Jurnal Ilmiah Pengabdian Masyarakat Bidang Kesehatan
Volume 4, Issue 3, 363-375
e-ISSN: 2986-2698
https://ejournal.gomit.id/index.php/abdigermas
DOI: 10.58723/abdigermas.v4i3.808

Development and Feasibility Evaluation of a Virtual Reality- and Artificial Intelligence–Based Mental Health Screening System (SMART) for Indonesian Migrant Caregivers in Japan

Andria Pragholapati* Universitas Pendidikan Indonesia INDONESIA
Herniwati Universitas Pendidikan Indonesia, INDONESIA
Dhika Dharmansyah Universitas Pendidikan Indonesia, INDONESIA
Lisda Nurjaleka Waseda University, JAPAN

Article Info

Article history:

Received: July 5, 2026
Revised: August 27, 2026
Accepted: November 9, 2026

Keywords:

Mental Health Screening;
Indonesian Migrant Caregivers;
Virtual Reality;
Artificial Intelligence;
Feasibility.

Abstract

Background: Indonesian migrant caregivers in Japan face high occupational demands, language barriers, and cultural adaptation, conditions associated with psychological distress. Culturally and linguistically appropriate mental health screening remains limited for this partner group, which may delay early identification of distress.
Aim: This program aimed to increase participants' mental health literacy and awareness of early psychological distress, and to assess the acceptability of a Virtual Reality–Artificial Intelligence (VR-AI)–assisted screening and support model.
Methods: A one-group pre-test–post-test community engagement design was used. The program ran from 5 to 12 June 2026 at three sites in Tokyo, the Ueno district, and Shizuoka Prefecture, with 20 Indonesian migrant caregivers recruited purposively. Activities comprised mental health education, VR-AI training, guided PHQ-4–based self-screening, and individual mentoring. Mental health literacy (0–20), technology acceptance (TAM, 1–7), and willingness to seek help (1–5) were measured before and after using Indonesian-language questionnaires and analyzed descriptively. Written informed consent was obtained.
Results: Mental health literacy rose from 9.5 ± 2.4 to 15.2 ± 1.8 (mean difference 5.7), technology acceptance from 3.7 ± 0.8 to 5.9 ± 0.6 (2.2), and willingness to seek professional support from 2.1 ± 0.9 to 3.8 ± 0.8 (1.7). Engagement and attendance also increased. No inferential testing was performed, so these findings are descriptive and indicate feasibility and acceptability rather than effectiveness.
Conclusion: VR-AI technology may be a feasible and acceptable complementary approach to mental health education and early psychological distress screening among Indonesian migrant caregivers in Japan. It does not replace professional clinical assessment, and controlled studies with larger samples are needed.

To cite this article: Pragholapati, A., Herniwati, H., Dharmansyah, D., & Nurjaleka, L. (2026). Development and Feasibility Evaluation of a Virtual Reality- and Artificial Intelligence–Based Mental Health Screening System (SMART) for Indonesian Migrant Caregivers in Japan. *Jurnal Ilmiah Pengabdian Masyarakat Bidang Kesehatan (Abdigermas), 4(3),* 363–375. https://doi.org/10.58723/abdigermas.v4i3.808

This article is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License ©2026 by author/s

INTRODUCTION

Recent migration and employment data indicate a substantial contribution of Indonesian workers to Japan's long-term care sector through bilateral labor cooperation schemes, including the Economic Partnership Agreement and the Specified Skilled Worker program. Within these schemes, Indonesian migrant caregivers work in clinically demanding settings that require advanced care competencies, cross-cultural adaptation, and Japanese language proficiency. Empirical studies of this population report associations between migrant care work and occupational stress, culture shock, anxiety, and psychological distress, although reported prevalence varies across samples and measurement instruments (Nugraha & Ohara-Hirano, 2016; Wu et al., 2022; Igarashi et al., 2024;

*** Corresponding author:**
Andria Pragholapati, Universitas Pendidikan Indonesia, INDONESIA. ✉ andria.pragholapati@upi.edu

Takeda & Fukuzaki, 2024). The evidence base is therefore consistent in direction but imprecise in magnitude, which argues for early screening rather than for confident prevalence claims.

Despite these documented psychosocial challenges, access to culturally and linguistically appropriate mental health screening remains an important concern for this partner group. Existing mental health assessment and service pathways in Japan may present linguistic and cultural barriers for Indonesian migrant workers, particularly when screening instruments and communication approaches do not adequately reflect migrants' cultural and linguistic contexts (Miller et al., 2019; Nagaya et al., 2024; Takubo et al., 2020). Four related but distinct issues should be separated here: the limited availability of culturally responsive screening instruments, language barriers within service encounters, restricted practical accessibility of services, and low utilization of services that are nominally available. This program addresses the first two directly and the remaining two only indirectly. Where these barriers persist, symptoms of psychological distress may remain unrecognized until they affect occupational functioning, interpersonal relationships, or overall well-being. The program was accordingly implemented among Indonesian migrant caregivers working in selected employment areas in Tokyo, including the Ueno district, and in Shizuoka Prefecture, where partner community networks were already established and able to host community-based activities.

