dental Sciences, Niigata, Japan.

⁸. Katsuyama Clinic, Yamanashi, Japan.

⁹. Division of Rheumatology, Department of Medicine, Showa University School of Medicine, Tokyo, Japan.

¹⁰. Uonuma City Hospital Koide Medical Center, Niigata, Japan.

¹¹. Niigata Prefectural Tokamachi Hospital, Niigata, Japan.

¹². Niigata Prefectural Tsugawa Hospital, Niigata, Japan.

¹³. Joetsu General Hospital, Niigata, Japan.

¹⁴. Department of Clinical Psychology, Faculty of Social Welfare, Psychology and Child Development, Niigata Seiryō University, Niigata, Japan.

¹⁵. Graduate School of Clinical Psychology, Niigata Seiryō University, Niigata, Japan.

¹⁶. Health Administration Center, Yamagata University, Yamagata, Japan.

¹⁷. Clinical psychology and school education course, Graduate School of Education, Hokkaido University of Education, Hokkaido, Japan.

¹⁸. Department of Health Policy and Management, Faculty of Medicine, Tohoku Medical and Pharmaceutical University, Miyagi, Japan.

¹⁹. Department of Cognitive Behavioral Physiology, Graduate School of Medicine, Chiba University, Chiba, Japan.

ABSTRACT

Objective: This research examines the diagnostic performance of the Japanese version of the GAD-7 (J-GAD-7) based on analysis of two independent studies. The first assesses the accuracy of the J-GAD-7 in identifying generalized anxiety disorder (GAD) in primary care (Study 1). Next, a meta-analysis of diagnostic accuracy studies in five language versions of the GAD-7 (English, Spanish, Dutch, Turkish, and German) is conducted (Study 2), the results of which are compared with those of the J-GAD-7.

Methods: In Study 1, subjects (n=161; male=55, female=106) returned the self-rating questionnaire (J-GAD-7) to their health care professional within 48 hours and underwent a diagnostic evaluation interview based on the Japanese version of M.I.N.I.-Plus. Most (n=151) also responded to the SF-8, and there were negligible missing data in either set of questionnaires. Of the 161 participants, 63 were diagnosed as suffering GAD. The meta-analysis of diagnostic test accuracy (MA-DTA) conducted in Study 2 focused on identifying GAD at the optimal cut-off value 10 using integration of the eight selected studies.

Results: The optimal cut-off value for the J-GAD-7 (Study 1) was determined as ≥ 10 , which displayed sensitivity of 70.6 %, specificity of 93.5 %, and positive likelihood ratio of 10.9. The area under the ROC curve (AUC) was 0.939, with an SE of 0.019 ($P < 0.0001$). Stratum-specific likelihood ratios (SSLRs) of the J-GAD-7 scores of 0-4, 5-9, 10-14, and 15-21 for GAD were 0.04 (95% CI: 0.01-0.14), 1.17 (95% CI: 0.68-2.01), 10.94 (95% CI: 4.29-27.8), and 10.94 (95% CI: 3.00-39.8), respectively. Findings of the Kruskal-Wallis test showed a significant relationship between the severity of the J-GAD-7 and MCS of the SF-8 ($\chi^2 = 71.73$, $df=3$, $p \leq 0.0001$). In Study 2, sensitivity was 0.751 (95% CI: 0.604-0.856), specificity was 0.816 (95% CI: 0.641-0.916), and AUC was 0.835. The heterogeneity τ^2 was 0.709 (95% CI: 0.064-1.626) for sensitivity and 1.153 (95% CI: 0.264-2.332) for specificity.

Conclusions: The Japanese version of the GAD-7 was validated as an adequate tool for identifying GAD in primary care. Moreover, its performance at the optimal cut-off value 10 was equivalent to that of the MA-DTA for five other language versions.



OPEN ACCESS

PUBLISHED

31 January 2025

CITATION

Muramatsu, K., Maeda, S., et al., 2025. Diagnostic Performance of the Japanese Version of the GAD-7 in Primary Care and Comparing its Accuracy with Other Language Versions. Medical Research Archives, [online] 13(1).

<https://doi.org/10.18103/mra.v13i1.6247>

COPYRIGHT

© 2025 European Society of Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI

<https://doi.org/10.18103/mra.v13i1.6247>

ISSN

2375-1924

Introduction

Anxiety disorders are common mental health disorders that are often seen in the general population and in primary care settings. However, because anxiety disorders often present as physical symptoms, primary care physicians may overlook them, and this can interfere with people's daily and social functioning.^{1,2} It has also been reported that people visit general physical medicine departments rather than mental health care institutions, which increases medical costs by physical medicine care use.

