Effectiveness of flipped classroom in medical education: Protocol study of a systematic review

Efficacité de la classe inversée dans l'enseignement médical : Protocole d'une revue systématique de littérature

Mariem Gaddour[1], Nedra Feni[2], Rihab Moncer[1], Rim Ghammem[3], Sonia Jemni[1], Helmi Ben Saad[4]

*MG, NF, RM, and RG contributed equally as first author
**SJ, and HBS contributed equally as senior authors
1. Faculty of Medicine of Sousse, university of Sousse, Sahloul University Hospital, Sousse, Tunisia
2. Faculty of Medicine of Sousse, university of Sousse, Ibn Jazzar University Hospital, Kairouan, Tunisia
3. Faculty of Medicine of Sousse, university of Sousse, Farhat Hached University Hospital, Sousse, Tunisia
4. Université de Sousse, Faculté de Médecine de Sousse, Hôpital Farhat Hached, Laboratoire de recherche LR12SP09 « Insuffisance cardiaque », 4000 Sousse, Tunisie

Abstract

Introduction: The flipped classroom (FC) integrates online content delivery with in-class activities focused on knowledge application, such as discussions and problem-solving. While increasingly prevalent in medical education field, the efficacy of FC remains a subject of debate.
Objective: This protocol outlines a systematic review (SR) to examine the scope and methodological quality of studies on the FC teaching approach in medical education.
Methods: The protocol was registered in PROSPERO (CRD420251008268) and adhered to the Preferred Reporting Items for SRs Protocols guidelines. A thorough literature search will be performed using Scopus, PubMed/Medline, the Cochrane Library, and Google Scholar. The review will consider comparative controlled studies that focus on FC in the medical learning field. Two independent reviewers will carry out a rigorous data extraction. The Medical Education Research Study Quality Instrument will be used to evaluate the study's quality. Synthetizing the findings of this SR will provide a thorough understanding of the impact of FC in academic performance and students' satisfaction in medical learning.
Expected results: This qualitative summary will offer a narrative overview of the included papers, emphasizing interventions designed to evaluate the impact of FC. Results will be discussed according to the review's objectives, evaluating both study quality and consistency of findings, and will be clearly illustrated using tables and graphs.
Conclusion: The SR developed from this protocol will undoubtedly provide evidence of the effectiveness of FC in medical education and will effectively contribute to increasing the present body of information on FC interventions for improving outcomes.
Registration: This protocol was duly registered in the International Prospective Register of SRs with the reference number CRD420251008268.
Keywords: Blended Learning, Efficacy, Flipped Classroom, Medical Learning, Postgraduate, Undergraduate

Résumé

Introduction: La classe inversée (CI) combine la diffusion de contenu en ligne avec des activités en classe axées sur l'application des connaissances, telles que des discussions et la résolution de problèmes. Bien que de plus en plus répandue dans l'enseignement médical, l'efficacité de la CI reste un sujet de débat.
Objectif: Ce protocole présente une revue systématique (RS) visant à examiner de façon exhaustive la qualité méthodologique des études publiées s'intéressant à l'application de la CI dans l'enseignement médical.
Méthodes: Ce protocole a été enregistré dans PROSPERO (CRD420251008268) et a respecté les lignes directrices PRISMA-P (Preferred Reporting Items for SRs Protocols). Une recherche bibliographique exhaustive sera effectuée dans Scopus, PubMed/Medline, la Cochrane Library et Google Scholar. La revue prendra en compte les études comparatives contrôlées portant sur la CI dans le domaine de la formation médicale. La synthèse des résultats de cette RS permettra de d'évaluer l'impact de la CI sur la performance académique et la satisfaction des étudiants en médecine.
Résultats attendus: Cette revue offrira une vue d'ensemble narrative des articles inclus. Les résultats seront discutés en fonction des objectifs de la revue, évaluant à la fois la qualité des études et la robustesse des conclusions, et seront clairement illustrés à l'aide de tableaux et de graphiques.
Conclusion: La RS élaborée à partir de ce protocole fournira certainement des preuves de l'efficacité de la CI dans l'enseignement médical et contribuera à enrichir les connaissances actuelles autour des effets de la CI pour l'amélioration des compétences médicales.
Enregistrement: Ce protocole a été dûment enregistré dans le Registre prospectif international des revues systématiques (PROSPERO) sous le numéro de référence CRD420251008268.
Mots-clés : Apprentissage hybride, Classe inversée, Efficacité, Pédagogie médicale

Correspondance
Mariem Gaddour
Faculty of Medicine of Sousse, university of Sousse, Sahloul University Hospital, Sousse, Tunisia
Email: gadmariem@gmail.com

