



Original Research Paper

Knowledge of Midwifery Instruments Following Use of the Informasi Instrumen Kebidanan Digital Book: A Quasi Experimental Study

Putri Yuniartis*, Wulan Anjani, Devi Devinda Novita Sari, Hidayanti Arifuddin, Charisma Destrikasari, Mita Herlina Sari, Nur Dwi Yulianti

Health Polytechnic of the Ministry of Health Jakarta I, Jakarta, Indonesia

Access this article online
Quick Response Code :



DOI : 10.22487/hj.v12i4.2312

Email Corresponding:
p.yuniartis@gmail.com

Page : 763-774

Article History:

Received: 2026-06-26

Revised: 2026-09-28

Accepted: 2026-10-06

Published by:

Tadulako University,
Managed by Faculty of
Medicine.

Website :

<https://jurnal.fk.untad.ac.id/index.php/hj/index>



This work is licensed under a
Creative Commons Attribution-
ShareAlike 4.0 International
License

Abstract

Background: A digital book may support learning instrument names and functions, but knowledge scores do not establish practical competence. **Objective:** To compare immediate posttest instrument knowledge between INSTADAN and lecturer PowerPoint groups after baseline adjustment. **Methods:** This nonrandomized study analyzed 161 first year midwifery students from four intact classes (INSTADAN, n=78; control, n=83). A 30 item knowledge test was administered before and after teaching. Individual level ANCOVA adjusted for pretest score, age, and information sources with HC3 standard errors. **Results:** Mean scores rose from 68.72 to 85.04 with INSTADAN and from 68.84 to 79.88 with PowerPoint. The adjusted difference was 6.08 points (95% CI, 3.99 to 8.17; P<0.001), equivalent to approximately 1.82 correct answers. Student stratified bootstrap gave a 95% interval of 4.05 to 8.14 points. Neither method adjusted for class dependence. **Conclusion:** INSTADAN was associated with higher immediate knowledge scores; four class allocation, baseline imbalance, and incomplete assessment documentation limit causal interpretation. Retention and practical competence were not examined.

Keywords: Digital book; INSTADAN; Knowledge; Midwifery education; Midwifery instruments.

Introduction

Midwifery education requires students to integrate factual knowledge, clinical reasoning, and practical skills. Early familiarity with instrument names, appearance, functions, and clinical context is one component of preparation for laboratory learning. Educational standards address broader professional competence, whereas a written knowledge test measures a narrower outcome^{1,2}. Digital resources may support this preparation by making explanatory material available beyond scheduled teaching. Reviews of learning through the internet and online education in the health professions have

examined knowledge, skills, attitudes, and satisfaction, with variation across interventions and comparators^{3,4}. A digital format should therefore be evaluated for its specific educational purpose rather than assumed to improve every dimension of competence.

First year students encounter unfamiliar instruments while developing the vocabulary and conceptual organization needed for clinical learning. A presentation used during a lecture may not provide the same opportunities for independent review as an accessible illustrated resource, but differences in teaching time, content, access, or lecturer support can also affect test scores. Evidence mapping identifies

substantial variation in the design and evaluation of digital education, including limitations in implementation and transfer across settings⁵. The relevant problem is whether an instrument focused resource adds knowledge beyond the material offered to a comparison group under the conditions studied. This requires a clearly defined comparator and a valid assessment, rather than a comparison of technology labels alone.

Informasi Instrumen Kebidanan, introduced here as INSTADAN, organizes instrument names, types, images, functions, and practical information for pregnancy, childbirth, postpartum, and newborn care. Its proposed contribution is the organization of instrument specific visual and verbal information for beginning students. Cognitive theory of multimedia learning provides a rationale for coherent integration of words and images, although it does not establish that every digital resource is effective⁶. Mobile learning research and nursing students' perspectives also indicate that access and use must be considered separately from learning outcomes^{7,8}. Relevant Indonesian studies include animated video with booklet education, digital comics, and educational video resources; these differ in audience, teaching design, and measured endpoint^{9,10,11}. The present study addresses knowledge of midwifery instruments in one institution and does not claim to establish the superiority of digital books in general.

The primary question was whether students receiving INSTADAN had a higher immediate posttest knowledge score than students receiving lecturer PowerPoint material after adjustment for pretest knowledge, age, and recorded information sources. The primary estimand was the adjusted mean difference in posttest score, expressed in points on a 0 to 100 scale. The corresponding null hypothesis was a zero adjusted difference. Secondary objectives

were to describe changes within each group, compare unadjusted gains, and examine the recorded Good knowledge category. These secondary outcomes provide complementary descriptions and do not replace the continuous primary outcome. Software choice and production of figures are analytical procedures rather than research objectives.

