Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 09 March 2026
Revised 12 April 2026
Accepted 25 May 2026

Natural Resources for Human Health 2026, Vol. 6 (5): 548–560

KEYWORDS:
Disaster reproductive health; Minimum Initial Service Package (MISP); reproductive health kits; midwifery education; focus group discussion; disaster preparedness; quasi-experimental study; Aceh; Indonesia.

Comparative Effectiveness of Lecture, Focus Group Discussion, and Seminar Methods for Improving Midwifery Students' Knowledge of Disaster Reproductive Health Kit Preparation: A Three-Group Quasi-Experimental Study in Aceh, Indonesia

Yushida[1]*, Rina Julianti[2], Satrinawati[3], Isnaini Putri[4]

[1]Midwifery Study Program of Meulaboh, Department of Midwifery, Poltekkes Kemenkes Aceh, Aceh Barat, Indonesia
[2]Midwifery Study Program of Meulaboh, Department of Midwifery, Poltekkes Kemenkes Aceh, Aceh Barat, Indonesia
[3]Midwifery Study Program of Meulaboh, Department of Midwifery, Poltekkes Kemenkes Aceh, Aceh Barat, Indonesia
[4]Midwifery Study Program of Banda Aceh, Department of Midwifery, Poltekkes Kemenkes Aceh, Indonesia

ABSTRACT:

Background: Indonesia is highly exposed to natural hazards, and Aceh Province has repeatedly experienced earthquakes, tsunamis, floods, and landslides. Reproductive health services are frequently disrupted during crises even as needs increase. The Minimum Initial Service Package requires Disaster Reproductive Health Kits to be ready pre-disaster, and midwives are the frontline personnel responsible for their preparation, yet evidence comparing teaching methods for building this competence remains scarce.

Objective: To compare the effectiveness of lecture, focus group discussion (FGD), and seminar methods on midwifery students' knowledge of disaster reproductive health kit preparation.

Methods: A three-group pretest–posttest quasi-experimental study was conducted among midwifery students at the Meulaboh Midwifery Study Program, Poltekkes Kemenkes Aceh, Indonesia, in 2024. Of 105 eligible students, 45 were allocated at the class level to three parallel, dose- and duration-matched 120-minute interventions (lecture, FGD, or seminar; n = 15 each). Knowledge was measured with a 20-item multiple-choice questionnaire pre- and post-intervention. Within-group change was analyzed with paired t tests; between-group differences with one-way ANOVA, Bonferroni post hoc tests, ANCOVA, and logistic regression, with effect sizes reported throughout.

Results: Knowledge increased significantly in all groups (all $P < .001$), with the largest gain in FGD ($\Delta = 33.33$; 95% CI, 29.06–37.61; $d = 4.32$), followed by lecture ($\Delta = 18.00$; $d = 2.48$) and seminar ($\Delta = 15.67$; $d = 1.47$); only FGD's posttest mean reached the "good" category (≥ 75). Posttest scores differed significantly among groups ($F(2,42) = 24.41$; $P < .001$; $\eta^2 = 0.538$), with FGD exceeding lecture and seminar (both $P < .001$), while lecture and seminar did not differ ($P = 1.000$). The FGD advantage persisted after adjustment for baseline score ($P < .001$) and corresponded to higher odds of achieving

Natural Resources for Human Health
eISSN: 2583-1194

"good" knowledge (AOR = 26.00; 95% CI, 2.55–265.12).

Conclusion: With content and duration held constant, FGD produced markedly greater knowledge gains than lecture or seminar. These findings support prioritizing structured small-group discussion in disaster-preparedness curricula in hazard-prone settings, pending confirmation in larger randomized trials

1. INTRODUCTION

Disasters constitute a major structural determinant of health in Indonesia, where geological, climatic, environmental, and anthropogenic factors interact to produce recurrent exposure to multiple hazards. Located at the convergence of three major tectonic plates and within the Pacific Ring of Fire, Indonesia is highly susceptible to earthquakes, tsunamis, volcanic eruptions, floods, landslides, droughts, and forest and land fires[1,2]. Hydrometeorological hazards, particularly floods, landslides, and extreme weather events, account for the majority of annual disasters and affect substantial proportions of the population[1–4] These events generate health consequences extending beyond mortality and injury, including displacement, disease transmission, nutritional insecurity, psychological distress, and disruption of essential health services[3,4]. Aceh Province represents a particularly important context for disaster preparedness. The 2004 Indian Ocean earthquake and tsunami profoundly affected Aceh's population, health infrastructure, and disaster preparedness systems.[5] Nevertheless, vulnerability has remained persistent because recurrent hydrometeorological disasters continue to affect communities. In 2021, flooding in Aceh Tenggara affected 1,804 households (6,571 people), while in early 2022 flooding was reported across 31 sub-districts in seven regencies/cities.[5,6] Aceh Barat, the setting of the present study, has similarly experienced recurrent flooding affecting urban and rural populations and disrupting essential services[6,7].

The consequences of disasters are particularly significant for reproductive health. Pregnancy, childbirth, menstruation, and other reproductive health needs continue during emergencies, whereas access to appropriate services may be severely constrained. National guidance estimates that approximately 4% of disaster-affected populations are pregnant women, with 15–20% potentially experiencing pregnancy or childbirth complications[8,9]. Displacement and weakened social protection may additionally increase risks of sexual and gender-based violence and sexually transmitted infections, including HIV.[10,11] Thus, maintaining reproductive health services during the initial emergency phase is essential for preventing avoidable maternal and neonatal morbidity and mortality[12,13]. The Minimum Initial Service Package (MISP), implemented in Indonesia as the *Paket Pelayanan Awal Minimum (PPAM) Kesehatan Reproduksi pada Krisis Kesehatan*, constitutes a standardized framework for ensuring the continuity of essential reproductive health services during humanitarian emergencies.[14,15] The framework prioritizes the prevention and management of sexual violence, prevention of sexually transmitted infections (STIs) and HIV transmission, prevention of maternal and neonatal morbidity and mortality, coordination of reproductive health responses, and preparation for the transition toward comprehensive reproductive health services.[16–18].