The novelty of this program lies in the integration of an immersive VR environment with AI-assisted interaction to support culturally responsive mental health self-screening and individualized preliminary guidance. The two technologies perform separate functions. VR delivers the screening environment: an Indonesian-language immersive setting in which participants complete a PHQ-4–based self-screening while viewing caregiving scenarios drawn from their own work context. This approach is intended to sustain engagement and reduce the sense of being examined. AI handles interaction and interpretation support: it conducts a conversational intake in Indonesian, scores the PHQ-4 responses, and generates a plain-language summary with tiered suggestions, including a referral prompt when responses reach the established cut-off. Unlike conventional paper-based or web-based screening, this configuration was designed to make the assessment interactive and user-centered while lowering the perceived barriers associated with conventional mental health services (Sutori et al., 2025; Torous et al., 2025). Cultural responsiveness was operationalized through Indonesian-language delivery, terminology matched to how participants describe emotional strain, and scenarios built from stressors specific to caregiving work in Japan. The system is referred to here as SMART (Screening for Mental Health with Adaptive and Responsive Technology). It has not undergone formal psychometric validation, and this program constitutes its first field evaluation.

This article maintains a clear distinction between preliminary screening and clinical assessment. The SMART system supports preliminary mental health screening and early identification of psychological distress; it does not perform clinical assessment or diagnosis and does not replace professional care. Where a participant's screening responses indicated possible distress, the mentoring session was used to explain the result in plain language and, where relevant, to discuss referral options.

The program pursued three measurable objectives: to increase participants' knowledge of the early signs of psychological distress, measured with a mental health literacy questionnaire administered before and after the program; to assess the acceptability of the VR-AI screening system, measured with a Technology Acceptance Model questionnaire; and to assess change in willingness to seek professional support, measured on a five-point help-seeking scale. Comfort and openness during the screening process were recorded through facilitator observation and mentoring notes rather than through a validated instrument, and are therefore reported as contextual observations only. The program was designed as a promotive and preventive model with potential for adaptation to other migrant communities facing comparable psychosocial conditions (Sato et al., 2016; Yektiningsih et al., 2021).

METHOD

The issue addressed by this international community engagement program was the psychological burden reported by Indonesian migrant caregivers in Japan in connection with

occupational demands, communication barriers, and socio-cultural adaptation. Many migrant caregivers have limited access to rapid, user-friendly, and stigma-free mental health screening. For the partner group, this gap in access to culturally responsive screening, rather than the stressors themselves, defined the entry point of the program. In the context of transnational labor migration, responsive and culturally appropriate support models are needed to facilitate early identification of psychological distress and timely follow-up (Nugraha & Ohara-Hirano, 2016; Gunawan et al., 2024; Takeda & Fukuzaki, 2024).

Because this activity was a community engagement program rather than a controlled study, its evaluation was guided by the following program evaluation questions:

1. What changes in participants' mental health knowledge were observed following participation in the program?
2. What level of acceptance of the Virtual Reality–Artificial Intelligence (VR-AI) screening system did participants report after using it?
3. To what extent were participants able to recognize early signs of psychological distress and identify the need for follow-up support after the program?

The program was evaluated with a one-group pre-test–post-test design. There was no control or comparison group and no randomization, so the evaluation documents feasibility, acceptability, and observed change rather than intervention effectiveness. Wording throughout this article is chosen to respect that boundary. The program was implemented at three activity sites in Japan: two in the Greater Tokyo area, one of which was located in the Ueno district, and one in Shizuoka Prefecture. Ueno is a district within Tokyo rather than a separate city; it is reported separately here only because it hosted a distinct participant group recruited through a different community network. These locations were selected because they contain substantial concentrations of Indonesian migrant caregivers and because partner organizations there were able to host community-based mental health promotion activities. The geographic distribution is shown in Figure 1.