An instrument focused digital book could provide structured preparation before demonstration sessions and an additional resource for independent review. Its potential value depends on accurate content, accessible design, actual use, and integration with supervised practice. A local comparison of knowledge scores can help determine whether further evaluation is justified, while observed handling skills, retention, and clinical performance require separate assessment. A study comparing video with simulation in HTJ illustrates an assessment of a practical skill rather than knowledge alone¹². Accordingly, the present findings are intended to inform educational development and the design of further evaluations, not to establish readiness for unsupervised clinical practice.

Materials and Methods

Study Design

The study used a quasi experimental design with pretest and posttest measurements and a nonequivalent control group at Poltekkes Kemenkes Jakarta I. Allocation followed existing first year class membership. Classes 1A and 1B received INSTADAN, while classes 1C and 1D received lecturer PowerPoint material. Thus, four classes, with two in each condition, constituted the allocation units; students were the units of measurement. Individual randomization was not reported. The intervention should be understood as the educational package delivered to each class, because the available description does not

isolate digital format from content or teaching exposure.

Reporting follows the transparency principles of the TREND statement for nonrandomized evaluations¹³. Recruitment, teaching, and assessment dates were not documented in the study materials available for this review. The posttest was described as immediate, but the interval after teaching and the length of digital book access were not specified. Consequently, the chronology of data collection relative to ethics approval cannot be verified from the available files.

Sample

The original Excel dataset contained 161 analyzable first year midwifery students, comprising 78 in the INSTADAN group and 83 in the control group. All records had nonmissing age, information source, correct answer counts, pretest and posttest scores, and recorded knowledge categories. The scores exactly matched the correct answer counts divided by 30 and multiplied by 100. Nevertheless, the dataset does not establish how many students were eligible, invited, declined, excluded, or absent from an assessment.

The supplied report did not describe an advance sample size calculation, recruitment strategy, or detailed eligibility rules beyond the student population and available measurements. The achieved sample is therefore described as the available analytic cohort. No retrospective power calculation was used to justify it. Because exposure was shared within classes, precision depends on the number and size of classes and their internal similarity as well as the number of students. With only two classes per condition, the nominal student sample cannot be interpreted as 161 independent intervention allocations.

Data Collection Technique

INSTADAN was described as a digital book containing instrument names, types, images, functions, and practical information covering pregnancy, childbirth, postpartum, and newborn care. Control classes received lecturer PowerPoint material. The report did not establish whether content, instructor contact, learning objectives, assessment preparation, or total learning time were matched. Table 1 distinguishes the available intervention description from information needed for replication, using the intervention reporting principles of TIDieR.¹⁴

Knowledge was assessed with the same 30 item test before and after the learning period. Scores were calculated as the number of correct answers divided by 30 and multiplied by 100. Each correct answer therefore contributed approximately 3.33 points. The report did not provide the item wording, response format, content blueprint, expert review procedure, pilot sample, item analysis, or reliability coefficient. Validity evidence concerns the interpretation and intended use of scores, and cannot be inferred from the number of items alone¹⁵. Internal consistency also does not, by itself, demonstrate that a test measures the intended construct¹⁶.

The recorded knowledge categories were Poor (Kurang), Moderate (Cukup), and Good (Baik). In the supplied records, a score of 73.33 was labeled Moderate whereas 76.67 was labeled Good; however, an authoritative classification rule and its source were not provided. These categories were therefore used as recorded and not interpreted as validated competency thresholds. Age and nonexclusive information sources, including books, practice sites, laboratory practicum, and social media, were recorded. Whether source exposure was measured strictly before the intervention needs confirmation, because postintervention sources

would alter the interpretation of covariate adjustment.

Table 1. Reported intervention and comparator characteristics

Component	INSTADAN digital book	Lecturer PowerPoint material
Learning medium	Illustrated digital book on midwifery instruments	Lecturer PowerPoint material
Content described	Names, types, images, functions and practical information	Detailed content and coverage not documented
Assigned classes	1A and 1B	1C and 1D
Analyzed students	78	83
Delivery and access	Device compatibility, file version, access method and instructions not documented	Presentation delivery, file access and instructions not documented
Dose and teaching conditions	Duration, number of sessions, instructor contact and usage logs not documented	Duration, teaching time and instructor comparability not documented
Contamination and fidelity	Sharing across classes and actual book use not documented	Access to the intervention book or other additional resources not documented
Assessment	Same 30 item pretest and posttest; exact timing not documented	Same 30 item pretest and posttest; exact timing not documented

Data Analysis Technique

The continuous posttest score was the primary outcome. The ANCOVA model, reproduced from individual Excel records, included group, pretest score, age, and three binary indicators for books or modules, practice sites, and social media. Laboratory practicum was omitted because all 161 students reported it. This additive specification estimated the group contrast conditional on these observed covariates. Adjustment for baseline score is appropriate for comparisons of follow up measurements but cannot remove confounding from nonrandom class allocation¹⁷. A group by pretest interaction was evaluated as a model diagnostic, not included in the primary additive model.