The effectiveness of these interventions depends substantially on the timely availability, appropriate allocation, and effective management of Inter-Agency Reproductive Health Kits. Midwives occupy a strategically important position within this emergency response system because of their extensive distribution across primary healthcare settings and their central responsibilities in maternal, neonatal, and reproductive health services. Indonesian Law No. 4 of 2019 on Midwifery, together with international competency frameworks for midwifery practice, recognizes disaster preparedness and emergency response as relevant dimensions of professional competency.[19,20] Strengthening disaster-related reproductive health competencies during pre-service education is therefore particularly important, especially in geographically and environmentally disaster-prone regions where midwives may constitute among the first available healthcare providers during an emergency. Despite this strategic role, substantial gaps in preparedness may persist among students entering the midwifery profession. A preliminary assessment involving 10 students from the Meulaboh Midwifery Study Program found that six students (60%) demonstrated limited understanding of reproductive health kit preparation, particularly with regard to the classification and intended utilization of the different kit blocks. This preliminary finding is consistent with previous evidence documenting persistent deficiencies in knowledge, preparedness, and implementation of the MISP within health-crisis settings in Indonesia.

These gaps highlight the need to identify pedagogical approaches capable of developing not only factual knowledge but also students' capacity to interpret, apply, and operationalize reproductive health preparedness principles in emergency contexts.

Lecture, seminar, and focus group discussion (FGD) represent three widely used instructional approaches, but they differ substantially in their underlying mechanisms of learning. Lectures are efficient for delivering standardized and systematically organized information to large groups; however, their predominantly one-directional structure may constrain opportunities for active cognitive processing and immediate feedback. Seminars provide greater opportunities for interaction with instructors and peers, although the depth of student engagement may vary according to individual participation and group dynamics. In contrast, FGD is inherently interactive and requires learners to articulate their understanding, examine alternative perspectives, evaluate information, and collaboratively negotiate meaning. These differences are theoretically consistent with the ICAP framework, which proposes that interactive engagement generates deeper learning than passive exposure to information because learners are required to co-construct knowledge through dialogue and cognitive elaboration[21,22]. Similarly, sociocultural and experiential learning perspectives emphasize the importance of social interaction, reflection, contextualized experience, and collaborative knowledge construction in promoting meaningful and transferable learning[23–25].

Empirical evidence from Indonesia has demonstrated the potential of FGD-based learning to improve disaster preparedness and educational achievement. Nevertheless, evidence directly comparing lecture, seminar, and FGD as instructional strategies for disaster-related reproductive health education remains limited[26,27]. In particular, there is a paucity of comparative evidence examining these three approaches under comparable instructional exposure and focusing specifically on students' ability to understand and prepare reproductive health kits within the MISP framework. This represents an important educational and practice-oriented evidence gap because the effectiveness of instructional methods may influence the extent to which future midwives are prepared to translate reproductive health knowledge into appropriate emergency response actions. Therefore, this study aimed to compare the effectiveness of lecture, FGD, and seminar methods in improving midwifery students' knowledge regarding disaster reproductive health kit preparation. Specifically, the study examined changes in knowledge from pre-intervention to post-intervention, evaluated within-group improvements, compared post-intervention outcomes across instructional groups, and identified the instructional approach associated with the greatest likelihood of achieving a good level of knowledge. By addressing these objectives, the study seeks to provide empirical evidence for strengthening disaster-oriented reproductive health education within midwifery curricula, particularly in settings where preparedness for reproductive health emergencies is an essential component of health-system resilience.

MATERIALS AND METHODS

Study Design

This study employed a quasi-experimental three-group pretest–posttest design without a passive control group to compare the relative effectiveness of three instructional modalities—lecture, focus group discussion (FGD), and seminar—in improving students' knowledge of disaster reproductive health kit preparation. The design was selected because the primary research question concerned the comparative effectiveness of alternative educational strategies rather than the effect of education versus no intervention. Accordingly, all three groups received an active intervention and served as comparators for one another. Individual-level randomization was not feasible because the intervention was delivered within pre-existing class structures; therefore, the study was classified as quasi-experimental. Reporting was guided by the *Transparent Reporting of Evaluations with Nonrandomized Designs* (TREND) statement to enhance methodological transparency and reproducibility.

Study Setting and Participants

The study was conducted at the Meulaboh Midwifery Study Program, Department of Midwifery, Poltekkes Kemenkes Aceh, Aceh Barat Regency, Aceh Province, Indonesia, in 2024. The institution was purposively selected because it is situated in an area with recurrent exposure to earthquakes, tsunamis, and floods, making disaster-related reproductive health preparedness particularly relevant to the competencies expected of its graduates. The target population comprised all 105 active students enrolled in the study program in 2024. Eligibility criteria included active enrollment, no previous formal training in the Minimum Initial Service Package (MISP) or disaster reproductive health kits, willingness to participate in the complete study protocol, and provision of informed consent. Students were excluded if they were absent during the intervention, had incomplete pretest or posttest data, or voluntarily withdrew from the study. A total of 45 students were included in the analytic sample and distributed equally across the three intervention groups, with 15 students per group. Equal allocation facilitated comparability across groups and allowed the FGD intervention to be conducted in appropriately sized small groups. A post hoc power calculation based on the observed effect size (Cohen's $f = 1.08$), $\alpha = 0.05$, three groups, and a total sample of 45 yielded statistical

power >0.99. Because this effect size was derived from the observed study data, however, the post hoc calculation is presented as descriptive information rather than as an *a priori* justification of sample size. The manuscript must explicitly document how the 45 participants were selected from the population of 105. The current source identifies an analytic sample of 45 but does not provide sufficient information to establish whether participants were selected by simple random sampling, systematic sampling, purposive selection, or another procedure. This information is essential for assessing selection bias and the external validity of the findings.

Group Allocation and Contamination Control

Participants were allocated at the class level to the lecture, FGD, or seminar group to minimize intergroup contamination. All three interventions were conducted on the same day, in separate rooms, and within the same time window. Participants were not informed of the instructional differences between groups, and cross-group discussion was restricted until completion of the posttest. These procedures were implemented to minimize information exchange between intervention groups that could attenuate the observed differences in instructional effectiveness.