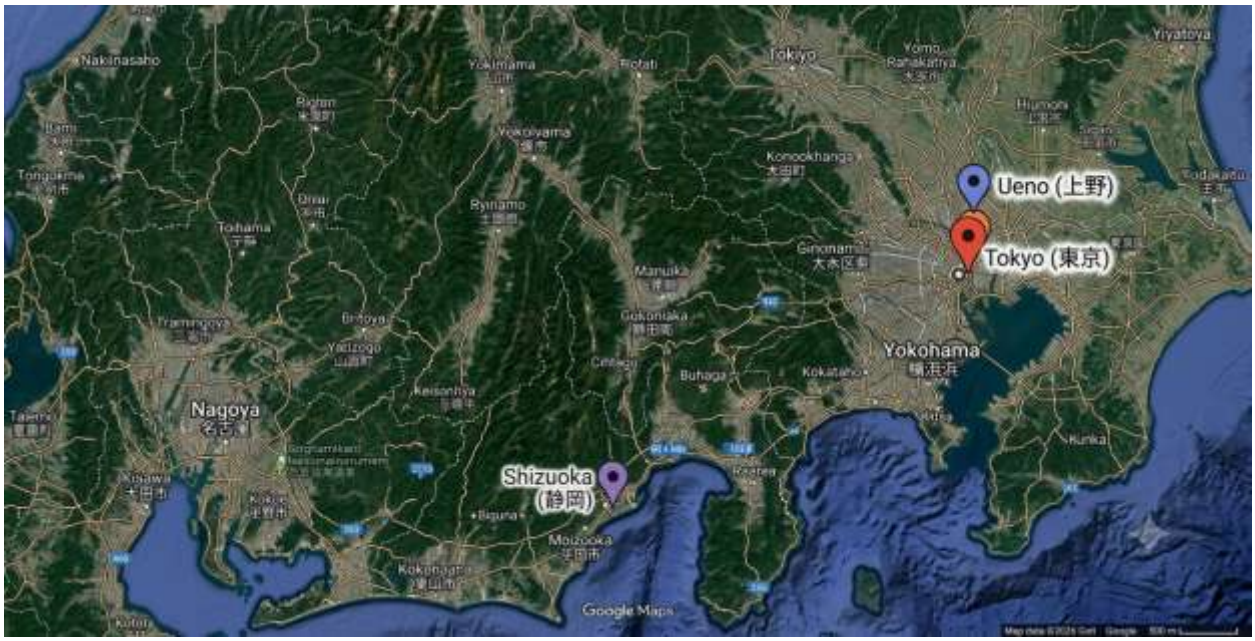


Figure 1. Activity sites of the international community engagement program: the Greater Tokyo area including the Ueno district, and Shizuoka Prefecture, Japan (Source: Google Maps, 2026).

The program employed a participatory approach consisting of health education, hands-on training, simulation, individualized mentoring, and evaluation. Twenty Indonesian migrant caregivers took part, recruited by purposive sampling through partner caregiving facilities and Indonesian community organizations at each site. Eligibility required Indonesian nationality, current employment as a caregiver in Japan, at least three months of work experience in the host

setting, and willingness to complete both the VR-AI session and the evaluation. Participants had a mean age of 31 years (range 27 to 31); 14 were women and 6 were men, and their mean length of employment in Japan was 8 years (range 4 to 8). The program ran from 5 to 12 June 2026 and was delivered as a structured session at each site covering the six phases summarized in Table 1. Each site received approximately 4 contact hours, of which the individual VR-AI screening took about 10 minutes per participant.

Table 1. Implementation Methods

Phase	Activity	Objective	Expected Outcome (Target)
1	Needs assessment	To identify the most common mental health challenges experienced by participants	Priority stressors and screening barriers mapped at all three sites; baseline (pre-test) data collected from 100% of participants.
2	Program orientation	To introduce the objectives, procedures, and expected outcomes of the program	100% of participants understand the program aims, procedures, and data-handling rules and give written informed consent before any screening.
3	VR-AI training	To familiarize participants with the operation of the VR-AI–based mental health assessment system	All participants complete at least one supervised practice run of the VR headset and the Indonesian-language AI interface.
4	Simulation and practical session	To provide hands-on experience using the VR-AI technology	VR-AI screening simulation completed by all participants; session attendance of at least 90%.
5	Individual mentoring	To discuss preliminary assessment results and identify participants' support needs	Every participant receives a plain-language explanation of their PHQ-4 result; all participants at or above the cut-off receive referral information.
6	Pre-test and post-test evaluation	To assess changes in participants' knowledge and acceptance of the VR-AI technology	Mean mental health literacy score at least 30% above baseline; mean TAM score of 5.0 or higher (1–7 scale); higher mean willingness to seek professional support than at baseline.

Note. Expected outcomes are stated as program targets. The values actually achieved are reported in the Results section (Table 2).

Implementation began with coordination meetings involving local community partners, followed by development of the educational materials and preparation of the VR-AI equipment. The activity was then delivered through program orientation, interactive training, VR-based simulation of the screening task, and individualized mentoring. The final stage consisted of post-test evaluation, activity documentation, and reporting.

All three instruments were administered in Indonesian immediately before and after the program. Mental health literacy was measured with a 20-item questionnaire covering recognition of stress symptoms, emotional exhaustion, and knowledge of help-seeking options, scored from 0 to 20, with higher scores indicating greater literacy. Technology acceptance was measured with a Technology Acceptance Model (TAM) questionnaire adapted for digital mental health tools, covering perceived usefulness and perceived ease of use, items rated on a seven-point Likert scale, and the responses were summarized and reported as a total mean score. Willingness to seek professional support was measured with a single five-point Likert item. Psychological distress was screened inside the VR-AI environment using the PHQ-4 (Kroenke et al., 2009), which was used to guide served individual feedback and referral discussion rather than functioning as a program outcome measure. All instruments were translated and back-translated into Indonesian and reviewed by the program team for comprehensibility. Formal psychometric validation in this population has not been carried out, and internal consistency was not computed for this activity; this limits how confidently the scores can be interpreted and is acknowledged in the Limitations.

Analysis was descriptive. Pre-test and post-test scores are reported as means and standard deviations together with the absolute mean difference and the percentage change relative to baseline. No inferential testing was performed, because the evaluation was designed to document feasibility and observed change within a single small group rather than to test a hypothesis. Percentage changes should therefore be read as descriptive summaries and not as evidence that the observed differences are statistically significant or clinically meaningful.

Because the activity involved screening for potentially sensitive psychological information, several safeguards were applied. Participation was voluntary and could be withdrawn at any point without any consequence for employment. Written informed consent was obtained after an explanation of the program's purpose, the nature of the VR-AI screening, data handling, and the limits of what a screening result can indicate. Responses were recorded without personal identifiers and stored on encrypted institutional storage accessible only to the program team. Participants whose PHQ-4 responses reached the established cut-off were informed individually during mentoring and given information on locally available mental health services and Indonesian-language support contacts. The activity was conducted under the community service assignment of Universitas Pendidikan Indonesia (contract No. 89/UN40.F7/AM.02.01/2026), and ethical clearance was obtained from STIKEP PPNI (Approval No. III/120/KEPK222 SLE/STIKEP/PPNI/JABAR/XII/2025).

