BUILDING SELF-CONFIDENCE IN EVIDENCE APPRAISAL AMONG PRECLINICAL MEDICAL STUDENTS THROUGH ABSTRACT INTERPRETATION: AN EXPERIENCE FROM AN INTRODUCTORY EPIDEMIOLOGY COURSE

Chayaporn Chanasuphaphathai, Ploypun Narindrarangkura***

***Medical student, Phramongkutklao College of Medicine, Bangkok, Thailand**
****Department of Military and Community Medicine, Bangkok, Thailand**

Abstract

Background: Evidence-based medicine (EBM) requires clinicians to identify, appraise, and apply research evidence in clinical decision-making, making research literacy a foundational competency in undergraduate medical education. However, these skills are rarely explicitly taught in medical curricula, leaving students underprepared to engage with scientific literature.

Objective: To evaluate changes in self-perceived academic reading competence among second-year medical students following a structured abstract interpretation assignment embedded in a short, intensive introductory epidemiology course.

Methods: A one-group pre–post design was employed with 105 second-year (preclinical) medical students in a compulsory one-credit epidemiology course in Bangkok, Thailand. The pre- and post-course assessments were administered approximately one week apart. Each student was assigned a peer-reviewed article, required to interpret the title and abstract using the PICO framework, and deliver a structured presentation assessed by a standardized rubric. Students' abstract interpretation performance was additionally assessed using the overall score from the faculty-assessed presentation rubric. A seven-item self-assessment instrument covering literature searching, article comprehension, summarization, critical analysis, and research confidence was administered before and after the course on a five-point Likert scale (maximum score = 35). Group differences were evaluated using independent-samples t-tests with Cohen's d as the effect size.

Results: Internal consistency was high at both time points (pre: $\alpha = 0.894$; post: $\alpha = 0.924$). The mean score for the abstract interpretation presentation was 7.20 (SD = 0.30) out of 7.50 points. The pre-course mean total score was 24.14 (SD = 3.92; 69.0% of maximum). Following the course, the mean increased to 30.70 (SD = 4.13), a gain of 6.55 points (27.1%), which was statistically significant ($t(208) = 11.80$, $p < .001$) with a large effect size (Cohen's d = 1.63). All seven competency items improved significantly ($p < .001$), with Cohen's d values ranging from 1.17 to 1.51.

Conclusion: A structured, individually accountable abstract-interpretation assignment embedded in a brief, intensive epidemiology course was associated with significant increases in students' self-perceived academic reading competence. These findings support integration of focused research literacy activities into early undergraduate medical curricula as a practical strategy for building EBM foundations.

Keywords: Medical Education, Evidence-Based Medicine, Epidemiology, Self-Assessment

J Southeast Asian Med Res 2026: 10: e0305
https://doi.org/10.55374/jseamed.v10.305

Correspondence to:
Narindrarangkura P, Department of Military and Community Medicine, 317 Ratchawithi Rd, Thung Phaya Thai, Ratchathewi, Bangkok 10400, Thailand
Email: n.ploypun@pcm.ac.th

Received: 17 June 2026
Revised: 3 August 2026
Accepted: 15 June 2026

Introduction

The principles of evidence-based medicine equip clinicians to identify, appraise, and integrate research evidence into medical decision-making, making research literacy a foundational competency for clinical practice.[1] For undergraduate medical students in particular, early engagement with the scientific literature has been shown to support the development of transferable skills, including information retrieval and academic writing, that are essential for both evidence-based practice and research.[2] Despite this recognized importance, academic reading skills are rarely taught explicitly within medical curricula. Medical education programs have traditionally included only limited research skills training, with little dedicated instruction in reading, comprehending, or critically evaluating published research.[3]

Critical appraisal, the systematic evaluation of research evidence to judge its validity, importance, and applicability to clinical practice, represents one of the highest-order demands placed on medical readers.[4] Clinical practice is increasingly dependent on the appropriate application of evidence-based medicine, and prospective medical professionals are expected to possess the skills to appraise research literature in order to make informed decisions for their patients.[5] Epidemiology, as the foundational science of population health, provides the scientific methods and conceptual framework for understanding disease patterns, causes, and their determinants.[6] Epidemiology courses at the undergraduate level offer students opportunities to build essential competencies, including critical thinking, quantitative literacy, and the ability to analyze public health information — skills that underpin evidence-based clinical practice and research throughout a student's career.[7]