Baseline differences were quantified using signed standardized mean differences calculated from the individual records. For continuous variables, group means were contrasted using the square root of the average group sample variances. For binary sources, proportions were contrasted using the square root of the average Bernoulli variances. Each difference was oriented as INSTADAN minus control. Laboratory practicum had no variation

and no defined standardized difference. Such summaries characterize baseline imbalance; nonsignificant baseline tests are not evidence of equivalence¹⁸.

The individual level ANCOVA used heteroskedasticity robust HC3 standard errors after the Breusch Pagan test indicated heteroskedasticity ($P=0.0011$). HC3 does not adjust for correlated outcomes among students within the same class¹⁹. The accompanying Python code specified 5,000 bootstrap resamples of students with replacement, stratified by intervention group, using random seed 20260822 and percentile 95% limits. We reproduced this resampling scheme and its interval, but it does not resample intact classes or provide a cluster adjusted inference.

Students within the same teaching class can share lecturers, schedules, peer interactions, and other conditions. Inference with very few clusters requires particular caution^{20,21}. Individual outcome data and regression outputs were available and reproduced; however, the Excel file identified only the combined intervention classes (1A and 1B) and control classes (1C and 1D), without a separate class identifier for each student. Class

specific sample sizes and intraclass correlations therefore could not be reconstructed. With only four intact classes, a standard cluster robust covariance or mixed model should not be considered a reliable correction by default. Including all class fixed effects would absorb the group indicator because treatment is constant within each class.

Model based adjusted means were calculated directly from the fitted additive ANCOVA by setting the group indicator to INSTADAN or control while holding the remaining covariates at their pooled sample means. The adjusted means were 85.52 points (HC3 95% CI, 84.29 to 86.74) for INSTADAN and 79.44 points (HC3 95% CI, 77.75 to 81.12) for control. Their difference is the adjusted contrast of 6.08 points. These model based intervals account for individual level heteroskedasticity but not class dependence; they should not be interpreted as class adjusted precision estimates.

Secondary analyses included paired t tests of within student change, an unadjusted Welch test of gain scores, a Mann Whitney sensitivity test, and logistic regression of the recorded Good category adjusted for baseline score and age. Dichotomization reduces information and the category threshold lacks independently documented validation, so categorical findings were treated as secondary²². Logistic odds ratios were not presented as risk ratios. Crude category proportions, their difference, and a risk ratio were calculated from observed counts. Although the individual records are available, a dependable class adjusted analysis of the binary outcome is not possible without separate class identifiers and a design appropriate for so few clusters²³. Within group improvement was distinguished from the between group contrast²⁴.

Analyses used a nominal two sided significance level of 0.05. All tables were

checked against the 161 individual records and the supplied Python code, including the original 5,000 iteration stratified student bootstrap. No multiplicity correction for secondary comparisons was prespecified in the materials. The figures display group means and sample standard deviations or regression coefficients with HC3 intervals; class adjusted confidence intervals and unrecorded outcomes were not fabricated.

Ethical Consideration

The source manuscript reported ethical exemption from the Health Research Ethics Committee of Poltekkes Kemenkes Jakarta I, No.065/KEPK PKJI/V/2026, valid from May 21, 2026, to May 20, 2027. It stated that participants received information about the objectives, procedures, voluntary participation, confidentiality, and withdrawal without academic consequences, and that data were analyzed anonymously and reported in aggregate. The original ethics document and recruitment dates were not included in the supplied files; these details require confirmation against the approved study documentation.

Result

The Excel dataset included 161 students: 78 (48.4%) assigned to the two INSTADAN classes and 83 (51.6%) assigned to two control classes. Table 2 displays baseline characteristics recalculated from those records. Pretest scores and ages were similar on average, but books or modules, practice sites, and social media were more commonly reported by control students. These clear differences in sources of information, together with possible unmeasured class factors, prevent treating the groups as exchangeable.