Knowledge Assessment

Knowledge was assessed using a structured 20-item multiple-choice questionnaire developed by the research team based on the Indonesian Ministry of Health's guidelines for implementing the MISP for reproductive health in health crises and the *Inter-Agency Reproductive Health Kits* manual. Each correct response received five points and each incorrect response received zero points, yielding a total score ranging from 0 to 100. Knowledge was categorized as good (≥75), moderate (56–74), or poor (<56). The instrument assessed six content domains: (1) fundamental concepts of disasters and health crises; (2) reproductive health MISP concepts; (3) the structure of disaster reproductive health kits; (4) the contents and functions of each kit block; (5) target populations and distribution to vulnerable groups; and (6) logistics, storage, and cold-chain management. This domain structure was designed to capture both conceptual knowledge and operational aspects of reproductive health kit preparedness. Content validity was evaluated by an expert panel comprising specialists in community midwifery, maternal–neonatal emergency care, and health disaster management. The instrument was subsequently pilot-tested among students who were not included in the study sample to evaluate item validity and internal consistency. However, the current manuscript does not provide the number of experts, content validity index (CVI), pilot-test sample size, item–total correlation range, critical *r* value, or Cronbach's alpha coefficient. These parameters should be retrieved from the original validation documentation and reported before submission, because the absence of psychometric evidence would substantially weaken the validity of inferences based on the knowledge scores.

Intervention Procedures

All three intervention groups received identical substantive content covering the MISP for reproductive health and disaster reproductive health kit preparation. The instructional dose was standardized at 120 minutes for each group. This standardization was intended to isolate the effect of instructional modality from differences in content exposure or intervention duration. The lecture intervention comprised 90 minutes of structured presentation, followed by 20 minutes of question-and-answer interaction and a 10-minute summary. The FGD intervention consisted of a 15-minute introduction, 75 minutes of structured case-based discussion, 20 minutes of intergroup presentation, and a 10-minute facilitator synthesis. The seminar intervention comprised 75 minutes of expert presentation, 30 minutes of panel discussion and question-and-answer interaction, and a 15-minute closing session. The interventions differed primarily in the degree and nature of learner engagement. The lecture emphasized information reception through listening and note-taking. FGD required students to articulate their understanding, critically respond to peers' arguments, develop kit requirements for disaster scenarios, and present collectively derived solutions. The seminar provided opportunities for students to obtain expert perspectives and participate through voluntary questioning. Within the ICAP framework, these modalities were operationalized as predominantly passive for lecture, interactive for FGD, and active-to-passive for seminar depending on the extent of voluntary participation.

Data Collection

The pretest was administered to all participants on the same day, approximately 30 minutes before the intervention, under supervised conditions and without access to reference materials. Immediately following the intervention, participants completed the posttest using the same instrument, with item order randomized to reduce potential order-recognition effects. Response

sheets were coded numerically without personal identifiers. Scoring was performed by a research assistant who was unaware of participants' group allocation, thereby reducing the potential for assessor bias. Data were independently entered by two personnel and discrepancies were reconciled before statistical analysis.

Variables and Operationalization

The independent variable was instructional modality, categorized as FGD, seminar, or lecture. The primary dependent variable was knowledge of the MISP and disaster reproductive health kits, encompassing understanding of MISP principles, the three-block kit structure, kit contents and functions, target populations, distribution procedures, and logistics management. The primary continuous outcome was the knowledge score ranging from 0 to 100. Knowledge change was operationalized as the difference between posttest and pretest scores for each participant. In addition, normalized learning gain was calculated as the proportion of the observed gain relative to the maximum possible gain and categorized as high (≥0.70), moderate (0.30–0.69), or low (<0.30).

Statistical Analysis

Statistical analyses were conducted using SPSS version [version number], with a two-sided significance threshold of $P < 0.05$. Continuous variables were summarized using means, standard deviations, and ranges. Distributional assumptions were evaluated using the Shapiro–Wilk test, while homogeneity of variance was assessed using Levene's test. Within-group changes in knowledge scores were assessed using paired-samples *t* tests when the assumptions for parametric analysis were satisfied. Differences in posttest scores among the three instructional groups were evaluated using one-way analysis of variance (ANOVA), followed by Bonferroni-adjusted pairwise comparisons when the omnibus test was statistically significant. Because baseline knowledge may differ between groups in a quasi-experimental design, analysis of covariance (ANCOVA), with pretest knowledge as a covariate, was prespecified as a supportive analysis to assess whether between-group differences persisted after adjustment for baseline variation. Effect sizes were reported alongside *P* values to quantify the magnitude of observed differences. Within-group changes were expressed using Cohen's *d*, with small-sample correction through Hedges' *g*. Between-group effects were quantified using eta-squared and omega-squared, with omega-squared emphasized because of its relatively lower bias in small samples. Normalized gain was additionally used to characterize the proportion of attainable learning improvement. Binary logistic regression was used to examine whether instructional modality was associated with attainment of the predefined good-knowledge category (score ≥75).

RESULT

Participant Characteristics

A total of 45 participants completed all study procedures, including the pretest, intervention, and posttest, with no missing outcome data. Baseline demographic, academic, disaster-experience, and prior-exposure characteristics according to intervention group are presented in Table 4.

Table 4. Baseline Characteristics of Participants According to Intervention Group (N = 45)

Characteristic	FGD (n = 15)	Seminar (n = 15)	Lecture (n = 15)	*P* value
Age, mean (SD), years	20.40 (0.83)	20.27 (0.70)	20.33 (0.72)	0.885[a]
Semester, n (%)				0.912[b]
└ Semester IV	9 (60.0)	10 (66.7)	9 (60.0)	
└ Semester VI	6 (40.0)	5 (33.3)	6 (40.0)	
GPA, mean (SD)	3.42 (0.21)	3.38 (0.19)	3.40 (0.22)	0.840[a]
Previous disaster-training experience, n (%)	2 (13.3)	3 (20.0)	2 (13.3)	0.856[b]
Previous experience of being affected by a disaster, n (%)	6 (40.0)	7 (46.7)	5 (33.3)	0.748[b]

Previous exposure to PPAM material, n (%)	0 (0.0)	0 (0.0)	0 (0.0)	—
Pretest knowledge score, mean (SD)	44.67 (10.09)	41.33 (6.66)	42.00 (5.63)	0.452[a]

Note: FGD = focus group discussion; GPA = grade point average; PPAM = *Paket Pelayanan Awal Minimum* (Minimum Initial Service Package); SD = standard deviation.