Evidence from intervention studies confirms that structured teaching can produce meaningful and measurable gains in students' academic reading and critical appraisal abilities. Interventions in undergraduate medical programs have resulted in significant improvements in knowledge of research design and evidence appraisal, with one synthesis reporting a mean knowledge gain of approximately 17% following instruction—a considerably larger effect than that observed at the residency level.[8] Critical appraisal skills training delivered through workshop-based and multimedia approaches has similarly been shown to enhance students' ability to evaluate information critically, with the most positive effects observed in students' knowledge of evidence-based principles and self-confidence.[9] Self-assessment has emerged as a valuable adjunct within such programs, providing a structured means for students to recognize their own competence gaps and monitor their progress over time. Studies have found that prompting self-assessment in medical students increases their awareness of learning needs and supports self-regulated learning, thereby facilitating the development of reflective practice, which is essential for lifelong professional competence.[10]

The present study was conducted in a compulsory introductory epidemiology course for second-year medical students in Bangkok, Thailand. However, less is known about whether a brief, structured, and individually accountable abstract-interpretation activity embedded in an intensive introductory epidemiology course can improve students' perceived competence in engaging with academic literature. Therefore, this study was designed to evaluate changes in

second-year medical students' self-perceived competence in searching, reading, summarizing, interpreting, and critically engaging with academic articles before and after participation in the course. In addition, students' performance in the abstract interpretation presentation was assessed using a faculty-assessed presentation rubric.

Methods

Study design and setting

This study employed a one-group pre–post design to evaluate changes in students' self-perceived academic reading competence following participation in a one-credit introductory epidemiology course. The course is delivered over one intensive week comprising eight hours of didactic instruction, 17 hours of practical activities, and six hours of self-directed learning. It covers the epidemiological approach to disease and intervention, measurement of disease occurrence, measures of association, classification of epidemiological study designs, and control of communicable and non-communicable diseases. Teaching methods include interactive lectures, text-based discussions, research article-based discussions, and flipped-classroom activities, delivered across multiple sessions.

Participants

This study enrolled 105 second-year medical students in the course during the 2025 academic year. Enrolment in the course was compulsory, and all enrolled students completed both assessments, yielding a 100% response rate at both time points. No exclusion criteria were applied.

Teaching process

The abstract interpretation activity was the key practical component examined in this study. It was part of the final session of the course and accounted for 7.5% of the overall course grade. Each student was individually assigned a peer-reviewed medical research article and required to read and interpret only the title and abstract. Students then prepared a six-slide presentation delivered the following day, covering: (1) analysis of the article title using the PICO framework — Population, Intervention, Comparison, and Outcome[11], (2) background and significance, (3) interpretation of measures of disease occurrence, (4) interpretation of measures of association, (5) conclusions, and (6) citations formatted in both APA or Vancouver styles. Each presentation was limited to five minutes, followed by three minutes of questions. Performance was assessed using a structured rubric evaluating PICO evaluation, background, measures of disease occurrence and association, conclusions, referencing, and presentation skills, each rated on a four-point scale (0 = not mentioned, 1 = poor, 2 = fair, 3 = good, 4 = excellent).

Data collection

A seven-item self-assessment instrument was administered anonymously to all students via an online form immediately before the first session of the course (pre-assessment) and immediately after the final session (post-assessment). The instrument assessed students' self-perceived competence in academic article reading across seven domains: (1) searching for relevant articles, (2) understanding article structure, (3) reading and grasping key points from an abstract, (4) summarizing main content, (5) analyzing and interpreting study findings, (6) questioning and commenting critically on readings, and (7) confidence in using academic articles for study or research. Each item was rated on a five-point Likert scale (1 = needs improvement, 2 = fair, 3 = moderate, 4 = good, 5 = excellent), yielding a maximum possible total score of 35. The pre- and post-assessment forms were identical in content. In addition to the self-assessment instrument, faculty members assessed students' performance in the abstract interpretation presentation using a standardized rubric. Each rubric item was scored on a five-level scale from 0 to 4 (0 = not mentioned, 1 = poor, 2 = fair, 3 = good, and 4 = excellent). The raw rubric score was converted to a weighted score out of 7.50 points, corresponding to the presentation component's 7.5% contribution to the total course assessment. The weighted presentation score was analyzed descriptively as a performance-based outcome.