Various estimates of the prevalence rate for general anxiety disorder (GAD) have been reported. Among the general population in the US, the rate was 5.1% in adults who had experienced it anytime within the last 12 months, including currently.³ Higher rates of GAD are observed in primary care patients, up to 8%.⁴ Muramatsu K et al reported that the occurrence of anxiety disorders in Japanese primary care is 8.5%.⁵ According to data from the National Comorbidity Survey Replication (NCS-R), the prevalence of GAD among US adults from 2001 to 2003 was 3.4% for women and 1.9% for men.⁶ Also, Misri et al noted that GAD has an estimated occurrence of 8.5% to 10.5% during pregnancy and 4.4% to 10.8% during the postpartum period.⁷

The US Preventive Services Task Force (USPSTF) commissioned a systematic review to evaluate the benefits and harms of screening for anxiety disorders in asymptomatic adults. The USPSTF concluded with moderate certainty that screening for anxiety disorders in adults, including pregnant and postpartum persons, has a moderate net benefit. In short, for the purpose of providing appropriate care for GAD, it is necessary to evaluate which of the questionnaires for assessing anxiety is most suitable in screening patients. The USPSTF reported that the most researched instruments were the GAD-7 and GAD-2, both of which demonstrate adequate sensitivity and specificity to detect GAD.¹ At the cut-off value 10 the GAD-7 had a pooled sensitivity of 0.79 (95%CI: 0.69 to 0.94) and specificity of 0.89 (95%CI: 0.83 to 0.94) across three studies.^{8,9,10}

Spitzer and Kroenke et al first developed the Patient Health Questionnaire (PHQ) of PRIME-MD. The original PHQ anxiety module focused on two diagnoses: panic disorder and other anxiety disorder.¹¹ In their work, the section of the questionnaire related to other anxiety disorders primarily included criteria for GAD, but its yes-no response format did not permit calculation of a severity score that had proved so useful with the PHQ-9. Hence, in 2006 they created a 7-item self-rating questionnaire for the identification of GAD named the GAD-7, and then conducted a study to validate its performance in US primary care patients.¹⁰ Muramatsu K translated the original GAD-7 from English into Japanese in 2008, after which Kroenke and Spitzer back translated the Japanese version (J-GAD-7) to confirm its accuracy. The following year, Muramatsu K et al presented their results on the validity and usefulness of J-GAD-7.¹²

Considering this background, this paper examines the diagnostic performance of the Japanese version of the GAD-7 based on the analysis of two independent studies.

First, we assess the performance of the J-GAD-7 for the identification of GAD in primary care (Study 1). Second, we conduct a meta-analysis on data derived from eight diagnostic accuracy studies comprising five other language versions of the GAD-7 (Study 2). In combining the results of the two studies, we validate the J-GAD-7 as a useful diagnostic tool in primary care.

Methods

The present research consists of two independent studies. First, a diagnostic performance assessment of the J-GAD-7 in primary care was conducted primarily by a team of primary care physicians. The second study comprises a meta-analysis of the diagnostic accuracy of the GAD-7 in five other language versions (English, Spanish, Dutch, Turkish, and German). It was conducted by a team of psychologists who attempted to compare the diagnostic accuracy results of Study 1 with those of the five language versions incorporated in Study 2.

STUDY 1

Participants

Specifics of the patient sample are outlined hereafter. Participants were enrolled between April and July of 2010 by eight physicians, one psychiatrist, three clinical psychologists, and a mental health social worker in four primary care settings, two general hospitals and one psychiatric hospital. Of the 176 subjects approached, 161 (91.5%) completed the questionnaire (J-GAD-7) with negligible levels of missing data. The mean age of subjects was 41.34 years (SD=14.38). Gender breakdown of the 161 participants was 55 male and 106 female. Additionally, most subjects (n=151) responded to the shortened version of the SF-36 Health Survey Scale (SF-8), and there were no significant missing data.

Procedure

Prior to the research, all interviewers (physicians, psychiatrist, clinical psychologists, and mental health social worker) had received training in administering the M.I.N.I.-Plus.¹³ Patients who were seeing their physician or mental health professional for routine medical appointments were asked to participate. Researchers briefly explained the purpose of the study and obtained written informed consent from all willing patients before having them complete the J-GAD-7. Most patients (>90%) also completed the SF-8.¹⁴ Participants were asked to return the questionnaires to their physician or mental health professional within 48 hours and undergo a diagnostic evaluation interview based on the Japanese version of the M.I.N.I.-Plus.¹⁵ The interviewer was kept blind to the results of the J-GAD-7 screening.

This protocol was approved by the Ethics Committee at the Faculty of Dentistry, Niigata University, and the other participating institutions in accordance with the Ethical Principles for Medical Research Involving Human Subjects (Declaration of Helsinki).

Japanese Version of the GAD-7 (J-GAD-7)

Kroenke and Spitzer et al developed a 7-item self-rating questionnaire for identifying GAD in 2006. They subsequently carried out a study of more than 2700 primary care patients to validate the questionnaire with a similar response set to the PHQ-9 to establish probable

diagnoses of GAD as well as to grade its severity. The GAD-7 scores can range from 0 to 21, with scores of 5, 10 and 15 representing mild, moderate and severe levels of anxiety symptoms. They reported that sensitivity and specificity at the optimal cut-off value 10 were 89.0% and 82.0%, respectively.¹⁰ The J-GAD-7 used in this research was translated by Muramatsu K and verified by Kroenke and Spitzer via back translation.¹²