RESULTS AND DISCUSSION

Results

Scores for participants' mental health knowledge increased between the pre-test and the post-test, particularly on items concerning recognition of stress symptoms, emotional exhaustion, and the value of early screening. Participants engaged actively during the Virtual Reality sessions and reported during mentoring that the immersive screening environment felt more approachable and less intimidating than conventional face-to-face assessment. These qualitative impressions were recorded in facilitator notes and group discussion rather than from a structured qualitative instrument, and are therefore presented as contextual observations rather than as findings (Yektiningsih et al., 2021; Sato et al., 2016).

Table 2. Program outcomes before and after the program (n = 20, descriptive)

Indicator	Instrument / Scale	Pre-test (Mean ± SD)	Post-test (Mean ± SD)	Change (%)	Interpretation
Mental health literacy	Mental health literacy questionnaire (0–20)	9.5 ± 2.4	15.2 ± 1.8	+60.0	The mean literacy score increased from limited to adequate levels.
Acceptance of VR-AI technology	Technology Acceptance Model (TAM) total (1–7)	3.7 ± 0.8	5.9 ± 0.6	+59.5	Acceptance of the VR-AI technology increased to a high level.
Willingness to seek professional support	Likert scale for help-seeking (1–5)	2.1 ± 0.9	3.8 ± 0.8	+81.0	Participants became more willing to seek professional mental health support.
Participant engagement	Engagement score (0–10) / attendance (%)	4.6 ± 1.5 / 72%	8.1 ± 1.2 / 95%	+76.1 / +31.9	Engagement and attendance increased; sessions became more interactive.

Read as absolute change, mental health literacy rose by 5.7 points on a 20-point scale, technology acceptance by 2.2 points on a 7-point scale, willingness to seek professional support by 1.7 points on a 5-point scale, and engagement by 3.5 points on a 10-point scale. Percentage change

is reported relative to a low baseline and therefore looks larger than the underlying point difference. No inferential statistics were computed, so it cannot be determined whether these differences exceed expected measurement variability. The results are descriptive and indicate feasibility and acceptability, not effect size. Measured against the targets in Table 1, the program reached its planned benchmarks. The literacy gain of 60.0% exceeded the 30% target, the post-test TAM mean of 5.9 exceeded the 5.0 benchmark, willingness to seek professional support rose above baseline, attendance reached 95% against a 90% target, and all 20 participants completed both the pre-test and the post-test. Meeting a target set for a single small group shows that the activity ran as intended; it does not show that the change would recur elsewhere. Implementation was also documented photographically across the educational sessions, VR-AI training, simulation, individual mentoring, and participant interaction. These records support the account of participation and delivery reported above.



Figure 2. Educational session on mental health literacy for Indonesian migrant caregivers in Japan
The session was held outdoors in a public park that served as the community gathering point for participants on their day off, an arrangement agreed with the partner community so that attendance would not compete with work shifts. Participants sat close to the facilitator in an open semicircle, which kept the session conversational and allowed questions drawn from their own caregiving work to be raised during the presentation rather than after it.



Figure 3. Participants engaging in the VR-AI–based mental health assessment and simulation
The VR-AI simulation took place in the same outdoor setting. Each participant used the headset individually while a facilitator stood alongside to help with the device and answer questions, and the other participants watched and waited for their turn. Seeing peers try the technology first made the initial exposure a shared activity rather than a private test.



Figure 4. Individual mentoring and discussion following the VR-AI assessment session



Figure 5. Group documentation of the international community engagement program conducted in Japan

Discussion

The results of this community engagement program suggest that pairing psychosocial education with VR-AI–based mental health screening is feasible in this setting and that participants responded positively to the approach, whose literacy, acceptance, and help-seeking readiness scores all moved in the intended direction. That pattern is relevant for Indonesian migrant caregivers in Japan, who commonly report language barriers, heavy workloads, discrimination, and social isolation, conditions linked to elevated anxiety and depressive symptoms (Asis & Carandang, 2020, 2024; Miller et al., 2019; Nagaya et al., 2024; Sato et al., 2016). The discussion that follows separates what was measured from what is inferred: with 20 participants, no control group, and no inferential testing, the observed changes cannot be attributed to the VR-AI component specifically, and the interpretations offered below are proposed as plausible explanations rather than established mechanisms (Munira et al., 2023; Takubo et al., 2020).

Linking the Score Changes to Field Implementation

The scores alone do not explain why they moved, so they need to be read against what happened on site. Three features of delivery plausibly carried most of the change, and each maps onto a different indicator. The first is interactive education. The material was built around situations drawn from caregiving work in Japan, and participants were asked to match warning signs of stress and emotional exhaustion to their own experiences. Knowledge that has already been applied once is easier to recognize when the same content reappears on a questionnaire, which may account for part of the 5.7-point literacy gain. Culturally adapted, user-centered formats have produced similar engagement in Indonesian mental health literacy work, precisely because they invite reflection instead of passive listening (Brooks et al., 2021, 2023).

The second is the staged introduction of the technology. Participants were oriented, then trained, then given a supervised practice run, and only after that used the system for their own screening (Table 1, phases 2 to 4). Acceptance under the Technology Acceptance Model rests heavily on perceived ease of use (Pavlopoulos et al., 2024; Torous et al., 2025), and this sequence works on exactly that perception, the device stops being unfamiliar before it is asked to handle anything sensitive. Watching peers use the headset in the same open setting likely lowered hesitation further. The rise in the TAM mean from 3.7 to 5.9 fits this sequence, although the program did not test it against an alternative format.