Table 2. Student characteristics and baseline standardized differences

Characteristic	INSTADAN (n = 78)	Control (n = 83)	Standardized difference
Age, years	18.95 (0.75)	18.95 (0.60)	-0.005
Pretest score, points	68.72 (7.92)	68.84 (7.11)	-0.016
Books or modules, n (%)	46 (59.0)	61 (73.5)	-0.311
Practice sites, n (%)	45 (57.7)	70 (84.3)	-0.614
Laboratory practicum, n (%)	78 (100.0)	83 (100.0)	Not defined
Social media, n (%)	16 (20.5)	49 (59.0)	-0.856
Pretest Poor category, n (%)	6 (7.7)	3 (3.6)	Not calculated
Pretest Moderate category, n (%)	58 (74.4)	68 (81.9)	Not calculated
Pretest Good category, n (%)	14 (17.9)	12 (14.5)	Not calculated

Continuous values are mean (SD); categorical values are n (%). Information sources permitted multiple responses. Age means were rounded; standardized differences used unrounded data.

Mean knowledge scores increased from 68.72 to 85.04 in the INSTADAN group and from 68.84 to 79.88 in the control group. Mean gains were 16.32 and 11.04 points, respectively (Table 3). These unadjusted summaries are not

interchangeable with covariate adjusted estimates. The outcome measures written knowledge of instruments; it does not directly assess handling skills, safe clinical performance, retention, or patient outcomes.

Table 3. Reported raw pretest, posttest, and change scores

Measure	INSTADAN (n = 78)	Control (n = 83)
Pretest mean (SD), points	68.7 (7.9)	68.8 (7.1)
Posttest mean (SD), points	85.0 (7.2)	79.9 (9.2)
Gain mean (SD), points	16.3 (5.6)	11.0 (8.3)
paired t test P value	<0.001	<0.001
paired standardized change	d = 2.90	d = 1.34

Primary adjusted comparison. The reproduced individual level ANCOVA estimated an INSTADAN minus control contrast of 6.08 points (HC3 95% CI, 3.99 to 8.17; $P < 0.001$; Table 4). The pretest coefficient was 0.65 points per point of baseline score (95% CI, 0.51 to 0.80). Model R^2 was 0.437 and adjusted R^2 was 0.415. These measures refer to in sample model fit, not external prediction. Because the analytic data lack separate class identifiers, these P values and confidence limits do not establish inference at the intervention allocation level.

The group by pretest interaction was 0.022 ($P = 0.892$), showing no statistical evidence of nonparallel fitted slopes in this diagnostic model, without proving the assumption. The reproduced stratified student bootstrap 95% percentile interval was 4.05 to 8.14 points, close to the HC3 interval. Both procedures

sampled or modeled individuals rather than the four teaching classes. Verified model based adjusted group means and their individual level HC3 confidence intervals are shown in Table 4.

Secondary continuous comparisons. The observed mean gain was 5.28 points larger with INSTADAN than with PowerPoint (Welch 95% CI, 3.09 to 7.47; $P < 0.001$). The corresponding unadjusted Hedges g was 0.74. A Mann Whitney test of gain also yielded $P < 0.001$ (exact software output $P = 8.94 \times 10^{-7}$). Paired tests within each group gave $P < 0.001$ (Table 3). These reproducible student level analyses agree in direction but do not establish robustness to the four class assignment mechanism.

Secondary categorical comparison. Recorded Good posttest knowledge occurred in 72 of 78 INSTADAN students (92.3%) and 61 of 83 controls (73.5%). The unadjusted

absolute proportion difference was 18.81 percentage points and the crude risk ratio was 1.26. A logistic regression controlling for pretest and age yielded an adjusted odds ratio of 7.90 (95% CI, 2.43 to 25.70; $P=0.0006$;

Table 5). As Good knowledge was common, this odds ratio must not be interpreted as an eightfold relative probability. These estimates neither validate the Good threshold nor account for class clustering.

Table 4. Primary ANCOVA contrast, verified model adjusted means, and diagnostics

Model component or summary	Estimate	95% CI	P value
INSTADAN versus control, points	6.08	3.99 to 8.17	<0.001
Pretest score, per point	0.65	0.51 to 0.80	<0.001
Age, per year	-1.05	-2.47 to 0.36	0.143
Books or modules: yes versus no	1.31	-0.68 to 3.30	0.195
Practice sites: yes versus no	0.57	-1.65 to 2.79	0.616
Social media: yes versus no	1.30	-1.04 to 3.64	0.277
Model adjusted mean: INSTADAN	85.52	84.29 to 86.74	Not applicable
Model adjusted mean: control	79.44	77.75 to 81.12	Not applicable
Diagnostic interaction: group $\times$ pretest	0.02	-0.30 to 0.34	0.892
Breusch Pagan test	22.20 (LM statistic)	Not applicable	0.0011
Student stratified bootstrap (5,000 resamples) $\ddagger$	6.08	4.05 to 8.14	Not supplied
R ² / adjusted R ² , additive model	0.437 / 0.415	Not applicable	Not applicable

Marginal means and their HC3 intervals are predictions at the pooled sample covariates. $\ddagger$ Bootstrap resampled students within study groups, not teaching classes. All intervals are conditional on individual level analysis.