SF-8 (Short Form 8 Health Survey) Japanese version

The SF-8 is a shortened form of the SF-36 Health Survey Scale that yields comparable eight-dimension health profiles and comparable estimates of summary scores for the physical and mental components of health.^{14,16} The SF-8 consists of the following items: physical function, role limitation due to physical problems, bodily pain, general health, energy/fatigue, vitality, social functioning, role limitation due to emotional problems, and mental health using Likert scales with five or six points. Using the standardized Japanese version of the SF-8, we calculated two summary scores, physical component summary (PCS) and mental component summary (MCS), in accordance with the scoring rules.¹⁴

Psychiatric interviews

The Mini International Neuropsychiatric Interview (MINI)-Plus was selected as the comparison standard for validation of the J-GAD-7.^{13,15} The MINI was originally developed as a brief, structured simple diagnostic interview compatible with DSM-IV-TR¹⁷ and ICD-10 criteria.¹⁸ Its algorithms and question formulations resemble those of the Composite International Diagnostic Interview (CIDI).^{19, 20} The MINI mainly focuses on current diagnoses and specifically explores lifetime diagnoses that are clinically relevant to the present condition. For most diagnostic sections, one or two screening questions are used to rule out a diagnosis when answered in the negative. The MINI includes nineteen disorders that are chosen as the most common disorders from data of epidemiological studies such as the Epidemiologic Catchments Area Study²¹ and the National Comorbidity Survey.²²

The MINI-Plus was designed as a more detailed measure of disorders, including 23 disorders that might reasonably be included in structured interviews in clinical research studies. Modules include those for somatization disorders, mixed anxiety-depressive disorder, and premenstrual dysphoric disorder. To diagnose GAD and at the same time rule out confounding disorders, we used the MINI-Plus.^{13,15}

Statistical analysis

The data from the J-GAD-7 validity study presented at an academic conference in 2009¹² has been supplemented with new data and re-analysed. Statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS 29) program. The MINI-Plus diagnosis of GAD-7 constituted the criterion standard. With respect to criterion validity, we investigated 'sensitivity', 'specificity', and 'likelihood ratio', which were calculated for Table 1.

To diagnose GAD and at the same time rule out confounding disorders, we used the MINI-Plus^{13,15} scores over various intervals. Stratum-specific likelihood ratios (SSLRs) of the GAD-7 were calculated via the spreadsheet program.²³ Receiver operating characteristics (ROC) analyses were conducted along with areas under the curves (AUCs). Construct validity of the J-GAD-7 as a measure of anxiety severity was also assessed.

STUDY 2

Using data extracted from studies included in a systematic review of the GAD-7 by Plummer et al,²⁴ we conducted a meta-analysis of the diagnostic accuracy of five language versions of the GAD-7 (English, Spanish, Dutch, Turkish, and German). Data from eight studies (each in one of the five languages) were selected and combined, and a meta-analysis of diagnostic test accuracy (MA-DTA) using R was performed on the accuracy of identifying anxiety disorders at the cut-off value 10.

Plummer et al assessed twelve studies using QUADAS-2²⁵ and conducted a systematic review and meta-analysis based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)²⁶ study selection process.²⁴ From that review, we extracted eight studies for which the data necessary for meta-analysis could be estimated. The remaining studies were removed as they did not allow for estimation of the data required. The meta-analysis using generalized linear mixed models was performed using R version 4.3.2, Rstudio 2023.12.1+402. package forestploter and altmeta.

Results

STUDY 1

The mean J-GAD-7 score of the 63 patients diagnosed by the physician or MHP as suffering GAD was 7.08 (SD=5.65). Table 1 shows the sensitivity, specificity, and likelihood ratios for the different J-GAD-7 thresholds in diagnosing GAD in the 161 patients. The optimal cut-off value ≥ 10 had sensitivity of 70.6%, specificity of 93.5%, and a positive likelihood ratio of 10.9 (95% CI: 2.1-23.3). Figure 1 displays the area under the ROC curve (AUC) was 0.939 (CI: 0.901-0.975), with an SE of 0.019 ($P < 0.0001$; Fig 1).

Table 2 illustrates the SSLRs of the J-GAD-7 in Japanese primary care. Spitzer et al¹¹ suggest that the GAD-7 scores of 5, 10 and 15 represent mild, moderate and severe, respectively. Using these classifications as a guide, we select four strata of J-GAD-7 scores, 0-4 (minimal), 5-9 (mild), 10-14 (moderate), and 15-21 (severe) for identifying GAD. The SSLRs of the J-GAD-7 scores of 0-4 (minimal), 5-9 (mild), 10-14 (moderate), and 15-21 (severe) for GAD are 0.04 (95% CI: 0.01-0.14), 1.17 (95% CI: 0.68-2.01), 10.94 (95% CI: 4.29-27.8), and 10.94 (95% CI: 3.00-39.8), respectively.

Figure 2 displays the relationship between the J-GAD-7 and the SF-8. Findings of the Kruskal-Wallis test show a significant correlation between the severity of the J-GAD-7 and MCS of the SF-8 ($\chi^2 = 71.73$, $df = 3$, $p \leq 0.0001$).