[a] One-way analysis of variance (ANOVA).

[b] Chi-square test.

The three intervention groups demonstrated comparable baseline profiles. The mean age ranged narrowly from 20.27 to 20.40 years, while the distribution of participants across semesters was broadly similar. Academic achievement was also comparable, with mean grade point averages (GPAs) ranging from 3.38 to 3.42. Likewise, the proportions of participants who had previously attended disaster-related training or had experienced a disaster were relatively balanced across the FGD, seminar, and lecture groups. None of the participants reported previous exposure to formal PPAM-related instructional material, indicating an absence of prior formal learning exposure to the intervention content across all three groups. Baseline knowledge scores were also relatively low across groups, with mean pretest scores of 44.67 in the FGD group, 41.33 in the seminar group, and 42.00 in the lecture group. The difference in baseline knowledge was not statistically significant ($P = 0.452$), suggesting that the groups entered the intervention with broadly comparable levels of prior knowledge. No statistically significant between-group differences were observed in age ($P = 0.885$), semester distribution ($P = 0.912$), GPA ($P = 0.840$), previous participation in disaster training ($P = 0.856$), or previous disaster experience ($P = 0.748$). These findings provide no statistical evidence of substantial baseline imbalance across the intervention groups. Importantly, however, the absence of statistically significant differences should not be interpreted as evidence that the groups were identical at baseline. Given the quasi-experimental design and the relatively small sample size, baseline characteristics may still contribute to residual confounding. Accordingly, baseline knowledge was retained as an important covariate for subsequent adjusted analyses.

Knowledge Outcomes Before and After the Intervention

At baseline, mean knowledge scores were low across all three instructional groups and showed only modest variation, ranging from 41.33 to 44.67 points. The absolute difference between the highest and lowest group means was 3.34 points, suggesting broadly comparable initial knowledge levels and providing a reasonable basis for evaluating differential post-intervention outcomes. Following the intervention, knowledge scores increased in all groups; however, the magnitude of improvement varied substantially according to instructional modality. The FGD group achieved the highest posttest mean (78.00 ± 8.62), crossing the predefined threshold for the "good" knowledge category, whereas the lecture (60.00 ± 8.45) and seminar (57.00 ± 9.60) groups remained within the "fair" category. These findings indicate that although all three approaches were capable of improving knowledge, FGD generated a markedly greater absolute and educationally meaningful improvement. Detailed results are presented in Table 5 and Figure 1.

Table 5. Descriptive Statistics of Knowledge Scores Before and After Intervention by Instructional Group (N = 45)

Instructional group	n	Pretest, mean (SD)	Posttest, mean (SD)	Posttest range	Mean improvement	95% CI	Posttest category
Focus group discussion	15	44.67 (10.09)	78.00 (8.62)	65–90	33.33	29.06–37.61	Good
Seminar	15	41.33 (6.66)	57.00 (9.60)	35–65	15.67	9.76–21.57	Fair
Lecture	15	42.00 (5.63)	60.00 (8.45)	45–75	18.00	13.97–22.03	Fair

Note. SD, standard deviation; CI, confidence interval. The ratio of the largest to smallest posttest SD was 1.14 (9.60/8.45), substantially below the commonly used threshold of 2.0, suggesting no meaningful heterogeneity of posttest variance across groups.

Natural Resources for Human Health
eISSN: 2583-1194

Stability of Individual Knowledge Ranking

The pretest–posttest correlations provided additional insight into the consistency of individual performance within each instructional group. A strong positive correlation was observed in the FGD group (r = 0.670, P = .006), indicating that participants with relatively higher baseline knowledge tended, on average, to retain their relative position at posttest. A moderate positive correlation was also observed in the lecture group (r = 0.528, P = .043). In contrast, the seminar group demonstrated a weak and statistically non-significant association (r = 0.178, P = .525). Thus, individual knowledge trajectories appeared more consistent following FGD and lecture, whereas the seminar intervention was associated with greater variability in the relative ordering of participants. Importantly, these correlations should not be interpreted as measures of intervention effectiveness; rather, they describe the extent to which individual differences in baseline knowledge were preserved following instruction.

Table 6. Pretest–Posttest Correlations by Instructional Group

Instructional group	n	r	P value	Interpretation
Focus group discussion	15	0.670	.006	Strong, significant positive association
Seminar	15	0.178	.525	Weak, non-significant association
Lecture	15	0.528	.043	Moderate, significant positive association

Within-Group Changes in Knowledge

Paired-samples t tests demonstrated statistically significant improvements in knowledge scores in all three instructional groups (all $P < .001$). Nevertheless, the magnitude of change differed considerably. The FGD group exhibited the largest increase, with a mean improvement of 33.33 points (95% CI, 29.06–37.61; $t(14) = -16.73$), followed by the lecture group with an 18.00-point increase (95% CI, 13.97–22.03; $t(14) = -9.59$) and the seminar group with a 15.67-point increase (95% CI, 9.76–21.57; $t(14) = -5.69$). The confidence interval for the FGD improvement was clearly separated from those of the lecture and seminar groups, whereas the intervals for lecture and seminar showed substantial overlap. Collectively, these results indicate that the three approaches were not equivalent in their capacity to increase knowledge, with FGD producing the most pronounced within-participant change.

Table 7. Paired-Samples t Test of Knowledge Scores Before and After Intervention

Group	Mean difference	SD of difference	SE	95% CI	t	df	P value
Focus group discussion	−33.33	7.72	1.99	29.06–37.61	−16.733	14	< .001
Seminar	−15.67	10.67	2.75	9.76–21.57	−5.688	14	< .001
Lecture	−18.00	7.27	1.88	13.97–22.03	−9.589	14	< .001

Note. SD, standard deviation; SE, standard error; CI, confidence interval. Negative mean differences reflect the pretest-minus-posttest calculation; confidence intervals are reported as the magnitude of the corresponding improvement.