LA TUNISIE MEDICALE-2025; Vol 103 (10): 1393-1397 DOI: 10.62438/tunismed.v103i10.6330

Introduction

Medical education refers to the organized system designed to equip qualified individuals with the knowledge and skills needed to become doctors (1). Graduates are expected to be able to handle the diverse responsibilities of a primary physician, such as patient care, medical practice, administrative roles, and adherence to ethical and legal standards (1). Recognizing the limitations of the traditional lecture in fostering deeper learning, many educational institutions have worked to reshape their class models to prioritize the application of knowledge and skill development (eg; critical thinking and communication) and adopt different teaching strategies (2). These strategies emphasize higher-order thinking and greater student participation (3). The flipped classroom (FC) is an example of an approach that has attracted significant attention (4). The traditional pedagogical approach relies on a teacher-centered lecture, where the instructor presents a lesson to an audience of students arranged in rows, attentively listening and taking notes (5).. Following the lesson, students are assigned homework or tasks to be accomplished outside the classroom. This method, often referred to as the 'sage-on-the-stage' model, emphasizes teacher-directed learning, and is also known as the transmission model (6). The limitations of the traditional didactic method become apparent when teachers are unable to dedicate sufficient time to individual student needs during class. This can result in students struggling to understand the material being left behind (2). The lecture-centered format, being predominantly a single-direction flow of information, severely restricts interaction between the teacher and the students, who tend to be passive learners. Additionally, students who become disengaged or unmotivated during the scheduled learning period pose a significant challenge (6). Besides, the accreditation process is placing a greater burden on institutions to demonstrate graduates' attainment of crucial skills, including the capacity to function in multidisciplinary teams, identify and solve problems, and communicate effectively (7). An innovative educational model, the FC, emphasizes active learning through collaborative activities held during class time, contrasting with traditional lecture-based instruction (7). Fundamentally, the FC reverses the typical allocation of in-class and self-study activities (8). In this approach, typical homework tasks like problem-solving and essay composition are transformed into interactive, teacher-facilitated classroom exercises, while lecture viewing or video study is completed at home (8).

Research on FC in health sciences has expanded in recent years, though it has largely concentrated on pharmacy and nursing disciplines (4), or in contrast, they focus on healthcare professionals in general (4). First, an interesting systematic review (SR) focusing on the effects of FC in medical education was published in 2017 (4). This SR established important knowledge, but it did not include more recent studies (ie; published after 2016) that could contribute to a revision or refinement of current understanding (4). The 2017 SR (4) included only 9 studies published between 2012 and 2016. Positive perceptions of the FC approach were noted, but its effects on knowledge and skill changes were inconclusive (4). Therefore, authors concluded that more substantial evidence is required to confirm its impact in these domains (4). Second, a 2025-SR and meta-analysis examining FC in medical field focused only on the impact of FC in clinical medicine, and the fundamental aspects were not explored (9).

Through a synthesis of current research, the present SR protocol aims to clarify the claims regarding the FC's ability to improve learning outcomes within medical education. The primary research question is what effect FC learning has on the academic achievement of medical professional students? The secondary study question focuses on the effects of FC learning on medical students' course satisfaction.

Methods

Study design

This SR's protocol is documented in PROSPERO, under the identification CRD420251008268. To ensure transparency and rigor, the SR protocol was constructed using the preferred reporting items for SRs and meta-analyses (PRISMA)-Protocol framework (Appendix A) (10). The final report will be structured according to the PRISMA guidelines (PRISMA statement) (11).

Study protocol

Review question

The SR will aim to answer the following primary research question: What is the impact of FC on medical students' learning outcomes?

Eligibility criteria

The PICOS (population, intervention, comparison, outcome, study design) model will be used to create research questions. The study's eligibility criteria will be defined as follows:

- Population (**P**): Medical students,
- Intervention (**I**): Implementation of the FC teaching methodology,
- Comparison (**C**): Traditional learning,
- Outcome (**O**): Students' academic performance and perception,
- Study design (**S**): Only comparative studies (ie; randomized clinical trial) will be included.

Only research published in English language will be considered. There will be no limits on the study setting or country of origin. Non-comparator observational studies, case reports, case series, conference abstracts, editorials, letters to the editor, comments, and studies that do not include a relevant population, intervention, or outcomes will be eliminated.

Search strategy

Four prominent databases (ie, PubMed/Medline, Scopus, Cochrane Library, and Google Scholar) will be systematically searched electronically, starting from 2012 publication dates. The search will be periodically updated

to guarantee that all pertinent data is included prior to the study's conclusion.