Statistical analysis

Descriptive statistics, including mean, standard deviation, minimum, and maximum, were calculated for each self-assessment item and for the total score at both time points. The overall abstract interpretation presentation score was summarized descriptively. Internal consistency was assessed using Cronbach's alpha for the pre- and post-assessment responses separately. Since the assessments were submitted anonymously and individual responses could not be matched across time points, group-level comparisons were performed using independent-samples t-tests to evaluate differences in item and total scores between the pre- and post-assessment groups. Effect size was calculated as Cohen's d and interpreted using conventional thresholds (small ≥ 0.20, medium ≥ 0.50, large ≥ 0.80). All analyses were conducted using R (version 4.4.0; R Foundation for Statistical Computing, Vienna, Austria)[12], with the psych package[13] used for reliability analysis and the effsize package[14] for effect size estimation. Statistical significance was set at $p < 0.05$.

Ethical considerations

This study was reviewed and approved by the Institutional Review Board, Royal Thai Army Medical Department (approval number M054h/68_Exp; approved December 30, 2025). All procedures were conducted in accordance with the Declaration of Helsinki. As the study used anonymous self-assessment data collected as part of routine educational evaluation, individual informed consent was not required. All data were de-identified at the point of collection, and participation did not affect students' academic standing or course grades.

Results

Instrument reliability

Internal consistency was acceptable at both time points. Cronbach's alpha was 0.894 for the pre-course assessment and 0.924 for the post-course assessment, indicating that the seven items formed a reliable scale on both occasions. The mean score for the abstract interpretation presentation was 7.20 (SD = 0.30) out of 7.50 points, corresponding to 96.0% of the maximum available marks. The narrow standard deviation indicates that students across the cohort performed the abstract interpretation task to a consistently high standard.

Pre-course self-assessment

Pre-course total scores ranged from 12 to 35 out of a possible 35, with a mean of 24.14 (SD = 3.92), representing 69.0% of the maximum. Only two students (1.9%) achieved the maximum total score of 35. Because this study used anonymized secondary educational data collected under the approved ethical protocol, individual pre- and post-assessment responses could not be linked. Therefore, it was not possible to determine whether these two students retained the maximum score in the post-course assessment. At the item level, mean scores ranged from 3.34 to 3.63 across the seven competency domains. The lowest-rated item was Item 6 (questioning and commenting critically on readings; Mean = 3.34, SD = 0.76), and the highest was Item 2 (understanding the structure of an academic article; Mean = 3.63, SD = 0.68). Full item-level descriptive statistics are presented in **Table 1**.

Post-course self-assessment

Following the course, total scores ranged from 21 to 35, with a mean of 30.70 (SD = 4.13), representing 87.7% of the maximum score. Thirty-five students (33.3%) achieved the maximum total score of 35. Item-level means increased across all seven domains, ranging from 4.27 to 4.60. Item 6 remained the lowest-rated domain (Mean = 4.27, SD = 0.82), while Item 2 remained the highest (Mean = 4.60, SD = 0.60).

Comparison of pre- and post-course scores

The mean total score increased by 6.55 points following the course, representing a relative gain of 27.1% over the pre-course mean. This difference was statistically significant ($t(208) = 11.80$, $p < .001$, Cohen's d = 1.63), indicating a large effect. All seven items showed statistically significant improvement ($p < .001$), with Cohen's d values ranging from 1.17 (Item 6) to 1.51 (Item 2), all exceeding the threshold for a large effect.

Table 1. Pre- and post-course self-assessment scores for each competency item (n = 105 per group)

Item	Competency	Pre- M (SD)	Post- M (SD)	Δ	t	Cohen's d
1	Searching for relevant articles	3.57 (0.65)	4.49 (0.65)	+0.91	10.19	1.41
2	Understanding article structure	3.63 (0.68)	4.60 (0.60)	+0.97	10.88	1.51
3	Reading and grasping abstracts	3.49 (0.72)	4.43 (0.68)	+0.94	9.68	1.34
4	Summarizing main content	3.41 (0.77)	4.41 (0.70)	+1.00	9.84	1.36
5	Analyzing and interpreting findings	3.37 (0.68)	4.28 (0.75)	+0.90	9.12	1.26
6	Questioning and commenting critically	3.34 (0.76)	4.27 (0.82)	+0.92	8.46	1.17
7	Confidence using articles for study/ research	3.36 (0.75)	4.31 (0.74)	+0.95	9.29	1.28
	Total score	24.14 (3.92)	30.70 (4.13)	+6.55	11.80	1.63

All comparisons $p < .001$. Scale: 1–5 per item; maximum total score = 35. Δ = mean difference (post minus pre).