Standardized Effect Size and Normalized Learning Gain

The magnitude of within-group change was further evaluated using standardized effect sizes and Hake's normalized learning gain. All three interventions demonstrated substantial standardized effects, although their magnitudes differed markedly. The largest effect was observed in the FGD group (Cohen's d = 4.32; Hedges' g = 4.09), followed by lecture (d = 2.48; g = 2.34) and seminar (d = 1.47; g = 1.39). The consistency of this ordering across both Cohen's d and Hedges' g strengthens the conclusion that FGD generated the largest educational effect. Normalized gain provided an additional perspective by accounting for the amount of knowledge improvement that remained theoretically attainable from each participant's baseline. The FGD group achieved a normalized gain of 0.602, corresponding to approximately 60.2% of the available learning space. Lecture achieved a gain of 0.310, whereas seminar achieved 0.267. Thus, FGD not only produced the largest absolute improvement but also enabled students to attain a substantially greater proportion of the knowledge that remained available for acquisition.

Table 8. Standardized Effect Sizes and Normalized Learning Gain

Group	Cohen's d	Hedges' g	Effect-size interpretation	Normalized gain	Gain category	Improvement room closed
Focus group discussion	4.32	4.09	Very large	0.602	Medium	60.2%
Lecture	2.48	2.34	Very large	0.310	Medium	31.0%
Seminar	1.47	1.39	Large	0.267	Low	26.7%

Note. Cohen's *d* was derived from the paired difference scores, while Hedges' *g* incorporated small-sample bias correction. Effect-size interpretation follows conventional thresholds of 0.20, 0.50, and 0.80 for small, medium, and large effects, respectively. Hake's normalized gain was interpreted as high (≥0.70), medium (0.30–0.69), or low (<0.30).

Between-Group Differences in Posttest Knowledge

A one-way ANOVA demonstrated a statistically significant difference in posttest knowledge among the three instructional groups, $F(2,42) = 24.41$, $P < .001$. The magnitude of the between-group effect was substantial, with $\eta^2 = 0.538$ and $\omega^2 = 0.510$. These estimates indicate that approximately 51%–54% of the variability in posttest knowledge was associated with instructional modality. Cohen's *f* was 1.08, considerably exceeding the conventional threshold for a large effect. Therefore, the observed differences were not only statistically significant but also substantial in practical magnitude.

Table 9. One-Way ANOVA of Posttest Knowledge Scores

Source	Sum of squares	df	Mean square	*F*	*P* value	Effect size
Between groups	3,870.000	2	1,935.000	24.405	< .001	$\eta^2 = 0.538$; $\omega^2 = 0.510$; $f = 1.08$
Within groups	3,330.000	42	79.286	—	—	—
Total	7,200.000	44	—	—	—	—

Bonferroni-adjusted post hoc comparisons demonstrated that the FGD group significantly outperformed both the seminar group (mean difference = 21.00 points, $P < .001$; $d = 2.36$) and the lecture group (mean difference = 18.00 points, $P < .001$; $d = 2.02$). In contrast, the lecture and seminar groups did not differ significantly (mean difference = 3.00 points, $P = 1.000$; $d = 0.34$). The posttest results therefore reveal a clear separation between FGD and the two predominantly transmission-oriented approaches, while lecture and seminar produced broadly comparable outcomes.

Table 10. Bonferroni-Adjusted Post Hoc Comparisons of Posttest Knowledge

Comparison	Mean difference	*P* value	Cohen's *d*	Interpretation
FGD vs seminar	21.00	< .001	2.36	FGD significantly superior
FGD vs lecture	18.00	< .001	2.02	FGD significantly superior
Seminar vs lecture	−3.00	1.000	0.34	No significant difference

Note. Between-group Cohen's *d* was calculated using the pooled within-group standard deviation derived from the ANOVA mean square error.

Adjusted Analysis of Posttest Knowledge

To determine whether the observed superiority of FGD could be explained by baseline differences in knowledge, an ANCOVA was performed with pretest knowledge included as a covariate. After adjustment for baseline scores, instructional modality

remained a highly significant predictor of posttest knowledge, $F(2,41) = 28.09$, $P < .001$, with a partial η^2 of 0.578. This represents a very large adjusted effect, indicating that the association between instructional modality and post-intervention knowledge persisted even after accounting for individual differences in baseline knowledge. Baseline knowledge itself was also significantly associated with posttest performance ($F = 12.68$, $P = .001$; partial $\eta^2 = 0.236$).

Table 11. ANCOVA of Posttest Knowledge with Pretest Knowledge as Covariate

Source	Type III SS	df	Mean square	*F*	*P* value	Partial η^2
Pretest knowledge	786.500	1	786.500	12.68	.001	0.236
Instructional modality	3,485.200	2	1,742.600	28.09	< .001	0.578
Error	2,543.500	41	62.037	—	—	—
Corrected total	6,815.200	44	—	—	—	—

Note. Interpretation of ANCOVA requires verification of the homogeneity-of-regression-slopes assumption. This assumption should be formally tested by examining the interaction between pretest knowledge and instructional modality before the adjusted estimates are considered definitive.

Predictors of Achieving the "Good" Knowledge Category

Binary logistic regression was subsequently used to examine whether instructional modality predicted attainment of the predefined "good" knowledge category (posttest score ≥75), using lecture as the reference category and adjusting for pretest knowledge. Participants receiving FGD had substantially greater odds of achieving the "good" knowledge category than those receiving lecture (adjusted odds ratio [AOR] = 26.00; 95% CI, 2.55–265.12; $P = .006$). This indicates a strong association between FGD participation and attainment of the desired knowledge threshold. However, the wide confidence interval reflects considerable statistical imprecision, which is expected given the small sample size and relatively limited number of participants achieving the outcome. Seminar was not significantly associated with attainment of the "good" category compared with lecture (AOR = 0.12; 95% CI, 0.003–4.48; $P = .252$). Pretest knowledge was independently associated with the outcome, with each one-point increase in baseline knowledge corresponding to approximately 16% higher odds of achieving the "good" posttest category (AOR = 1.16; 95% CI, 1.00–1.34; $P = .044$).