To personalize search strategies for each database, we will use a combination of medical subject headings and synonyms relating to medical education, students, and FC. The Boolean operators 'AND' and 'OR' will be utilized.

Appendix B (12) contains the whole database search approach. In Google Scholar, search results will be ordered by relevancy, and just the first ten pages will be examined using the same criteria. In addition to exploring electronic databases, the reference lists of all papers included will be thoroughly examined to identify any additional relevant studies.

Study selection

Using bibliographic software (ie; Zotero), search results from all databases will be integrated and duplicates will be eliminated. Two reviewers (MG and NF on the authors' list) will separately screen the research, with any disagreements resolved by discussion or, if necessary, by a third author (RM).

The initial assessment will include a title review to determine its relevance to the research issue. Abstracts will then be examined to verify they fit the inclusion criteria. Full-text reviews will be carried out on publications whose abstracts match these criteria to determine their eligibility for the SR. The complete research selection technique will be illustrated using the PRISMA flow diagram (Figure 1).

The domain refers to the reported learning outcomes in medical education: The four levels of educational outcome evaluation are reaction (satisfaction and perceived outcomes), learning (direct measures of outcomes such as knowledge, skills, and attitudes), behavior (behavioral changes resulting from the intervention), and results (organizational changes resulting from the intervention) (13).

Data extraction

No artificial intelligence tools will be used for screening or data extraction. The latter will be performed manually by two reviewers (MG and NF in the authors' list) ensuring the rigor and reliability of our work.

For each study, we will extract the following data:

- Study identification: Authors, year of publication, country.
- Participants' characteristics including number, levels.
- Study design: Sample size calculation, setting, randomization method, blinding technique.
- Intervention details: Students access course content through pre-recorded lectures or online resources. Interactive exercises, collaborative projects, discussions, and the practical application of knowledge will be the focus of face-to-face class time (14).
- Comparison group (traditional classroom): During class, teachers use lectures to passively transmit foundational knowledge to students, who will then use and apply their knowledge outside of the classroom setting (4).
- Outcomes: Effectiveness of FC and effect size of the main outcomes.

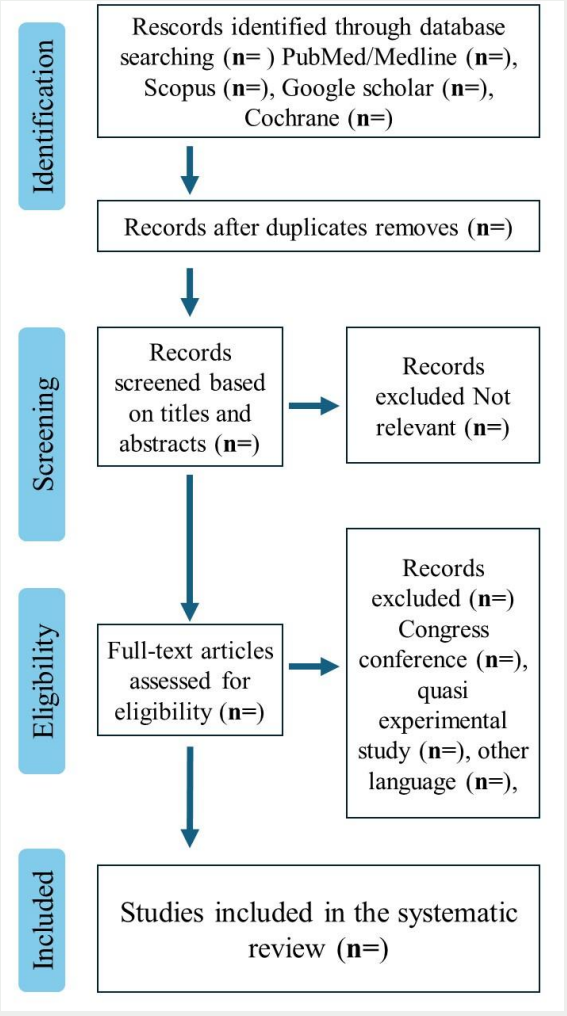

Figure 1. PRISMA flow diagram.

The modified Kirkpatrick framework will be used to classify research outcomes of educational interventions based on their impact level (17). The modified Kirkpatrick classification contains six levels:

- Level 1: learner perceptions
- Level 2a: self-reported changes in opinions
- Level 2b: modifications in knowledge or skills
- Level 3: improvements in learner behavior
- Level 4a: shifts in professional practice; and
- Level 4b: changes in patient condition.