Table 1. Operating characteristics of various J-GAD-7 cut-off scores for diagnosing GAD

Generalized Anxiety Disorder (n=68)			
J-GAD-7 Score	Sensitivity (%)	Specificity (%)	Likelihood Ratio 95% CI
≥8	85.3	89.2	7.9 4.4-14.1
≥9	79.4	92.5	10.6 5.3-21.2
≥10	70.6	93.5	10.9 2.1-23.3
≥11	63.2	94.6	11.8 5.1-27.0
≥12	55.9	95.7	13.0 5.2-32.8
≥13	41.2	97.8	19.1 5.5-67.1

Figure 1. Receiver operating characteristics (ROC) curve of the J-GAD-7

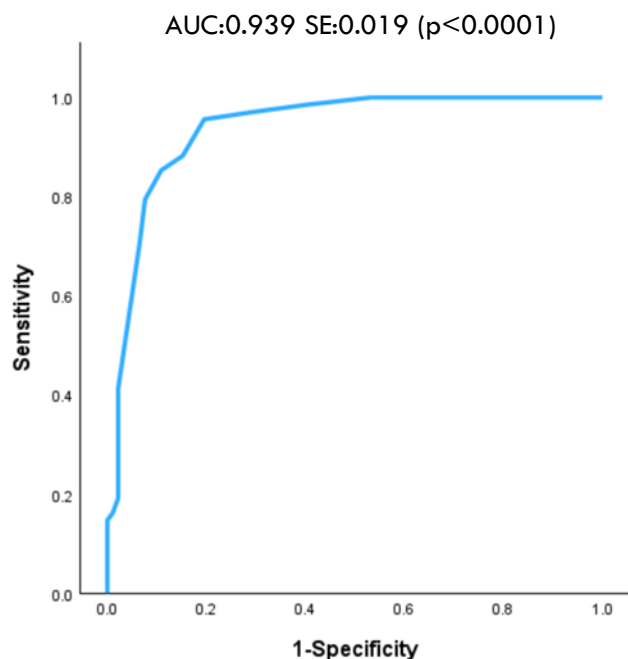
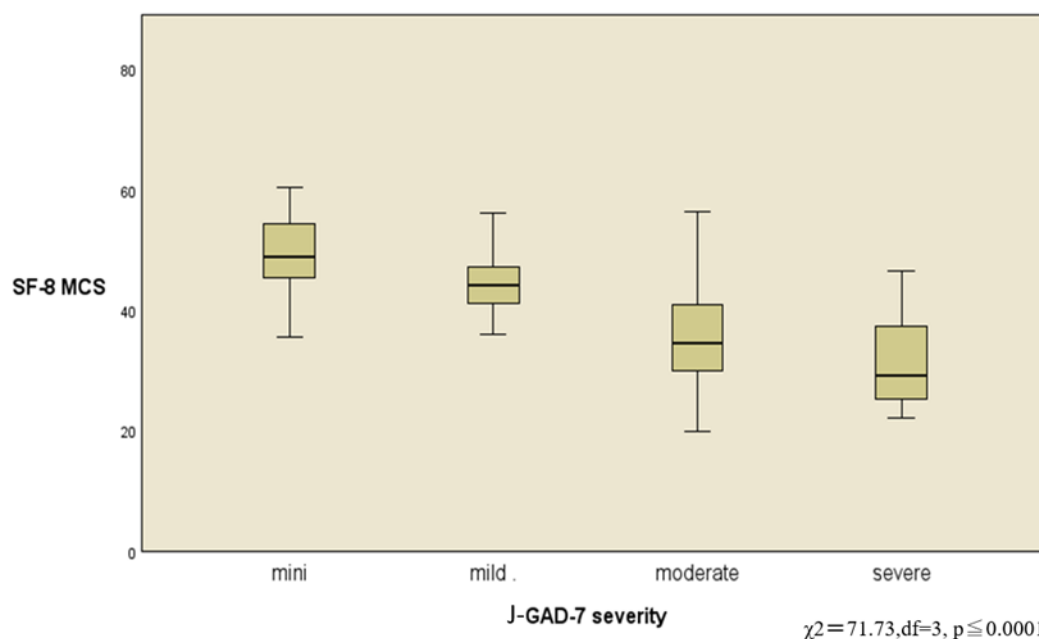


Table 2. Operating characteristics of J-GAD-7 severity scores at different levels using SSLR

GAD-7 Score	MINIGAD (+) n=68	MINIGAD (-) n=93	SSLR 95%CI
Minimal (0-4)	2	66	0.04 0.01-0.14
Mild (5-9)	18	21	1.17 0.68-2.01
Moderate (10-14)	32	4	10.94 4.29-27.8
Severe (15-21)	16	2	10.94 3.00-39.8

Figure 2. Relationship between severity of J-GAD-7 and MCS of SF-8 using the Kruskal-Wallis test



STUDY 2

Table 3 lists the eight diagnostic accuracy studies of the GAD-7 used in the meta-analysis.^{10, 27-34} Figure 3 shows forest plots and integrated estimates of sensitivity and specificity. Table 4 illustrates the diagnostic accuracy of the GAD-7 for identifying anxiety disorders at the cut-off value 10. Sensitivity is 0.751, specificity is 0.816, and the ACU is 0.835 in the generalized linear mixed model diagnostic accuracy meta-analysis.