Table 12. Binary Logistic Regression Predicting Achievement of the "Good" Knowledge Category

Predictor	B	SE	Wald	P value	AOR	95% CI
FGD vs lecture	3.258	1.185	7.56	.006	26.00	2.55–265.12
Seminar vs lecture	−2.120	1.850	1.31	.252	0.12	0.003–4.48
Pretest knowledge	0.145	0.072	4.05	.044	1.16	1.00–1.34
Constant	−8.725	3.420	6.51	.011	—	—

The statistical evidence demonstrates a coherent pattern across complementary analyses: FGD produced the largest absolute knowledge improvement, the greatest standardized effect, the highest normalized learning gain, the highest posttest performance, and substantially greater odds of attaining the predefined "good" knowledge category. The persistence of the instructional effect following adjustment for baseline knowledge further supports the interpretation that the superiority of FGD was associated with the instructional modality itself rather than solely with initial differences in participant knowledge.

4. DISCUSSION

The present study demonstrates that lecture, focus group discussion (FGD), and seminar methods all significantly improved midwifery students' knowledge of disaster reproductive health kit preparation, although the magnitude of improvement differed

substantially. Under equivalent instructional content and duration, FGD produced the greatest gain, nearly twice that of lecture and more than twice that of seminar. Notably, only the FGD group achieved a posttest mean within the predefined "good" knowledge category, whereas lecture and seminar remained within the "fair" category and did not differ significantly. The superiority of FGD was consistently reflected in mean change, confidence intervals, standardized effect size, and normalized gain, indicating that the observed advantage represents a meaningful educational effect rather than an artifact of a single statistical indicator. The stronger effect of FGD can be interpreted through the ICAP framework, which distinguishes passive, active, constructive, and interactive forms of cognitive engagement[28–30]. FGD most closely represents interactive learning because students must articulate ideas, evaluate peer responses, explain reasoning, correct misconceptions, and collaboratively construct solutions. These processes promote retrieval, elaboration, self-explanation, and conceptual restructuring, which are generally more conducive to durable learning than passive information exposure.[38] The finding is also consistent with peer-instruction evidence showing that structured discussion can produce greater conceptual learning than unidirectional information transmission[39] Several mechanisms may explain the magnitude of the FGD effect. First, FGD incorporates retrieval practice because students must recall the three-block structure and identify appropriate kit components in response to disaster scenarios. Second, immediate peer feedback enables misconceptions regarding kit classification, target populations, and levels of care to be corrected during learning. Third, small-group interaction creates social accountability, increasing the likelihood that each student actively contributes. Fourth, peer discussion may facilitate cognitive scaffolding within the zone of proximal development, whereby students with stronger understanding help peers interpret complex concepts[31,32]. These mechanisms are particularly relevant to disaster reproductive health preparedness because students must transform technical information into operational decisions. The content itself may further explain the advantage of FGD. Disaster reproductive health kit preparation is not merely a matter of memorizing inventory. Students must determine which kit is appropriate, for whom, at what level of care, for what duration, and under which emergency circumstances. This represents conditional and classificatory knowledge requiring "if–then" reasoning.[40] Scenario-based FGD allows learners to apply PPAM and reproductive health principles to contextualized problems, thereby connecting declarative knowledge with decision making.

Lecture nevertheless produced a substantial improvement of 18.00 points with a large standardized effect (d = 2.48), demonstrating that conventional instruction remains valuable. Its principal strength is efficient and systematic transmission of foundational concepts. However, its normalized gain of 0.310 suggests that lecture alone enabled students to acquire a relatively limited proportion of the knowledge that remained available for improvement. Lecture may therefore be best positioned as a foundational component followed by interactive application. Seminar generated the smallest improvement (15.67 points), standardized effect (d = 1.47), and normalized gain (0.267; low category), and did not significantly outperform lecture. Although seminars provide access to expert experience and opportunities for questions, participation may remain predominantly receptive when interaction is voluntary. Thus, expert presence alone does not guarantee active cognitive engagement. The weak pretest–posttest correlation in the seminar group (r = 0.178, P = 0.525), compared with FGD (r = 0.670) and lecture (r = 0.528), may indicate heterogeneous responses to the intervention. However, this interpretation is exploratory because the small sample and potential regression to the mean may have influenced the association. The findings are consistent with Indonesian studies reporting beneficial effects of FGD on earthquake preparedness and nursing students' learning achievement.[32] More importantly, the present study directly compares three instructional approaches under equivalent content and instructional exposure, providing stronger evidence regarding their relative effectiveness. The findings also correspond with broader evidence demonstrating advantages of active learning over conventional lecture-based instruction.[41] The study therefore extends active-learning evidence to disaster reproductive health preparedness among future midwives, a domain requiring integration of conceptual knowledge and contextual decision making.

From a reproductive-health preparedness perspective, these findings have implications beyond classroom achievement. The Minimum Initial Service Package (MISP) and disaster reproductive health kits are operational components of emergency health-system preparedness, whose implementation depends substantially on competent frontline personnel[33–35]. Strengthening preparedness during pre-service midwifery education may therefore constitute an upstream strategy for improving the human-resource component of disaster readiness. The observed superiority of FGD suggests that curricula should emphasize scenario interpretation, justification of decisions, collaborative problem solving, and correction of misconceptions rather than relying predominantly on information transmission. Several educational implications follow. FGD should be considered a core instructional strategy for PPAM and reproductive health kit preparation rather than merely a supplementary activity. A concise lecture can first establish essential concepts, followed by structured scenario-based FGD requiring students to apply them to