In fact, competence in medical training is a multifaceted concept (16). It encompasses a range of skills, knowledge, and attitudes necessary for effective clinical practice (16). Therefore, measuring competence is challenging because it is not simply a quantifiable trait. Consequently, a single exam mark is insufficient to truly evaluate competence (16). Therefore, we adopted the modified Kirprick frame to measure the effectiveness of FC.

In case of no information is provided, we will try to extract the information required to calculate the associated effect sizes for each outcome.

Initially, two authors (MG and NF in the authors' list) will extract data independently using a predesigned data extraction form in Microsoft Excel to assure precision and consistency. Disagreements between the two authors will be resolved by debate or, if necessary, by a third author (RM on the author list).

Methodological quality assessment

The medical education research study quality instrument (MERSQI) will be used to assess the quality of retained studies (Table 1) (17).

Table 1. Medical education research study quality instrument (MERSQI) items.

Domain: Item	Response options: Scores
MERSQI	
Study design	• Non randomized, 2 groups: 2 • Randomized controlled trial: 3
Sampling: institutions	• 1 institution: 0.5 • 2 institutions: 1 • ≥3 institutions: 1.5
Sampling: Rresponse rate	• Not applicable • < 50% or not rated: 0.5 • 50%-74%: 1 • ≥75%: 1.5
Type of data	• Assessment by participants: 1 • Objective: 3
Validity evidence for evaluation institution scores	• Not applicable • Content: 1 • Internal structure: 1 • Relashionships to other variables: 1
Data analysis: Sophistication	• Descriptive analysis only: 1 • Beyond descriptive analysis: 2
Data analysis: Appropriate	• Data analysis appropriate for study design and type of data: 1
Outcome	• Satisafction, attitudes, perceptions, opinions, general facts: 1 • Knowledge, skills: 1.5 • Behaviors: 2 • Patient/health care outcome: 3

This 10-item scale measures the methodological quality across six domains: study design, sampling, type of data, validity evidence, data analysis, and results (17). The "sampling" and "data analysis" domains contain two items each. Each of the three sources of validity evidence (content, internal structure, and relationships to other variables) is considered as a separate item. Total MERSQI scores span from 2 (low-quality research) to 18 (high-quality research). Two authors (MG and NF in the authors' list) independently will score each retained article, and in case of disagreements, a third author (RM in the authors' list) will take the decision.

Missing data

If an article lacks or contains insufficient data, the retained paper's corresponding author will be emailed. If contact is not established after two attempts, or if the data remains insufficient, the paper (or the data concerned) will be omitted from the final analysis. However, if appropriate, an imputation approach will be used to address missing data.

Statistical analysis

The Data from the included studies will be synthesized quantitatively using a random-effects model to account for variability both within and between trials. The effect sizes will be calculated for each outcome, such as satisfaction and perception of students, and given as standardized mean differences with 95% confidence intervals.

Expected results

This qualitative synthesis will offer a narrative overview of the included studies, emphasizing interventions designed to evaluate the impact of FC. Methodological characteristics of studies, such as study design, topic, and level, will be summarized in Table 2.

Table 2. Summary of controlled studies on flipped classroom in medical education (n=).

Author	Year	Topic	Design	Intervention	Level	Kirkpatrick measure of effectiveness*	Findings on effectiveness in kirkpatrick framework	Effect size [95% CI] † on changes of knowledge and skills (level 2b)

Kirkpatrick Framework (levels):
1: Perceptions of intervention,
2a: Attitude changes,
2b: Changes of knowledge and skills,
3: Changes in behaviors,
4a: Changes in professional practice,
4b: Changes in patient outcome.
† Effect sizes in Cohen's d will be calculated for changes of knowledge and skills (level 2b).
95% CI = 95% confidence interval of the d estimate.

The synthesis will detail the interventions and identify patterns, similarities, and differences across studies. Results will be discussed in relation to the SR's objectives, considering study quality and consistency of findings, and will be presented clearly using tables and graphs.

This SR will likely demonstrate whether FC has a major impact on medical student's learning outcomes and does effectively improve their academic scores. It is expected that FC, being a competency-based education model, will align well with the characteristics of medical students, who are generally driven and self-directed learners. This pedagogical approach fosters in-depth concept discussions and practical knowledge application to clinical scenarios, thereby encouraging advanced cognitive skills and replicating real-world clinical settings.

Ethics and dissemination

The SR will utilize publicly available literature and data. Hence, this research is exempt from ethical review. The findings will be published in a peer-reviewed journal. Any changes to the protocol will be documented in the PROSPERO database, and the updated version will be included in the final SR. This study is currently in the pre-initiation phase. The results will be published in a peer-reviewed journal. Any changes to the protocol will be registered in the PROSPERO database, and the updated version will accompany the final SR. This study is still in its pre-initiation phase. We plan to start the initial stage on 1 September 2025. We expect to complete the search, screening, data extraction, and synthesis processes by late December 2025.