There is no established method for quantifying the degree of heterogeneity in diagnostic test accuracy meta-analyses.^{26,35,36} We calculate the "heterogeneity t -squared" method, which quantifies the same level of heterogeneity as the diagnostic accuracy meta-analysis of the PHQ-9 by the Depression Screening Data (DEPRESSD) PHQ Group project.³⁵ The heterogeneity τ^2 of sensitivity and specificity is 0.709 (95% CI:0.064-1.626) and 1.153 (95% CI:0.264-2.332), respectively.

The integrated SROC for the anxiety disorder identification system with the cut-off value 10 is displayed in Figure 4. Integrated sensitivity is indicated by the square (■). The estimate false positive rate is obtained by subtracting the specificity from 1. The circle (○) indicates the sensitivity and false positive rate for each study, the straight line indicates hierarchical summary receiver operating characteristic (HSROC), and the dotted line indicates the confidence interval for the estimates.

In summary, for the eight studies carried out in five languages, the identification accuracy of GAD at the cut-off value 10 is verified to have good diagnostic performance based on the AUC results, with estimated values for sensitivity and specificity being sufficiently high. Heterogeneity τ^2 in sensitivity and specificity is observed.

Table 3. List of research on the diagnostic accuracy of GAD-7 (5 language versions) for identifying GAD

Country	Language	Study	Setting	Reference test
Australia	English	Christensen et al. (2011) ²⁷	General population	MINI
UK	English	Delgadillo et al. (2012) ²⁸	Community drugs treatment service	CIS-R
Netherland	Dutch	Donker et al. (2011) ²⁹	General population	CIDI
Spain	Spanish	Garcia-Campayo et al. (2010) ³⁰	Primary care	MINI
USA	English	Kertz et al. (2013) ³¹	Secondary care : psychiatric hospital	MINI
Turkey	Turkish	Konkan et al. (2013) ³²	Secondary care: psychiatric hospital	SCID
USA	English	Spitzer et al. (2006) ¹⁰ & Kroenke et al. (2007) ³³	Primary care	SCID
Germany	German	Wild et al. (2014) ³⁴	General population	SCID

Figure 3. Forest plots of sensitivity and specificity estimates for cut-off value 10 for the GAD-7 among 5 language participants.

(Number of studies = 8 ; Number of participants = 2657; Number with GAD =440)

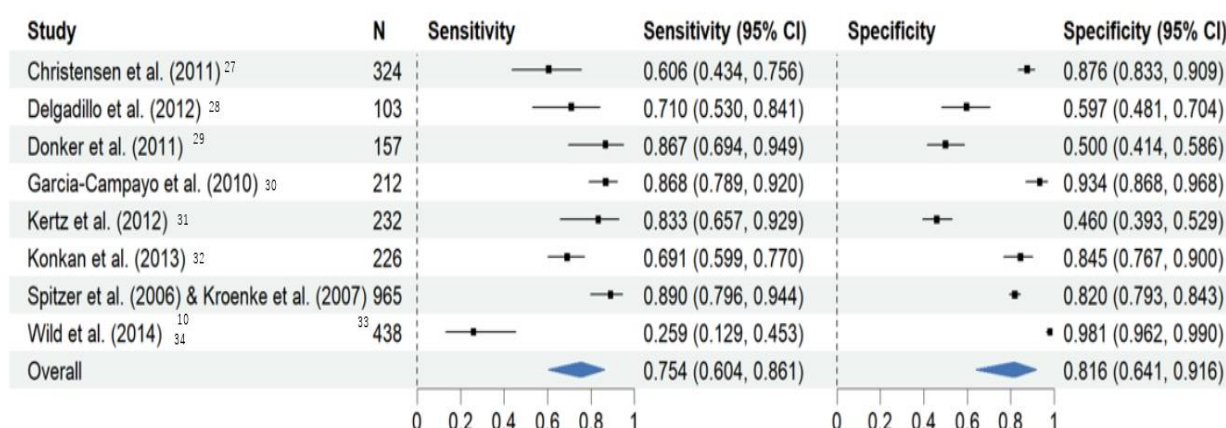
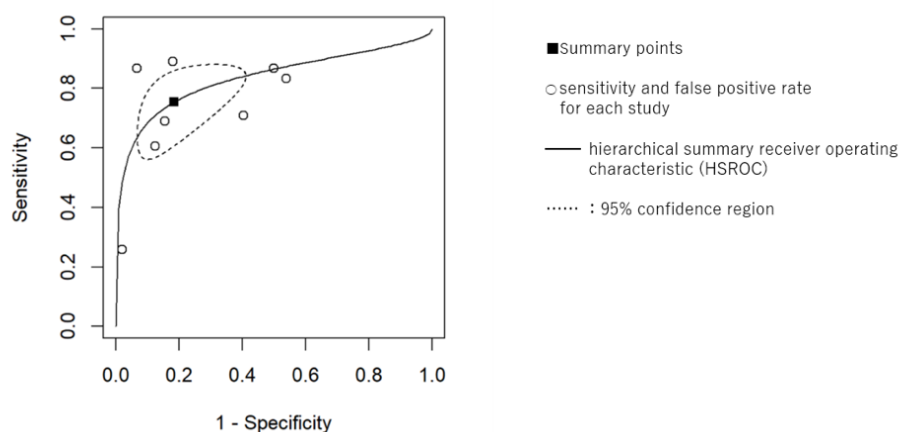