The mean gain per item ranged from +0.90 (Item 5, analyzing and interpreting study findings) to +1.00 (Item 4, summarizing main content). **Table 1** summarizes item-level comparisons.

Discussion

This study demonstrated that participation in a one-week introductory epidemiology course incorporating an individual abstract interpretation assignment was associated with significant improvements in medical students's self-perceived academic reading competence across all seven assessed domains. Mean total score increased from 24.14 to 30.70 out of 35, a gain of 6.55 points (27.1%), with a large overall effect size (Cohen's d = 1.63). All item-level comparisons were statistically significant ($p < .001$), with effect sizes ranging from d = 1.17 to d = 1.51. These findings suggest that structured, individually accountable engagement with research literature, even within a short intensive unit, can produce meaningful gains in students' self-perceived reading competence.

The effect sizes observed in this study considerably exceed benchmarks reported in the broader reading self-efficacy literature. A meta-analysis of reading self-efficacy interventions across student populations reported a pooled pre–post effect size of g = 0.36 (95% CI 0.16-0.57).[15] The substantially larger effects observed here may reflect the individually accountable nature of the assignment. Each student was personally assigned an article, prepared an individual presentation, and responded to questions, which demanded genuine engagement with the text rather than passive reception of instruction. This is consistent with evidence that active, student-centered learning approaches produce larger gains in EBM-related competency than traditional lecture-based formats.[2]

The largest gains were in understanding article structure (Item 2, Δ = +0.97) and summarizing main content (Item 4, Δ = +1.00), competencies directly practiced during the assignment. Students were required to identify PICO elements from the article title and interpret the abstract section by section, providing structured practice in recognizing the organization and key components of academic articles. The high post-course mean for Item 2 (4.60/5) may also reflect a ceiling effect, as students approached the upper limit of the rating scale. This alignment between task demands and observed improvement is consistent with findings from critical appraisal training studies, in which the greatest gains are often observed in competencies most directly targeted by the instructional intervention.[9]

In contrast, Item 6, questioning and commenting critically on readings, had the lowest pre-course mean and remained the lowest-rated domain post-course. However, it still showed a substantial increase (Δ = +0.92). Critical que stioning is a higher-order analytical skill that

requires students to evaluate evidence, identify limitations, and formulate reasoned judgments rather than simply understanding or summarizing information. These skills may require repeated practice, feedback, and exposure to multiple research articles over time. Therefore, the relatively smaller gain observed for Item 6 may reflect the intervention's limited duration and single-activity nature rather than an absence of improvement. The pattern suggests that brief, structured activities may be particularly effective for foundational academic reading skills, whereas higher-order critical appraisal skills may require longitudinal, repeated instructional reinforcement.[3]

The pattern and magnitude of improvement are broadly consistent with evidence from journal club-based interventions in undergraduate medical education, which have reported significant gains in self-confidence, literature searching, and comprehension following structured article reading and presentation activities.[16] The abstract interpretation exercise shares structural features with journal clubs, individual assignment, structured presentation format, and peer questioning, but is more tightly scoped, targeting only the title and abstract rather than the full article. The fact that broad gains were nonetheless observed across all seven domains, including searching and confidence in using articles for research (Items 1 and 7), suggests that even constrained engagement with a single article can generate meaningful improvements in perceived competence beyond the immediate skills practiced. This study is also consistent with Norman and Shannon's finding that undergraduate-level instruction in critical appraisal yields approximately 17% knowledge gains, with undergraduate students responding more strongly to instruction than residents.[8]

This study achieved 100% participation from the entire cohort, avoiding selection bias, and the instrument demonstrated high internal consistency at both time points (Cronbach's α = 0.894 pre-course; α = 0.924 post-course). The item content was directly aligned with the assignment's targeted competencies, enabling meaningful interpretation of domain-specific change patterns. The study also adds to a limited evidence base on academic reading skill development in medical education in Southeast Asia, where English is the language of academic publishing, and limited prior exposure to research literature may pose distinctive challenges.[3]