realistic emergencies. Seminars should likewise be redesigned to require active participation through case-based questions, peer discussion, and structured responses. Because lecture and seminar groups remained below the "good" threshold, a single 120-minute session may be insufficient for mastery; criterion-based remediation and repeated practice may therefore be warranted. Knowledge acquisition should subsequently be followed by simulation-based learning involving kit identification, assembly, allocation, and distribution under realistic disaster conditions. The findings also have institutional implications. Midwifery education institutions in disaster-prone areas could strengthen partnerships with reproductive health and disaster-management authorities so that classroom learning is connected to actual preparedness systems. Such collaboration could expose students to operational processes involving commodity coordination, storage, requesting, distribution, and deployment during emergencies, thereby improving alignment between educational outcomes and health-system requirements. Several strengths support interpretation of the findings. The three modalities used equivalent substantive content and instructional duration, reducing differential exposure as an explanation for the observed differences. Baseline knowledge was broadly comparable, while the use of confidence intervals, standardized effect sizes, and normalized gain enabled assessment of both statistical and practical improvement. Blinded scoring further reduced assessor-related bias. Nevertheless, several limitations require consideration. The absence of a no-treatment control prevents complete separation of intervention effects from testing and temporal effects. Class-level rather than individual allocation may have introduced clustering and group-level confounding. The sample of 15 students per group was small, making estimates and particularly large effect sizes potentially unstable. Immediate posttesting also prevents conclusions about retention. Furthermore, the outcome measured knowledge rather than actual competence in assembling or distributing kits. The single-institution setting limits generalizability, while differences between the FGD facilitator and seminar speaker introduce potential instructor effects.

CONCLUSION

This three-group quasi-experimental study shows that lecture, focus group discussion, and seminar methods all significantly improved midwifery students' knowledge of disaster reproductive health kit preparation, but not to an equivalent degree. With content dose and duration held constant, FGD produced a knowledge gain nearly twice that of lecture and more than twice that of seminar, was the only method whose posttest mean reached the "good" competency threshold, and remained the strongest predictor of knowledge achievement after adjustment for baseline differences. Seminar and lecture did not differ significantly from one another, indicating that access to an expert speaker adds limited value when learners remain passive recipients. Because allocation was performed at the class level within a single institution and group sizes were modest, these results should be regarded as preliminary evidence rather than definitive proof of superiority, warranting confirmation through cluster-randomized trials with longer-term retention outcomes. Nonetheless, given the consistency of the FGD advantage across every analytic approach used, midwifery education programs in disaster-prone regions such as Aceh have a reasonable evidence base to prioritize structured, scenario-based small-group discussion as the core method for building disaster reproductive health preparedness competence, while retaining lecture to establish conceptual grounding and seminar to enrich field-based context.

DATA AVAILABILITY

The data presented in this study are available on request from the corresponding author.

CONFLICTS OF INTEREST

The authors declare no conflict of interest and no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGMENTS

The authors thank the Director of Poltekkes Kemenkes Aceh and the Head of the Meulaboh Midwifery Study Program for permission and support in conducting this research, for their willingness to lead the intervention sessions, and all the students who gave their time and shared their experience as participants.

AUTHOR CONTRIBUTIONS (CREDIT)

Author Contributions (CRediT): Yushida: Conceptualization, Methodology, Investigation (data collection lead), and Writing – original draft. Rina Julianti: Methodology (instrument development) and Investigation (intervention delivery). Satrinawati:

Formal analysis, Data curation, and Writing – review & editing (statistical interpretation). Isnaini Putri: Validation and Writing – review & editing (critical manuscript review).

REFERENCES

[1] Sulistiadi W, Wasir R, Thalib W, Ayuningtyas D, Bawazier N, Buskens E. Building health systems resilience: understanding the social, economic, and cultural impacts of climate change from stakeholders' perspectives in Indonesia. Archives of Public Health 2024;82:168. https://doi.org/10.1186/S13690-024-01403-4

[2] Heo S, Park S, Lee DK. Multi-hazard exposure mapping under climate crisis using random forest algorithm for the Kalimantan Islands, Indonesia. Sci Rep 2023;13:13472. https://doi.org/10.1038/S41598-023-40106-8

[3] Karnaji K, Susanti E, Ariadi S, Saud M. Social impacts and post-disaster management in disaster-prone areas of East Java, Indonesia. Jàmbá: Journal of Disaster Risk Studies 2024;16:1747. https://doi.org/10.4102/JAMBA.V16I1.1747

[4] Solekhah N, Hakim FN, Wahyono E, Prayoga RA, Fatimah S, Purbandini L, et al. Learning from volcanic eruptions: Co-production of knowledge at Merapi and Kelud, Indonesia. Jàmbá: Journal of Disaster Risk Studies 2025;17:1881. https://doi.org/10.4102/JAMBA.V17I1.1881

[5] Doocy S, Rofi A, Moodie C, Spring E, Bradley S, Burnham G, et al. Tsunami mortality in Aceh Province, Indonesia. Bull World Health Organ 2007;85:273. https://doi.org/10.2471/BLT.06.033308

[6] Lépine A, Restuccio M, Strobl E. Can we mitigate the effect of natural disasters on child health? Evidence from the Indian Ocean tsunami in Indonesia. Health Econ 2021;30:432–52. https://doi.org/10.1002/HEC.4202

[7] Doocy S, Gorokhovich Y, Burnham G, Balk D, Robinson C. Tsunami Mortality Estimates and Vulnerability Mapping in Aceh, Indonesia. Am J Public Health 2007;97:S146. https://doi.org/10.2105/AJPH.2006.095240

[8] Chi PC, Urdal H, Umeora OUJUJ, Sundby J, Spiegel P, Devane D. Improving maternal, newborn and women's reproductive health in crisis settings. Cochrane Database Syst Rev 2018;2018:CD011829. https://doi.org/10.1002/14651858.CD011829.PUB2

[9] Sheikh J, Allotey J, Kew T, Khalil H, Galadanci H, Hofmeyr GJ, et al. Vulnerabilities and reparative strategies during pregnancy, childbirth, and the postpartum period: moving from rhetoric to action. EClinicalMedicine 2023;67:102264. https://doi.org/10.1016/J.ECLINM.2023.102264

[10] Perez M, Galang RR, Snead MC, Strid P, Bish CL, Tong VT, et al. Emergency Preparedness and Response: Highlights from the Division of Reproductive Health, 2011–2021. J Womens Health (Larchmt) 2021;30:1673. https://doi.org/10.1089/JWH.2021.0553

[11] Cangöl E, Yıldırım ES. The effect of disaster preparedness training on knowledge and preparedness levels of pregnant women: a randomized controlled study. BMC Pregnancy Childbirth 2026;26:162. https://doi.org/10.1186/S12884-026-08635-Y