Appendices

*Appendix A. PRISMA-P for the study titled "Effectiveness of flipped classroom in medical education: Protocol study of a systematic review". 2025. Link: https://zenodo.org/records/15545974 (Last visit: September 8, 2025).

*Appendix B. Search strategy for the systematic review protocol titled "Effectiveness of flipped classroom in medical education: Protocol study of a systematic review". 2025. Link: https://doi.org/10.5281/zenodo.17076919 (Last visit: September 8, 2025).

CONFLICTS OF INTEREST. No competing interests were disclosed

ACKNOWLEDGMENTS. The authors would like to express their sincere gratitude to the reviewer for his/her excellent feedback, which has substantially improved the quality of this work. His/her insightful comments and constructive suggestions were invaluable in refining our manuscript (18). Authors express their gratitude to their families and especially their mothers for their unwavering support and sacrifices. The authors acknowledge the use of artificial intelligence tools (eg; Gemini, ChatGPT Ephemeral, and QuillBot) to support the refinement of language throughout the manuscript. These tools were used solely to enhance clarity and coherence in the writing, without modifying the scientific content or producing original material (19).

References

1. Shafei I. Medical Education. In: Loh E, Long PW, eds. Textbook of medical administration and leadership. Singapore: Springer Nature Singapore 2023:413-27.
2. Adam S, Katsiris I, Michalis L, et al. The applied model approach in a skills-based education. EJENG 2023;1(CIE):78-86
3. Haryadi R, Situmorang R, Khaerudin K. Enhancing students' high-order thinking skills through stem-blended learning on kepler's law during Covid-19 outbreak. J Penelit Pembelajaran IPA 2021;7(2): 168-92.
4. Chen F, Lui AM, Martinelli SM. A systematic review of the effectiveness of flipped classrooms in medical education. Med Educ 2017;51(6):585-97.
5. Ghafar ZN. The teacher-centered and the student-centered: A comparison of two approaches. IJAH 2023;1(1):18-23.
6. Morrison CD. From 'Sage on the stage' to 'guide on the side': a good start. Int J Scholarsh Teach Learn 2014;8(1):1-15.
7. Moquin R, Dewey M, Weinhold A, et al. The flipped classroom approach in a pediatric anesthesiology fellowship curriculum. Cureus 2023;15(8):e43979.
8. Vaddatti T, Chaitra B, Renuka IV, et al. The perceptions of II MBBS students toward flipped classroom as an active method of learning. Dent Med Res 2022;10(2):60-3.
9. Shi X-Y, Yin Q, Wang Q-W, et al. Is the flipped classroom more effective than the traditional classroom in clinical medical education: a systematic review and meta-analysis. Front Educ 2025;9:1485540.
10. Gaddour, M. Appendix A. PRISMA-P for the study titled "Effectiveness of flipped classroom in medical education: Protocol study of a systematic review". 2025. Link: https://zenodo.org/records/15545974 (Last visit: September 8, 2025).
11. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. PLoS Med 2021;18(3):e1003583.
12. Gaddour, M. Appendix B. Search strategy for the systematic review protocol titled "Effectiveness of flipped classroom in medical education: Protocol study of a systematic review". 2025. Link: https://doi.org/10.5281/zenodo.17076919 (Last visit: September 8, 2025).
13. Holtschneider ME, Park CW. Evaluating simulation education in the workplace: Kirkpatrick's levels and beyond. J Nurses Prof Dev 2019;35(1):44-5.
14. Ben Kahla K, Khanchel H. Assessing the effects of flipped classroom at the Tunisian university. Indep J Manag Prod 2021;12(5):1413-35.
15. DeSilets LD. An update on Kirkpatrick's model of evaluation: Part two. J Contin Educ Nurs 2018;49(7):292-3.
16. Ng IKS, Mok SF, Teo D. Competency in medical training: current concepts, assessment modalities, and practical challenges. Postgrad Med J 2024: 38376156.
17. Cook DA, Reed DA. Appraising the quality of medical education research methods: the medical education research study quality instrument and the Newcastle-Ottawa scale-education. Acad Med 2015;90(8):1067-76.
18. Hidouri S, Kamoun H, Salah S, Jellad A, Ben Saad H. Key guidelines for responding to reviewers. F1000Res. 2024 Sep 20;13:921.
19. Dergaa I, Zakhama L, Dziri C, Ben Saad H. Enhancing scholarly discourse in the age of artificial intelligence: A guided approach to effective peer review process. Tunis Med. 2023;101(10):721-6.