Table 4. Accuracy of the GAD-7 (at cut-off value 10) in identifying GAD: Meta-analysis of diagnostic accuracy using generalised linear mixed models

sensitivity (95%CI)	specificity (95%CI)	AUC	Heterogeneity τ^2	
			sensitivity(95%CI)	specificity (95%CI)
0.751 (0.604 - 0.856)	0.816 (0.641 - 0.916)	0.835	0.709(0.064-1.626)	1.153(0.264-2.332)

Figure 4. GAD-7 for identifying GAD: SROC plot for cut-off value 10

Discussion

This research consists of two studies, with the aim of clarifying the diagnostic accuracy and performance of the J-GAD-7. To compare the diagnostic accuracy of the J-GAD-7 with those of other language versions, Study 2 was included to complement Study 1. It does not in any way constitute a systematic review and/or meta-analysis of the GAD-7. Instead, we merely compare the diagnostic accuracy of MADTA in Study 2 with the results for the J-GAD-7 derived from Study 1.

Our analysis on the utility of the J-GAD-7 were first presented at the 20th World Congress of Psychosomatic Medicine in 2009.¹² Since then, numerous diagnostic accuracy studies encompassing other language versions of the GAD-7 have been reported. There had been no recommendations related to screening for anxiety disorders in the US prior to 2022. That year, however, the USPSTF first issued recommendations on screening for anxiety disorders which emphasized the importance of screening for GAD. In response to this recommendation, some mental health professionals in Japan believed there was a need to validate the diagnostic accuracy of J-GAD-7 as a screening tool for GAD. With the above background in mind, we have decided to publish a report that integrates the results of both studies.

In the diagnostic performance assessment (Study 1), we obtained sensitivity of 70.6% at the optimal cut-off value 10, with specificity of 93.5 %, and a positive likelihood ratio of 10.9 (95%CI:2.1-23.3). Moreover, the area under the curve (AUC) was 0.939 (95% CI:0.901-0.971), with an SE of 0.019 ($P < 0.0001$). These results were deemed acceptable because the J-GAD-7 is defined as a screening tool in Japanese primary care.

Considering the diagnostic performance of the GAD-7 in various clinical settings from previous studies around the world, we applied stratum-specific likelihood ratios (SSLRs) analysis²³ to our J-GAD-7 study. Though our sample is comprised mostly of outpatients of general practitioners and general hospitals, some patients are from private mental clinics. Thus, the sensitivity and specificity are independent of disorder prevalence.

The SSLRs approach was proposed by Furukawa et al,²³ which is advantageous over the dichotomous fixed

threshold approach in that it is independent of the prevalence of target diseases as its case for predictive positive values. Also, SSLRs can indicate multi-level likelihood ratios that provide predictive information as to the probability of GAD in primary care settings.

We calculated SSLRs of the J-GAD-7 using the Furukawa group's spreadsheet program. They are 0.04 (95% CI: 0.01-0.14) when J-GAD-7 scores are between 0 and 4, 1.17 (95% CI: 0.68-2.01) for scores 5 to 9, 10.94 (95% CI: 4.29-27.8) for scores between 10 and 14, and 10.94 (95% CI: 3.00-39.8) for scores ranging from 15 to 21. Interestingly, the strata of 10-14 points and 15-21 points of the J-GAD-7 share the same SSLR of 10.94. Therefore, if the J-GAD-7 score of a respondent is 10 or more, it is most likely that GAD is detectable. When the J-GAD-7 score is 10 points or higher in primary care, general physicians should suspect GAD and request diagnosis of the patients by psychiatrists, who can in turn provide an appropriate treatment.

Muramatsu K et al reported that major depressive disorder is more likely to be comorbid with GAD and other anxiety disorders than with panic disorder in both PRIME-MD and PHQ studies in Japan.⁵ Therefore, general physicians should ask psychiatrists to confirm the coexistence of other psychiatric disorders, such as depression. Moreover, physicians should monitor anxiety symptoms carefully when the score lies in the stratum of 5-9. It may be helpful to take a psychological approach for this patient group. The stratum of 0-4 points has an SSLR of 0.04, thus GAD can be ruled out if the respondent's J-GAD-7 score is 4 or less.

Furthermore, Study 1 clarifies the relationship between the severity of J-GAD-7 and the MCS of the SF-8. There are significant differences in the MCS of the SF-8 among the four groups of the J-GAD-7 scores, termed minimal, mild, moderate, severe anxiety.

The data uncovers a relationship between the severity of anxiety and quality of life in addition to validating the optimal cut-off value 10 of the J-GAD-7 according to severity of anxiety.

Since there are very few examples of evidence-based research on the relationship between anxiety disorders

and QOL, we designed and conducted Study 1 to provide new evidence on the differences in QOL according to the severity of GAD. A score of 10 or more was used to distinguish between the mild and moderate groups. Furthermore, the significant difference in mental QOL scores between the minimal and mild groups, which are below the threshold of 10 points, suggests that mental QOL can be improved for the mild group through psychological approaches.