[12] Okyay EK, Dinçer M, Bal Z. Development and psychometric evaluation of the post-disaster women's sexual and reproductive health assessment scale. BMC Womens Health 2026;26:56. https://doi.org/10.1186/S12905-025-04229-6

[13] Zotti ME, Williams AM, Wako E. Post-disaster Health Indicators for Pregnant and Postpartum Women and Infants. Matern Child Health J 2015;19:1179. https://doi.org/10.1007/S10995-014-1643-4

[14] Onyango MA, Hixson BL, McNally S. Minimum Initial Service Package (MISP) for reproductive health during emergencies: Time for a new paradigm? Glob Public Health 2013;8:342–56. https://doi.org/10.1080/17441692.2013.765024

[15] Lisam S. Minimum initial service package (MISP) for sexual and reproductive health in disasters. J Evid Based Med 2014;7:245–8. https://doi.org/10.1111/JEBM.12130

[16] Broaddus-Shea ET, Kobeissi L, Ummer O, Say L. A systematic review of monitoring and evaluation indicators for sexual and reproductive health in humanitarian settings. Confl Health 2019;13:43. https://doi.org/10.1186/S13031-019-0221-1

[17] Newman L, Rowley J, Hoorn S Vander, Wijesooriya NS, Unemo M, Low N, et al. Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting. PLoS

One 2015;10. https://doi.org/10.1371/JOURNAL.PONE.0143304

[18] Looker KJ, Magaret AS, Turner KME, Vickerman P, Gottlieb SL, Newman LM. Global estimates of prevalent and incident herpes simplex virus type 2 infections in 2012. PLoS One 2015;10. https://doi.org/10.1371/JOURNAL.PONE.0114989

[19] Graham K, Colquhoun S, LaBond C, Housen T, Mohale H, Campbell S, et al. The preparedness and readiness of rural and remote primary care midwives working in low- and middle-income countries: A scoping review. Women and Birth 2025;38:101866. https://doi.org/10.1016/J.WOMBI.2024.101866

[20] Pantaleo F, Stievano A, Mastroianni C, Petrucci G, Mazzitelli N, Piredda M, et al. The Nursing Student Licensure Examination: A Scoping Review. Nurs Rep 2025;15:299. https://doi.org/10.3390/NURSREP15080299/S1

[21] Shimojo S, Hayashi Y, Shimojo S, Hayashi Y. Age-related differences in explanatory activities during collaborative learning with concept maps: experimental investigation using epistemic network analysis. International Journal of Computer-Supported Collaborative Learning 2025 20:4 2025;20:491–518. https://doi.org/10.1007/S11412-025-09456-5

[22] Thurn CM, Edelsbrunner PA, Berkowitz M, Deiglmayr A, Schalk L. Questioning central assumptions of the ICAP framework. NPJ Sci Learn 2023;8. https://doi.org/10.1038/S41539-023-00197-4

[23] Archer RM, Wagner M, Snell L. A theoretical framework for experiential learning: Creating Active Knowledge Experiences (CAKE). Advances in Health Sciences Education 2026 2026:1–31. https://doi.org/10.1007/S10459-026-10508-0

[24] Kong Y. The Role of Experiential Learning on Students' Motivation and Classroom Engagement. Front Psychol 2021;12:771272. https://doi.org/10.3389/FPSYG.2021.771272

[25] Boke H, Aygun Y, Tufekci S, Yagin FH, Canpolat B, Norman G, et al. Effects of cooperative learning on students' learning outcomes in physical education: a meta-analysis. Front Psychol 2025;16:1508808. https://doi.org/10.3389/FPSYG.2025.1508808

[26] Peer KS, McClendon RC. Sociocultural Learning Theory in Practice: Implications for Athletic Training Educators. J Athl Train 2002;37:S-136.

[27] Lai WF. Integrating sociocultural perspectives into a university classroom: A case study of students' experience. Heliyon 2023;9:e17228. https://doi.org/10.1016/J.HELIYON.2023.E17228

[28] Chi MTH, Adams J, Bogusch EB, Bruchok C, Kang S, Lancaster M, et al. Translating the ICAP Theory of Cognitive Engagement Into Practice. Cogn Sci 2018;42:1777–832. https://doi.org/10.1111/COGS.12626

[29] Thurn CM, Edelsbrunner PA, Berkowitz M, Deiglmayr A, Schalk L. Questioning central assumptions of the ICAP framework. NPJ Sci Learn 2023;8:49. https://doi.org/10.1038/S41539-023-00197-4

[30] Chi MTH, Adams J, Bogusch EB, Bruchok C, Kang S, Lancaster M, et al. Translating the ICAP Theory of Cognitive Engagement Into Practice. Cogn Sci 2018;42:1777–832. https://doi.org/10.1111/COGS.12626

[31] Arja SB, Ponnusamy K, Kottathveetil P, Ahmed TFA, Fatteh R, Arja SB. Effectiveness of Small Group Discussions for Teaching Specific Pharmacology Concepts. Med Sci Educ 2020;30:713. https://doi.org/10.1007/S40670-020-00938-9

[32] Kraatz E, Nagpal M, Lin TJ, Hsieh MY, Ha SY, Kim S, et al. Teacher Scaffolding of Social and Intellectual Collaboration in Small Groups: A Comparative Case Study. Front Psychol 2020;11:587058. https://doi.org/10.3389/FPSYG.2020.587058

[33] Myers A, Sami S, Onyango MA, Karki H, Anggraini R, Krause S. Facilitators and barriers in implementing the Minimum Initial Services Package (MISP) for reproductive health in Nepal post-earthquake. Confl Health 2018;12. https://doi.org/10.1186/S13031-018-0170-0

[34] Myers A, Sami S, Onyango MA, Karki H, Anggraini R, Krause S. Facilitators and barriers in implementing the Minimum Initial Services Package (MISP) for reproductive health in Nepal post-earthquake. Confl Health 2018;12:35. https://doi.org/10.1186/S13031-018-0170-0

[35] Greer A, Pai R, Tanabe M, Martin L, Tran NT. Evolution of the humanitarian roster system for sexual and reproductive health. Sex Reprod Health Matters 2026;33:2686033. https://doi.org/10.1080/26410397.2026.2686033