The third is individual mentoring, which turned a screening score into a conversation. A PHQ-4 result on its own tells a caregiver little about what to do next. In mentoring, facilitators explained the result in Indonesian, related it to the participant's working situation, and, where needed, named local services and Indonesian-language contacts. This step is the most likely driver of the change in willingness to seek professional support, the indicator with the largest relative increase (+81.0%). Help-seeking among migrant workers is constrained less by missing information than by uncertainty about where to go and concern about consequences at work (Miller et al., 2019; Munira et al., 2023), and mentoring addressed both in a private, one-to-one setting. These explanations remain interpretive. The program did not record which component each participant found most useful, so the relative weight of education, staged training, and mentoring cannot be separated from these data. A short post-session item asking participants to rate each component would allow future evaluations to check these field observations against participants' own accounts.

Observed Change in Mental Health Literacy

The observed improvement in mental health literacy suggests that combining psychosocial mentoring with interactive digital simulation may be a promising approach to mental health education in this population. Because the program used a single-group pre–post design without a comparison condition, the independent contribution of the VR-AI component cannot be determined; part of the change may reflect the mentoring, the group setting, or familiarity with a questionnaire completed twice within a short interval. One possible explanation for the size of the shift is that participants were not only receiving information passively but also working through VR scenarios that rendered abstract ideas about stress symptoms and emotional exhaustion as concrete, personally recognisable situations. User-centered and culturally adapted digital interventions have been reported as more meaningful to underserved populations because they support reflection, relevance, and active learning rather than didactic instruction alone (Brooks et al., 2021, 2022, 2023). On this reading the mentoring component may have reinforced interpretation and discussion while the VR-AI component supported experiential learning, though this decomposition was not tested directly (Bell et al., 2020; Wang et al., 2024).

VR-AI Technology Acceptance

The post-intervention TAM score increased to 5.9 on a seven-point scale, indicating a relatively favorable level of technology acceptance. Interpretation should remain cautious for four reasons: the instrument was an adapted TAM covering perceived usefulness and perceived ease of use rather than a fully specified validated version; its Indonesian translation has not been psychometrically validated in this population; subscale scores were not reported separately; and no established threshold defines what counts as high acceptance on this scale. A plausible reading is that a digital tool feels less confrontational and more self-paced than conventional counselling for a population that may associate face-to-face services with stigma, judgment, or language-related discomfort. Acceptance research consistently identifies perceived usefulness, effort expectancy, and user-centered design as central drivers of adoption, and these factors matter particularly where engagement is voluntary and trust must be established quickly (Pavlopoulos et al., 2024; Torous et al., 2025). The immersive format may also have narrowed the perceived distance between screening and self-exploration, but this mechanism was not measured here (Bell et al., 2020; Blackmore et al., 2024).

Technology Acceptance and Help-Seeking Readiness

The observed increase in willingness to seek professional support indicates improved help-seeking readiness following the program. Two limits on that statement matter. First, actual service utilization was not measured, so the findings cannot establish whether increased willingness translated into help-seeking behaviour; the gap between stated intention and subsequent action is well documented and should not be assumed away. Second, no correlation or regression analysis was performed between technology acceptance and help-seeking scores, so any link between the two remains a hypothesis for future testing rather than a finding of this program. One possible explanation is that the VR-AI screening normalized emotional difficulty and lowered hesitation around disclosure, which would be consistent with evidence that acceptability and perceived usefulness precede engagement with digital mental health services (Munira et al., 2023; Torous et al., 2025). For migrant caregivers working in a linguistically demanding environment, a lowered disclosure threshold could support earlier referral when distress is detected, though this proposition requires direct testing (Miller et al., 2019; Nagaya et al., 2024).

Implications

The findings provide preliminary evidence supporting the feasibility of integrating VR-AI technology into community-based mental health promotion for Indonesian migrant caregivers (Gunawan et al., 2024; Nagaya et al., 2024; Takeda & Fukuzaki, 2024). The approach addresses several challenges this group commonly faces, including language barriers, occupational stress, discrimination, and socio-cultural adaptation, by offering a semi-anonymous, Indonesian-language environment for education and preliminary screening (Asis & Carandang, 2020, 2024; Miller et al., 2019; Nugraha et al., 2017; Sato et al., 2016). Whether the model can be scaled is a separate question that this activity cannot answer. Further research is required to determine scalability, cost-effectiveness, sustainability, and implementation feasibility across settings with different technical resources and staffing (Bell et al., 2024; Torous et al., 2025; Wang et al., 2024). Subject to those conditions, the experience reported here may inform community health initiatives in both host and sending countries.

Research Contribution

This program contributes to the digital mental health literature by documenting the feasibility and acceptability of integrating virtual reality and artificial intelligence within a culturally responsive psychosocial activity for Indonesian migrant caregivers, a population that remains underrepresented in this field (Bell et al., 2024; Hidayat & Muda, 2026; Jecker et al., 2026). Unlike work centered on generic digital screening tools, it provides preliminary context-specific evidence regarding the feasibility and acceptability of combining mental health education with PHQ-4–based psychological screening inside an immersive VR-AI platform tailored to the occupational and migration-related experience of Indonesian caregivers in Japan (Nugraha & Ohara-Hirano, 2016; Nugraha et al., 2017; Su et al., 2022). It also supports the argument that psychological well-being in this group is best understood through the interaction of occupational demands, migration experience, cultural adaptation, and technology acceptance rather than through symptom-based screening alone (Gunawan et al., 2024; Igarashi et al., 2024; Takubo et al., 2020). The value of these observations is as an initial foundation for the design of culturally adaptive digital mental health interventions, not as evidence of effect.