Table 5. Secondary recorded categories and reproduced logistic model

Outcome or model term	INSTADAN n (%) / estimate	Control n (%) / 95% CI	P value
Recorded Good posttest, n (%)	72 (92.3%)	61 (73.5%)	Not calculated
Recorded Moderate posttest, n (%)	6 (7.7%)	22 (26.5%)	Not calculated
Recorded Poor posttest, n (%)	0 (0.0%)	0 (0.0%)	Not calculated
Crude proportion difference, percentage points	18.81	Not estimated	Not calculated
Crude risk ratio	1.26	Not estimated	Not calculated
Adjusted OR: INSTADAN versus control	7.90	2.43 to 25.70	0.0006
Adjusted OR: pretest, per point	1.20	1.11 to 1.29	<0.001
Adjusted OR: age, per year	0.77	0.35 to 1.67	0.510
Reported apparent AUC	0.851	Not supplied	Not applicable

Figure 1 summarizes verified individual record group means and standard deviations before and after the intervention. It does not depict paired individual trajectories. Figure 2

plots the reproduced ANCOVA coefficients with HC3 confidence intervals. Neither visualization accounts for class level assignment uncertainty.

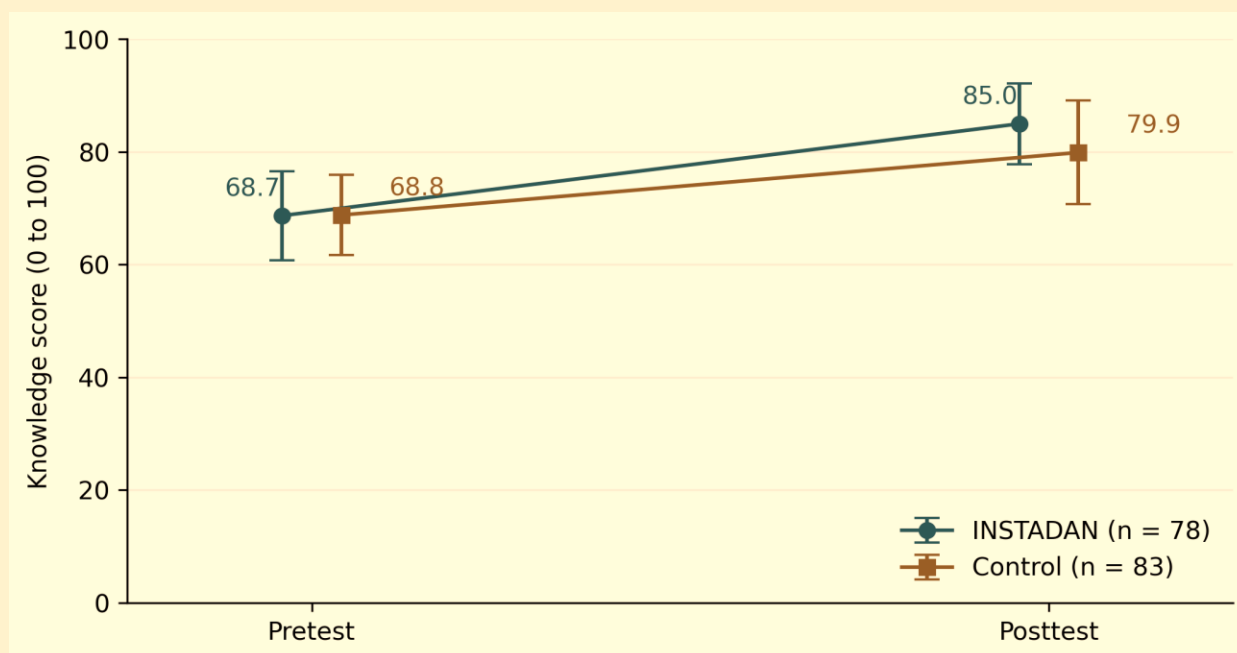


Figure 1. Mean knowledge scores calculated from the 161 Excel records before and after learning. Error bars represent one sample standard deviation, not confidence intervals. Lines join group means rather than individual trajectories; n=78 for INSTADAN and n=83 for control.