This study encountered several limitations. The single-group pre–post design without a control group precludes causal inference. Although the study included both self-assessment and faculty-assessed performance data, these outcomes captured different aspects of learning. The self-assessment instrument measured changes in students' perceived competence before and after the course, whereas the presentation rubric provided a single post-intervention assessment of performance. Although the two outcomes could not be directly compared, the high rubric performance provides preliminary objective corroboration that the course supported genuine engagement with academic literature. Future studies should incorporate linked, objective pre- and post-performance measures to determine whether self-perceived gains correspond to measured competence. The outcome measure was self-assessed rather than objectively measured competence, and self-assessment in medical students is known to be subject to systematic bias, including the tendency for lower-performing students to overestimate their abilities.[17] The ceiling effect observed post-course, with one-third of students achieving the maximum score, suggests the scale may have lacked sensitivity at the upper range. Because the study used anonymized secondary educational data, individual pre–post responses could not be linked. This data structure, established in accordance with the approved ethical procedures, prevented individual-level longitudinal analysis and precluded determining whether the two students who achieved the maximum pre-course score also achieved the maximum post-course score. Finally, as a single-institution study, generalizability to other settings is uncertain, and the durability of the observed gains remains unknown.

Conclusion

This study provides evidence that a structured abstract-interpretation assignment, embedded in

a short, intensive introductory epidemiology course, is associated with significant improvements in medical students' self-perceived competence in searching, reading, and critically appraising academic literature. Gains were observed consistently across all seven assessed domains, with the most pronounced improvements in skills most directly practiced during the assignment. These findings support integrating focused, individually accountable academic reading activities into early undergraduate medical curricula as a practical, time-efficient strategy for building research literacy. Future studies employing objective outcome measures, control groups, and longitudinal follow-up are needed to confirm whether these self-perceived gains translate into durable improvements in evidence-based practice competency.

References

1. Ilic D, Maloney S. Methods of teaching medical trainees evidence-based medicine: a systematic review. Med Educ 2014; 48: 124–35.
2. Zee M, de Boer M, Jaarsma ADC. Acquiring evidence-based medicine and research skills in the undergraduate medical curriculum: three different didactical formats compared. Perspect Med Educ 2014; 3: 357–70.
3. Kling J, Larsen S, Thomsen SF. The need for focused literacy training in the medical school curriculum: a cross-sectional study of undergraduate students. Shankar PR, editor. Education Research International 2017; 2017: 7273824.
4. Greenhalgh T. How to read a paper: The basics of evidence-based medicine. 5th ed. John Wiley & Sons; 2014. 285 p.
5. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996 13; 312:71–2.
6. Gordis L. Epidemiology. 5th ed. Elsevier Saunders; 2014. 392 p.
7. Goldmann E, Stark JH, Kapadia F, McQueen MB. Teaching epidemiology at the under graduate level: considerations and approaches.Am J Epidemiol 2018; 187: 1143–8.
8. Norman GR, Shannon SI. Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: a critical appraisal. CMAJ. 1998; 158: 177–81.
9. Sasannia S, Amini M, Moosavi M, Askarinejad A, Moghadami M, Ziaee H, et al. Critical appraisal skills training to undergraduate medical students: A randomized control study. J Adv Med Educ Prof 2022; 10: 253–8.
10. Kennedy G, Rea JNM, Rea IM. Prompting medical students to self-assess their learning needs during the ageing and health module: a mixed methods study. Med Educ Online 2019; 24: 1579558.
11. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club 1995; 123: A12-13.
12. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for statistical computing; 2024 [cited 2026 Jun 17]. Available from: https://www.r-project.org/
13. Revelle W. psych: Procedures for psychological, psychometric, and personality research [Internet]. 2026 [cited 2026 Jun 17]. Available from: https://cran.r-project.org/web/packages/psych/index.html
14. Torchiano M. effsize: Efficient effect size computation [Internet]. 2020 [cited 2026 Jun 17]. Available from: https://cran.r-project.org/web/packages/effsize/index.html
15. Unrau NJ, Rueda R, Son E, Polanin JR, Lundeen RJ, Muraszewski AK. Can reading self-efficacy be modified? a meta-analysis of the impact of interventions on reading self-efficacy. review of educational research. 2018; 88:167–204.
16. Zhang H, Hu X, He C, Wang L, Yang M, Lu D. A journal club format that combines literature reading and clinical cases is more conducive to improving senior undergraduate medical students' overall competence. Adv Med Educ Pract 2025; 16: 279–86.
17. Knof H, Berndt M, Shiozawa T. Prevalence of Dunning-Kruger effect in first semester medical students: a correlational study of self-assessment and actual academic performance. BMC Med Educ 2024; 24: 1210.