Limitations

Several limitations bound what can be concluded. The absence of a control or comparison group is the principal methodological constraint: with a single-group pre–post design and 20 participants, the observed changes cannot be attributed to the VR-AI component, and no causal claim is defensible (Bell et al., 2024). No inferential testing was performed, so the reported changes are descriptive and their statistical and clinical significance is unknown. The short evaluation window says nothing about durability of any change in literacy, distress, or help-seeking, a weakness shared with much of the VR-AI mental health evidence base (Blackmore et al., 2024; Torous et al., 2025; Wang et al., 2024). Reliance on self-report, including the PHQ-4 and the acceptance measures, introduces reporting bias (Kroenke et al., 2009).

Four further sources of bias deserve explicit mention. Selection bias is likely: participants volunteered through partner networks and may already have held favorable attitudes toward mental health or technology, which would inflate both literacy gains and acceptance scores. The Hawthorne effect is plausible because participants knew they were taking part in a program being evaluated. Social desirability bias is a particular concern here, since mental health and help-seeking are sensitive topics in migrant working communities where disclosure may be perceived as carrying employment risk. Measurement reactivity is also possible, as the same questionnaires were administered twice within a short interval, so familiarity with the items may account for part of the post-test improvement.

The outcome set itself is a limitation. The strongest results concern literacy, acceptance, willingness, and engagement, none of which is a clinical outcome. Although the PHQ-4 was administered inside the screening environment, change in psychological distress was not analyzed as a program outcome, so this activity provides no evidence that the intervention altered participants' actual distress. Cultural adaptation and psychometric validation of digital mental health tools for Indonesian migrant populations also remain limited, which affects the precision and transferability of the scores reported here (Brooks et al., 2023; Gondek & Bernardi, 2023). Ethical questions surrounding AI-assisted screening, including privacy, transparency, accountability, and algorithmic bias, require attention before wider implementation (Alhuwaydi, 2026; Jecker et al., 2026). Finally, the activity ran within a specific institutional and community context supported by dedicated technological and human resources, so the experience may not transfer to other migrant caregiver populations or implementation settings without adaptation (Miller et al., 2019; Takubo et al., 2020).

Suggestions

Future work should begin with the analytical gap in this activity: pre–post data should be tested with paired methods, such as the paired t-test or the Wilcoxon signed-rank test depending on distributional assumptions, and reported with mean differences, 95% confidence intervals, and effect sizes. Larger and more diverse samples across regions, occupational settings, and host countries are needed to improve external validity (Bell et al., 2024; Su et al., 2022). Longitudinal studies using randomised or quasi-experimental designs with an active comparison condition are required to separate the contribution of VR-AI from that of mentoring and to evaluate durability and clinical impact (Blackmore et al., 2024; Gunawan et al., 2024; Wang et al., 2024). Instrument development should include formal psychometric validation of the Indonesian versions used here, with reliability and construct validity reported. Participatory co-design involving Indonesian migrant caregivers throughout development would strengthen cultural relevance, linguistic appropriateness, and usability (Brooks et al., 2021, 2023; Nagaya et al., 2024). Further platform work should add multilingual support, adaptive features, and links to occupational health services and migrant support networks, while parallel efforts should establish ethical frameworks covering privacy, informed consent, fairness, transparency, and responsible AI governance (Alhuwaydi, 2026; Jecker et al., 2026; Pavlopoulos et al., 2024). Collaboration among policymakers, employers, healthcare professionals, and community organizations will determine whether such services can be sustained equitably.

CONCLUSION

This international community engagement program showed that a VR-AI–assisted approach to mental health education and screening is feasible and acceptable to Indonesian migrant caregivers in Japan. Implemented in Tokyo, including the Ueno district, and in Shizuoka Prefecture from 5 to 12 June 2026, the activity was associated with higher post-test scores for mental health literacy (9.5 to 15.2 on a 20-point scale), technology acceptance (3.7 to 5.9 on a 7-point scale), and willingness to seek professional support (2.1 to 3.8 on a 5-point scale). These are descriptive changes in a single group of 20 participants without a comparison condition or inferential testing, and they should not be read as evidence that the VR-AI component caused the improvement.

What the activity does support is narrower and still useful: immersive digital technologies can complement community-based mental health promotion for migrant workers, and participants were willing to use them. By combining interactive learning, guided self-screening, and individual mentoring, the SMART model offers an accessible and culturally responsive entry point to psychological support for Indonesian migrant caregivers abroad. It complements professional mental health services and does not replace them. Future initiatives should involve larger and more diverse participant groups, apply controlled designs with inferential analysis and effect size reporting, and follow participants long enough to observe whether changes in literacy and help-seeking readiness persist or translate into service use. That evidence is what would justify scaling the model into routine digital mental health services for migrant workers.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Universitas Pendidikan Indonesia. This community service program is funded by the Universitas Pendidikan Indonesia through the community service program (contract No. 89/UN40.F7/AM.02.01/2026). This support was instrumental in the successful implementation of the program.