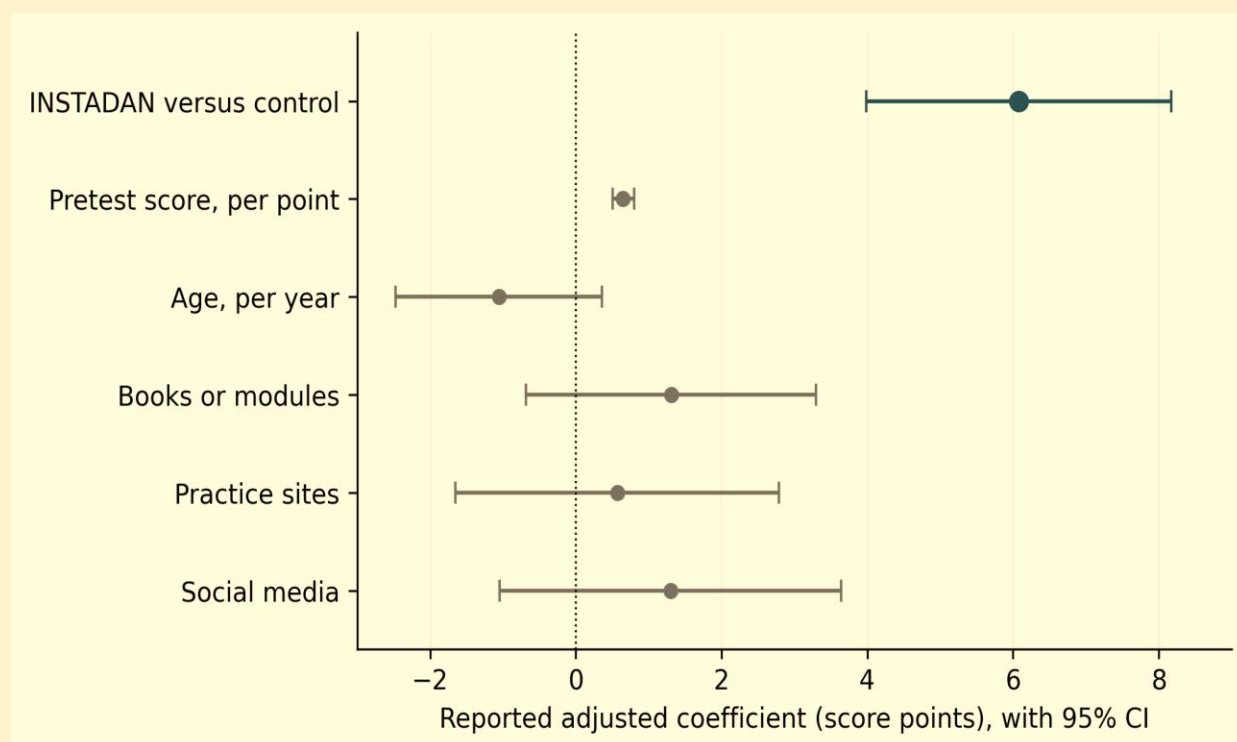


Figure 2. Additive ANCOVA coefficients reproduced from the individual Excel records with heteroskedasticity robust HC3 95% confidence intervals. The INSTADAN coefficient is the main adjusted contrast. These confidence intervals are not adjusted for clustering within the four classes.

Both groups improved on the written knowledge measure. The adjusted individual level contrast favored INSTADAN by 6.08 points, and the secondary Good category was more frequent in that group. The adjusted odds ratio appears much larger than the crude risk ratio because odds and risks are different scales. No established minimally important difference or validated clinical competence threshold was available; cluster dependent uncertainty remains unresolved.

Discussion

The reproduced primary ANCOVA estimate favored INSTADAN by 6.08 points after adjustment for measured baseline knowledge, age, and information sources. Direct model based standardized means were 85.52 and 79.44 points; their HC3 confidence intervals quantify individual level, not class level, uncertainty. On a 30 item test, the adjusted group difference corresponds to about 1.82 additional correct answers (6.08 divided by 100/30). This translation does not establish a minimally important educational difference or safe clinical competence. The larger unadjusted gain in the INSTADAN arm supports a consistent descriptive pattern, but improvement within either arm cannot isolate a treatment effect. The common Good category also makes its adjusted odds ratio unsuitable as a literal relative probability.

The direction of the knowledge comparison is compatible with research examining digital education in health professions, but reviews include varied instructional methods and cannot be treated as direct replications of this intervention^{3,4,5}. Recent syntheses of visual interventions and video approaches also concern diverse populations and educational outcomes^{25,26}. The four relevant HTJ studies provide useful distinctions. Adelia and colleagues evaluated animated video with a booklet against leaflet

education and reported a larger gain in parental knowledge, with nonrandom site allocation limiting causal interpretation⁹. Azizah and colleagues compared video and simulation using an observed cardiopulmonary resuscitation skill outcome, which differs from the written knowledge outcome here¹². Purwati and Andiani did not demonstrate a significant difference in interest in breast self examination with digital comics; interest, knowledge, and behavior should not be treated as interchangeable¹⁰. Subekti and colleagues described literacy following educational video exposure without a baseline or concurrent comparison in the reported design, so that report cannot establish a comparative learning effect¹¹. Together, these examples support attention to instructional design and outcome definition rather than an assumption that a visual medium is inherently effective.