The second study comprised a meta-analysis on the results of diagnostic accuracy studies in five other GAD-7 language versions (English, Spanish, Dutch, Turkish, and German). The diagnostic accuracy of the GAD-7 for identifying GAD at cut-off value 10 was sufficiently high in both sensitivity and specificity, and the AUC results verified good diagnostic performance for the eight GAD-7 studies selected. Heterogeneity in the results was also observed in the sensitivity and specificity of the R meta-analysis. These results were almost identical to those in a previous report by Plummer et al.²⁴

Importantly, when the J-GAD-7 validation study results are compared with those of the meta-analysis, it can be suggested that the J-GAD-7 performs adequately in identifying GAD at the cut-off value 10.

It must be noted that this research comes with limitations though. A single cross-sectional survey was selected and applied in Study 1 due to financial restraints. Next, despite the variety of medical settings used, sites were not randomly selected throughout Japan. In the future, a more randomized cohort investigation encompassing a larger sample size and area is required to further validate the suitability of the J-GAD-7 in primary care. Besides, as this paper is centred on GAD and its treatment in only primary care, a broader study judging the performance of the J-GAD-7 in primary care, secondary care and the general population would certainly provide a clearer picture of the effects of GAD on contemporary

Japanese society. Finally, although the J-GAD-7 can be considered an effective questionnaire tool for a physician and co-medical staff in the initial assessment of GAD in Japanese primary care, it cannot be used to confirm a clinical diagnosis.

Moreover, although Study 2 was added to this research to play a supplementary role, its inclusion allowed for the comparison of diagnostic accuracies of the J-GAD-7 and other language versions. In the future, a systematic review of the diagnostic accuracy of the GAD-7 needs to be carried out. Such a review should be based on a research protocol in line with PRISMA-DTA (Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy)²⁶ guidelines. Besides, each of the eight GAD-7 diagnostic accuracy studies used in the meta-analysis was published in or prior to 2014. Further research to confirm the performance of GAD-7 as a diagnostic tool requires analysis of diagnostic accuracy studies published after 2014 and use of PRISMA-DTA.²⁶

Conclusion

We validated the performance of the J-GAD-7 for identifying general anxiety disorder in primary care. Furthermore, our meta-analysis utilizing data derived from eight diagnostic accuracy studies outside Japan confirmed that the diagnostic accuracy of the J-GAD-7 in identifying GAD at the optimal cut-off value 10 was equivalent to the MA-DTA of five other language versions of the GAD-7 questionnaire.

Conflict of interest

No competing interests are declared by authors.

Role of funding source

This research was supported by AMED under Grant Number JP23rea522113h.

References

1. US Preventive Services Task Force (2023). Screening for Anxiety Disorders in Adults US Preventive Services Task Force Recommendation Statement. *JAMA*, 329(24):2163-2170. doi:10.1001/jama.2023.9301
2. O'Connor E, Henninger M, Perdue L, Coppola EL, Thomas R, Gaynes BN (2023). Screening for Depression, Anxiety, and Suicide Risk in Adults: A Systematic Evidence Review for the US Preventive Services Task Force. Evidence Synthesis No. 223. Agency for Healthcare Research and Quality, AHRQ publication 22-05295-EF-
3. Harvard Medical School, 2007. National Comorbidity Survey (NCS). (2017, August 21). Retrieved from <https://www.hcp.med.harvard.edu/ncs/index.php>. Data Table 2: 12-month prevalence DSM-IV/WMH-CIDI disorders by sex and cohort
4. Wittchen HU (2002). Generalized anxiety and depression in primary care: prevalence, recognition, and management. *Depression and Anxiety*, 16:162-171. doi:10.1002/da.10065
5. Muramatsu K (2021). Prevalence and Comorbidity of Anxiety and Depressive Disorders in Studies of PRIME-MD and PHQ (Patient Health Questionnaire) in Japan. *Anxiety Disorders - The New Achievements*, Edited by Vladimir V. Kalinin, Cicek Hocaoglu and Shafizan Mohamed. Chapter 7, pp93-106. doi: 10.5772/intechopen.97218
6. Harvard Medical School, 2007. National Comorbidity Survey (NCS). (2017, August 21). Retrieved from <https://www.hcp.med.harvard.edu/ncs/index.php>. Data Table 1: Lifetime prevalence DSM-IV/WMH-CIDI disorders by sex and cohort
7. Misri S, Abizadeh J, Sanders S, Swift E (2015). Perinatal generalized anxiety disorder: assessment and treatment. *J Women's Health (Larchmt)*, 24(9): 762-770. doi:10.1089/jwh.2014.5150
8. Ahn JK, Kim Y, Choi KH. The Psychometric Properties and Clinical Utility of the Korean Version of GAD-7 and GAD-2. *Front Psychiatry*. 2019;10:127. doi:https://dx.doi.org/10.3389/fpsy.2019.00127
9. Kujanpaa T, Ylisaukko-Oja T, Jokelainen J, et al. Prevalence of anxiety disorders among Finnish primary care high utilizers and validation of Finnish translation of GAD-7 and GAD-2 screening tools. *Scand J Prim Health Care*. Jun 2014;32(2):78-83. doi:https://dx.doi.org/10.3109/02813432.2014.920597
10. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. May 22 2006;166(10):1092-7. doi:https://doi.org/10.1001/archinte.166.10.1092
11. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire*. *Jama*. 1999/11/24 ed1999. p. 1737-44
12. Muramatsu K, Muramatsu Y, Miyaoka H et al: Validation and utility of a Japanese version of the GAD-7. *PANMINERVA MEDICA 20th World Congress on Psychosomatic Medicine Abstracts Book*, 51 (Suppl 1 to No 3):79, 2009.
13. Sheehan DV, Lecrubier Y, Sheehan, KH, Sheehan, HK, Amorim P, Janavas J et al (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 59 (suppl.20), 22-33.
14. Fukuhara S, Suzukamo Y. *Manual of the SF-8 Japanese Version*. Institute for Health Outcomes & Process Evaluation Research; Kyoto. 2004.
15. Otsubo T, Tanaka K, Koda R, Shinoda J, Sano N, Tanaka S et al (2005). Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry and Clinical Neuroscience*. 59:517-26.
16. Fukuhara S, Ware JE Jr, Kosinski M, Wada S, Gandek B (1998). Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. *J Clin Epidemiol*. 51:1045-53
17. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV*. 4th Edition (2000). American Psychiatric Association; Washington DC.
18. World Health Organization. *International Classification of Diseases (ICD)(1992)*. Geneva.
19. World Health Organization. *The composite international diagnostic interview (CIDI)*. (1990) Geneva.
20. Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Sheehan HK et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.) - a short diagnostic structured interview: reliability and validity according to the CIDI (1997). *European Psychiatry*, 12: 224-31.
21. Regier DA, Myers JK, Kramer M, Robins LN, Blazer DG, Hough RL et al (1984). The NIMH Epidemiologic Catchment Area Program: historical context, major objectives and study population characteristics. *Archives of General Psychiatry*, 41:934-41.
22. Kessler RC, McGonagle KM, Zhao S, Nelson M, Hughes S, Eshleman H et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the US: results from the National Comorbidity Survey (1994). *Archives of General Psychiatry*, 51:8-20
23. Furukawa TA, Strauss S, Bucher HC, Guyatt G - diagnostic tests. Guyatt G, Drummond R, Meade MO, Cook DJ - editors. *Users' guides to the medical literature: a manual for evidence-based practice* (2008). 2nd edition. New York: The McGraw-Hill Companies, Inc.
24. Plummer F, Manea L, Trepel D, McMillan D (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic meta-analysis. *General Hospital Psychiatry*, 39:24-31. <http://dx.doi.org/10.1016/j.genhosppsych.2015.11.005>
25. Hitting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, 155:529-36.
26. Salameh JP, Bossuyt PM, McGrath TA, et al (2020). Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and