AUTHOR CONTRIBUTION STATEMENT

AP conceptualized the program, designed the methodology, coordinated project implementation, analyzed and interpreted the data, and prepared the original manuscript. HW contributed to program implementation, participant coordination, data collection, and manuscript review. DD assisted with the development and implementation of the VR-AI intervention, data management, and critical revision of the manuscript. LN contributed to international collaboration, participant recruitment in Japan, project supervision, and critical review and editing of the manuscript. All authors read and approved the final version of the manuscript.

AI DISCLOSURE STATEMENT

The authors used ChatGPT (OpenAI) during the preparation of this manuscript to assist with English language editing, grammar refinement, and improvement of academic writing style. All AI-generated suggestions were carefully reviewed, verified, and substantially edited by the authors. The authors take full responsibility for the accuracy, integrity, and originality of the content presented in this publication.

CONFLICTS OF INTEREST

The authors declare that there are no financial, institutional, professional, or personal conflicts of interest that could have influenced the design, implementation, analysis, interpretation, or publication of this study.

REFERENCES

Alhuwaydi, A. M. (2026). Artificial intelligence-integrated virtual reality in mental health care: A scoping review of evidence, clinical applications, and future directions. *Journal of Clinical Medicine*. 15(11), 3993. https://doi.org/10.3390/jcm15113993

Asis, E., & Carandang, R. R. (2020). The plight of migrant care workers in Japan: A qualitative study of their stressors on caregiving. *Journal of Migration and Health, 1–2*, 100001. https://doi.org/10.1016/j.jmh.2020.100001

Asis, E., & Carandang, R. R. (2024). Racial microaggressions and positive coping strategies among migrant elderly care workers in Japan. *Social Sciences and Humanities Open*. https://doi.org/10.1016/j.ssaho.2024.100812

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience*. 22(2), 169–177. https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia

Bell, I. H., Pot-Kolder, R., Rizzo, A., et al. (2024). Advances in the use of virtual reality to treat mental health conditions. *Nature Reviews Psychology, 3*, 552–567. https://doi.org/10.1038/s44159-024-00334-9

Best, P., Meireles, M., Schroeder, F., Montgomery, L., Maddock, A., Davidson, G., Galway, K., Trainor, D., Campbell, A., & Van Daele, T. (2022). Freely available virtual reality experiences as tools to support mental health therapy: A systematic scoping review and consensus-based interdisciplinary analysis. *Journal of Technology in Behavioral Science, 7*(1), 100–114. https://doi.org/10.1007/s41347-021-00214-6

Blackmore, R., Giles, C., Tremain, H., & Thomas, N. (2024). Examining the use of virtual reality to support mindfulness skills practice in mood and anxiety disorders: Mixed methods study. *Journal of Medical Internet Research*. https://doi.org/10.2196/45640

Brooks, H., Irmansyah, I., Syarif, A. K., & Bee, P. (2023). Evaluating a prototype digital mental health literacy intervention for children and young people aged 11–15 in Java, Indonesia: A mixed methods, multi-site case study evaluation. *Child and Adolescent Psychiatry and Mental Health*. 17(1), 79. https://doi.org/10.1186/s13034-023-00608-9

Brooks, H., Prawira, B., Windfuhr, K., & Bee, P. (2022). Mental health literacy amongst children with common mental health problems and their parents in Java, Indonesia: A qualitative study. *Global Mental Health*. 9, 72–83. https://doi.org/10.1017/gmh.2022.5

Brooks, H., Syarif, A. K., Pedley, R., & Bee, P. (2021). Improving mental health literacy among young people aged 11–15 years in Java, Indonesia: The co-development of a culturally-appropriate, user-centred resource (The IMPeTUs intervention). *Child and Adolescent Psychiatry and Mental Health*. 15(1), 56. https://doi.org/10.1186/s13034-021-00410-5

Goldsmith, E. S., Zerzan, N., Greer, N., et al. (2025). *Virtual reality for treating mental health disorders and prevention of suicide: A systematic review*. Department of Veterans Affairs (US). https://www.ncbi.nlm.nih.gov/books/NBK622013/

Gondek, D., & Bernardi, L. (2023). Mental health and wellbeing of population with migrant background in Switzerland: A scoping review and evidence map of quantitative evidence. *Journal of Immigrant and Minority Health*. 25(5), 1108-1117. https://doi.org/10.1007/s10903-023-01490-5

Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health: A review of the literature. *Social Psychiatry and Psychiatric Epidemiology, 42*(5), 343–354. https://doi.org/10.1007/s00127-007-0173-4

Gunawan, J., Efendi, F., Tsujita, Y., & Oda, H. (2024). Caring across borders: Indonesian care workers' experiences of working in Japan. *International Journal of Migration, Health and Social Care, 20*(2), 181–198. https://doi.org/10.1108/IJMHSC-05-2023-0043

Hatta, M. H., Sidi, H., Siew Koon, C., Che Roos, N. A., Sharip, S., Abdul Samad, F. D., Wan Xi, O., Das, S., & Mohamed Saini, S. (2022). Virtual reality (VR) technology for treatment of mental health problems during COVID-19: A systematic review. *International Journal of Environmental Research and Public Health, 19*(9), 5389. https://doi.org/10.3390/ijerph19095389

Hidayat, E. Y., & Muda, A. K. (2026). From internal accuracy to clinical readiness: A systematic review of artificial intelligence-based mental health studies in Southeast Asia. *International Journal of Medical Informatics*. https://doi.org/10.1016/j.ijmedinf.2026.106558