An illustrated book could support recognition by placing an instrument image beside a concise explanation of its name and function. Multimedia learning theory offers a plausible account of how coherent visual and verbal information may aid learning, but the current study did not measure cognitive load, engagement, navigation, study time, or the mechanisms responsible for the score difference⁶. In practice, a lecturer could assign a small set of instruments before a demonstration, ask students to explain their intended functions, and then observe instrument selection and handling in a supervised setting. This is a proposed teaching application, not an intervention component verified by the present report. Mobile access may facilitate review, while device availability and use require assessment rather than assumption^{7,8}. Evidence on simulation and the distinction between knowing and performing supports retaining practical assessment as a separate educational requirement^{27,2}. The book should therefore supplement demonstration,

feedback, and safety instruction rather than serve as evidence of clinical competence or patient benefit.

Strengths include complete paired records for 161 students, a concurrent comparison, pretest adjustment, a continuous outcome, and reproducible individual level analyses. The accuracy of score calculations was confirmed against all 30 item correct answer counts. Nevertheless, important baseline differences in information sources and assignment by four intact classes create risks of confounding. The HC3 and student bootstrap intervals do not account for within class dependence, and two classes per condition cannot reliably characterize between class variation^{20,21}. The available materials do not document recruitment flow, delivery dates, matched teaching time, fidelity or contamination, instrument content validity, item reliability, or an established category cutoff. Reuse of the test may cause familiarity effects, and assessor blinding was not described. Information source timing was uncertain. The spreadsheet contains individual records but no student specific class identifiers; this prevents calculating class sizes or assessing intraclass correlation. Only immediate written knowledge from one institution was measured, so causal, retention, and clinical competence claims remain unsupported.

Before final submission, the authors should confirm recruitment and ethics chronology, intervention delivery and comparator equivalence, source measurement timing, digital book exposure, assessment instrument validation, and the originally intended knowledge category thresholds. If a confidential class identifier can be recovered, a statistical reviewer should examine defensible sensitivity analyses acknowledging the four cluster design; any resulting interval will still be uncertain with so few classes. Future studies should allocate a larger number of classes,

prospectively specify analyses and a meaningful educational threshold, document access and fidelity, and independently measure retained knowledge and observed instrument selection, handling, and safety. Secondary categorical results should remain subordinate to the continuous score and distinguish risks from odds^{22,23}.

Conclusion

In this nonrandomized four class study, INSTADAN exposure was associated with higher immediate posttest knowledge scores than lecturer PowerPoint material. The adjusted individual level difference was 6.08 points (HC3 95% CI, 3.99 to 8.17), approximately 1.82 of the 30 test items. Although the raw data and supplied analysis were reproducible, uncertainty remains because treatment was allocated by class, separate class identifiers were unavailable, and the instrument and educational importance threshold were not fully documented. These findings justify further evaluation of the resource, not claims of proven clinical skill, lasting knowledge, or patient benefit.

Acknowledgment

The authors thank Poltekkes Kemenkes Jakarta I, the Health Research Ethics Committee, lecturers, and students of the Midwifery Department for their support. They acknowledge the team's contributions to digital book development, instrument preparation, analysis planning, and manuscript preparation.

References

1. Vermeulen J, Luyben A, Jokinen M, Matintupa E, O'Connell R, Bick D. Establishing a Europe wide foundation for high quality midwifery education: the role of the European Midwives Association (EMA). *Midwifery*. 2018;64:128–131. doi:10.1016/j.midw.2018.06.009

2. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(9 suppl):S63–S67. doi:10.1097/00001888-199009000-00045
3. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet based learning in the health professions: a meta analysis. JAMA. 2008;300(10):1181–1196. doi:10.1001/jama.300.10.1181
4. George PP, Papachristou N, Belisario JM, et al. Online eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health. 2014;4(1):010406. doi:10.7189/jogh.04.010406
5. Tudor Car L, Poon S, Kyaw BM, et al. Digital education for health professionals: an evidence map, conceptual framework, and research agenda. J Med Internet Res. 2022;24(3):e31977. doi:10.2196/31977
6. Mayer RE. Applying the science of learning to medical education. Med Educ. 2010;44(6):543–549. doi:10.1111/j.1365-2923.2010.03624.x
7. Dunleavy G, Nikolaou CK, Nifakos S, Atun R, Law GCY, Tudor Car L. Mobile digital education for health professions: systematic review and meta analysis by the Digital Health Education Collaboration. J Med Internet Res. 2019;21(2):e12937. doi:10.2196/12937
8. O'Connor S, Andrews T. Smartphones and mobile applications (apps) in clinical nursing education: a student perspective. Nurse Educ Today. 2018;69:172–178. doi:10.1016/j.nedt.2018.07.013
9. Adelia D, Indarwati F, Rahmah R, Suminar IT. The effect of educational interventions through animated video and booklet on parents' knowledge regarding cognitive stimulation in preschool aged children. Healthy Tadulako Journal (Jurnal Kesehatan Tadulako). 2026;12(3):444–454. doi:10.22487/atvb7n86
10. Purwati Y, Andiani S. Penerapan komik digital pada edukasi kanker payudara terhadap minat pemeriksaan payudara sendiri. Healthy Tadulako Journal (Jurnal Kesehatan Tadulako). 2023;9(3):273–278. doi:10.22487/htj.v9i3.796
11. Subekti AT, Atmoko D, Miftah M. The effectiveness of video database based education in improving occupational safety and health literacy among construction sector MSME workers. Healthy Tadulako Journal (Jurnal Kesehatan Tadulako). 2026;12(2):307–315. doi:10.22487/dthpsz98
12. Azizah NPN, Adhi MHP, Suparti S, Fatwati NF. Comparison of self directed video and simulation methods on hand only CPR skills. Healthy Tadulako Journal (Jurnal Kesehatan Tadulako). 2025;11(2):265–274. doi:10.22487/htj.v11i2.1630
13. Des Jarlais DC, Lyles C, Crepaz N; TREND Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health. 2004;94(3):361–366. doi:10.2105/AJPH.94.3.361
14. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687. doi:10.1136/bmj.g1687
15. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J

- Med. 2006;119(2):166.e7–166.e16.
doi:10.1016/j.amjmed.2005.10.036
16. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53–55. doi:10.5116/ijme.4dfb.8dfd
17. Vickers AJ, Altman DG. Analysing controlled trials with baseline and follow up measurements. *BMJ.* 2001;323(7321):1123–1124.
doi:10.1136/bmj.323.7321.1123
18. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity score matched samples. *Stat Med.* 2009;28(25):3083–3107.
doi:10.1002/sim.3697
19. MacKinnon JG, White H. Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties. *J Econom.* 1985;29(3):305–325. doi:10.1016/0304-4076(85)90158-7
20. Cameron AC, Miller DL. A practitioner's guide to cluster robust inference. *J Hum Resour.* 2015;50(2):317–372.
doi:10.3368/jhr.50.2.317
21. Huang FL, Li X. Using cluster robust standard errors when analyzing group randomized trials with few clusters. *Behav Res Methods.* 2022;54(3):1181–1199.
doi:10.3758/s13428-021-01627-0
22. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ.* 2006;332(7549):1080.
doi:10.1136/bmj.332.7549.1080
23. Zou G. A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol.* 2004;159(7):702–706.
doi:10.1093/aje/kwh090
24. Morris SB. Estimating effect sizes from pretest posttest control group designs. *Organ Res Methods.* 2008;11(2):364–386.
doi:10.1177/1094428106291059
25. Galmarini E, Marciano L, Schulz PJ. The effectiveness of visual based interventions on health literacy in health care: a systematic review and meta analysis. *BMC Health Serv Res.* 2024;24:718.
doi:10.1186/s12913-024-11138-1
26. Morgado M, Botelho J, Machado V, Mendes JJ, Adesope O, Proença L. Video based approaches in health education: a systematic review and meta analysis. *Sci Rep.* 2024;14:23651. doi:10.1038/s41598-024-73671-7
27. Cook DA, Hatala R, Brydges R, et al. Technology enhanced simulation for health professions education: a systematic review and meta analysis. *JAMA.* 2011;306(9):978–988.
doi:10.1001/jama.2011.1234

Conflict of Interest Statement

The author(s) declare no commercial, financial, or personal conflicts of interest related to this research. All authors approved the final manuscript and consented to its publication in *Healthy Tadulako Journal*.

Copyright and Licensing

© Healthy Tadulako Journal. This open-access article is licensed under the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0), allowing use, distribution, and reproduction with proper attribution.

Publisher's Note

Healthy Tadulako Journal, a peer-reviewed open access journals prone published by the Quality Assurance Unit, Faculty of Medicine, Tadulako University, Indonesia.