- checklist. *BMJ* 2020;370:m2632.
doi:10.1136/bmj.m2632
27. Christensen H, Batterham PJ, Grant JB, Griffiths KM, Mackinnon AJ. A population study comparing screening performance of prototypes for depression and anxiety with standard scales. *BMC Med Res Methodol* (2011). <http://dx.doi.org/10.1186/1471-2288-11-154>.
28. Delgadillo J, Payne S, Gilbody S, Godfrey C, Gore S, Jessop D, et al (2012). Brief case finding tools for anxiety disorders: validation of GAD-7 and GAD-2 in addictions treatment. *Drug Alcohol Depend*,125:37–42.
29. Donker T, van Straten A, Marks I, Cuijpers P (2011). Quick and easy self-rating of Generalized Anxiety Disorder: validity of the Dutch web-based GAD-7, GAD-2 and GAD-SI. *Psychiatry Res*,188:58–64.
30. Garcia-Campayo J, Zamorano E, Ruiz MA, Pardo A, Perez-Paramo M, Lopez Gomez V, et al (2010). Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. *Health Qual Life Outcomes*.1186/
<http://dx.doi.org/10.1477-7525-8-8>.
31. Kertz S, Bigda-Peyton J, Bjorgvinsson T (2013). Validity of the Generalized Anxiety Disorder- 7 Scale in an acute psychiatric sample. *Clin Psychol Psychother*, 20:456-464.
32. Konkan R, Senormanci O, Guclu O, Aydin E, Sungur MZ (2013). Validity and reliability study for the Turkish adaptation of the Generalized Anxiety Disorder-7 (GAD-7) scale. *Arch Neuropsychiatry*, 50:53–8.
33. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Lowe B (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med*.146:317–25.
34. Wild B, Eckl A, Herzog W, Niehoff D, Lechner S, Maatouk I, et al (2014). Assessing Generalized Anxiety Disorder in elderly people using the GAD-7 and GAD-2 scales: results of a validation study. *Am J Geriatric Psychiatry*, 22:1029–38.
35. Negeri ZF, Levis B, Sun Y, He C, Krishnan A et al (2021). Accuracy of the Patient Health Questionnaire-9 for screening to detect major depression: updated systematic review and individual participant data meta-analysis. *BMJ*, 374: n2183.
<http://dx.doi.org/10.1136/bmj.n2183>
36. Macaskill P, Gatsonis C, Deeks JJ, et al. Analyzing and Presenting Results (2010). In: Deeks JJ, Bossuyt PM, Gatsonis C, eds. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0*. Cochrane Collaboration.