Igarashi, N. H., Kiriya, J., Carandang, R. R., & Jimba, M. (2024). Securing human resources for Japan's aging population: A mixed methods study of job satisfaction and well-being among Southeast Asian long-term care workers in Japan. *BMC Health Services Research*. 24(1), 11-34. https://doi.org/10.1186/s12913-024-11572-1

Jecker, N. S., Ozeki-Hayashi, R., Nakazawa, E., & Fujita, T. (2026). Designing chatbots and social robots for mental health to be culturally competent and people-centered. *Philosophy and Technology*. 39. https://doi.org/10.1007/s13347-026-01073-w

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics*, *50*(6), 613–621. https://doi.org/10.1176/appi.psy.50.6.613

Miller, R., Tomita, Y., Ong, K. I. C., & Jimba, M. (2019). Mental well-being of international migrants to Japan: A systematic review. *BMJ Open*. https://doi.org/10.1136/bmjopen-2019-029988

Munira, L., Liamputtong, P., & Viwattanakulvanid, P. (2023). Barriers and facilitators to access mental health services among people with mental disorders in Indonesia: A qualitative study. *Belitung Nursing Journal*, 9(2), 110–11. https://doi.org/10.33546/bnj.2521

Nagaya, Y., Gillin, N., & Smith, D. (2024). South-East Asian nurses' experiences under the Economic Partnership Agreement (EPA) in Japan: How language ability affects self-confidence and interpersonal relationships. *Journal of Transcultural Nursing*, 35(6), 465-474. https://doi.org/10.1177/10436596241271133

Nugraha, S., & Ohara-Hirano, Y. (2016). The mental health predictors of Indonesian healthcare workers in Japan. *Japanese Psychological Research, 58*(1), 85–96. https://doi.org/10.1111/jpr.12100

Nugraha, S., Hirano, Y., Sumihisa, H., et al. (2017). The change in mental health status of Indonesian health care migrant workers in Japan. *Kesmas: Jurnal Kesehatan Masyarakat Nasional, 12*(2), 53–59. https://doi.org/10.21109/kesmas.v12i2.1698

Pavlopoulos, A., Rachiotis, T., & Maglogiannis, I. (2024). An overview of tools and technologies for anxiety and depression management using AI. *Applied Sciences*. 14(19), 9068. https://doi.org/10.3390/app14199068

Sato, F., Hayakawa, K., & Kamide, K. (2016). Investigation of mental health in Indonesian health workers immigrating to Japan under the Economic Partnership Agreement. *Nursing & Health Sciences, 18*(3), 342–349. https://doi.org/10.1111/nhs.12275

Shinohara, A., Kuwano, N., Shinagawa, Y., & Ohnishi, M. (2025). Development of a health support competence scale for supervisors of female migrant workers in Japan. *BMC Public Health*, 25(1), 2688. https://doi.org/10.1186/s12889-025-23837-3

Su, H.-C., Hu, S. H., Chi, M.-J., Lin, Y.-K., Wang, C.-Y., Nguyen, T. V., & Chuang, Y.-H. (2022). Prevalence of depressive symptoms and associated factors among foreign caregivers: A cross-sectional study. *Nursing Open*. 10(3):1693-1703. https://doi.org/10.1002/nop2.1424

Sutori, S., Eliasson, E. T., Mura, F., Ortiz, V., Catrambone, V., Hadlaczky, G., Todorov, I., Alfeo, A. L., Cardi, V., Cimino, M. G. C. A., Mioni, G., Raya, M. A., Valenza, G., Carli, V., & Gentili, C. (2025). Acceptability, usability, and insights into cybersickness levels of a novel virtual reality environment for the evaluation of depressive symptoms: Exploratory observational study. *JMIR Formative Research, 9*, e68132. https://doi.org/10.2196/68132

Takeda, S., & Fukuzaki, T. (2024). Association between workplace interpersonal relationships and psychological distress among care workers at elder care facilities. *Psychogeriatrics, 24*(4), 847–853. https://doi.org/10.1111/psyg.13130

Takubo, Y., Nemoto, T., Iwai, M., & Mizuno, M. (2020). Demographic and clinical characteristics of foreign nationals accessing psychiatric services in Japan: A multicentre study in a metropolitan area. *BMC Psychiatry*, 20(1), 569. https://doi.org/10.1186/s12888-020-02951-z

Torous, J., Linardon, J., Goldberg, S. B., & Firth, J. (2025). The evolving field of digital mental health: Current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. *World Psychiatry*. 24(2), 156–174. https://doi.org/10.1002/wps.21299

Wang, P., Ai, X., Zhang, X., Ma, F., Zhuang, Y., & Wang, S. (2024). Evaluating virtual reality technology in psychotherapy: Impacts on anxiety, depression, and ADHD. *Frontiers in Psychiatry, 15*, 1480788. https://doi.org/10.3389/fpsyt.2024.1480788

Wu, Q., Yamaguchi, Y., & Greiner, C. (2022). Factors related to mental health of foreign care workers in long-term care facilities in Japan during the COVID-19 pandemic: A comparative study. *International Journal of Environmental Research and Public Health, 19*(24), 16491. https://doi.org/10.3390/ijerph192416491

Yektiningsih, E., Risnasari, N., & Tri Wijayanti, E. (2021). Association between personality traits toward culture shock among Indonesian caregivers in Japan under the Economic Partnership Agreement. *Journal of Nursing Practice, 5*(1), 146–154. https://doi.org/10.30994/jnp.v5i1.165