



## OPEN ACCESS

## EDITED BY

Ana Teresa Ferreira Oliveira,  
Instituto Politécnico de Viana do Castelo,  
Portugal

## REVIEWED BY

Jessica Borger,  
Monash University,  
Australia  
Keerti Singh,  
The University of the West Indies, Cave Hill,  
Barbados

## \*CORRESPONDENCE

Mona Mlika  
mouna.mlika@fmt.utm.tn

## SPECIALTY SECTION

This article was submitted to  
Higher Education,  
a section of the journal  
Frontiers in Education

RECEIVED 04 September 2022

ACCEPTED 18 October 2022

PUBLISHED 03 November 2022

## CITATION

Mlika M, Naceur A, Dziri C, Loueslati MH,  
Mezni F, Zakhama L, Labbene I and  
Jouini M (2022) Critical appraisal of  
medical literature in undergraduate and  
postgraduate medical students.  
*Front. Educ.* 7:1036627.  
doi: 10.3389/feduc.2022.1036627

## COPYRIGHT

© 2022 Mlika, Naceur, Dziri, Loueslati,  
Mezni, Zakhama, Labbene and Jouini. This  
is an open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Critical appraisal of medical literature in undergraduate and postgraduate medical students

Mona Mlika<sup>1\*</sup>, Abdelmajid Naceur<sup>2</sup>, Chadli Dziri<sup>3</sup>,  
Mohamed Hédi Loueslati<sup>3</sup>, Faouzi Mezni<sup>1</sup>, Lilia Zakhama<sup>3</sup>,  
Iheb Labbene<sup>3</sup> and Mohamed Jouini<sup>3</sup>

<sup>1</sup>Department of Pathology, Abderrahman Mami Hospital, University of Tunis El Manar, Research Unit: LR18SP06, Faculty of Medicine, Tunis, Tunisia, <sup>2</sup>Virtual University of Tunis, Research Unit ECOTIDI (UR 16E S10), Tunis, Tunisia, <sup>3</sup>University Tunis El Manar, Faculty of Medicine, Tunis, Tunisia

**Background:** Critical appraisal of medical literature is a challenging competency which is mandatory in the practice of evidence-based medicine. Achieving this competency has been reported to vary according to many factors including the students' background, gender, or specialty. The majority of the articles published in this field mainly assessed the students' satisfaction or their cognitive knowledge. The authors aimed to assess the possibility of performing the same training of critical appraisal of medical literature to students with different backgrounds to highlight the link between the critical appraisal practice and the reflexive, cognitive, attitudes and conative profile of the students according to their level.

**Methods:** Students in the second year (SYME) and third year (TYME) of medical education and students performing continuing medical education (CME) and family doctors (FD) were invited to voluntarily participate in the training. This study occurred from September 2020 to June 2022. Before the workshop, the students were invited to answer a self-assessment questionnaire and a prerequisite test. Two original manuscripts dealing with COVID-19 were analyzed using published checklists. At the end of the workshops, all of the students fulfilled a second self-assessment questionnaire and a Fresno-adapted final test assessing their knowledge. Both self-assessment questionnaires were conceived by an expert committee according to the different steps of questionnaires' validation. They assessed the cognitive, conative, motivational, and reflexive students' profile using Likert scale questions. Also, the participants were invited to answer a semi-structured interview.

**Results:** Ninety-five participants were included with 6 participants in SYME, 54 in TYME, 16 in CME, and 19 FD. The results highlighted the possibility of using the same program independently from the students' level when using manuscripts dealing with a widely known and confusing disease such as COVID-19. The analysis of the different pre- and postworkshop questionnaire scores highlighted mild or negative scores in the TYME and intermediate or good scores in the other levels. The variation of the mean scores according to the students' level revealed a significant difference in prerequisite, final test, pre and postworkshop attitudes, postworkshop cognitive, postworkshop conative and satisfaction scores. 11 students from different levels were interviewed. The

content analysis highlighted 4 themes including the teaching organization, the institutional assessment, the impact on the research practice and the impact on the patients' management with the emphasis of the TYME students on the necessity of integrating the teaching and the assessment of the EBM principles in the curriculum.

**Conclusion:** Our results put emphasis on the necessity of introducing the teaching of critical appraisal of medical literature early in the curriculum in order to avoid negative behaviorism during the clerkship period.

#### KEYWORDS

evidence-based medicine, evidence-based medicine practice, critical appraisal of medical literature, motivational determinants, students profile

## Background

Evidence-based medicine (EBM) is a widely used concept that aims to practice medicine according to the literature data (Parkes et al., 2001; Jouquan, 2009; Goichot and Meyer, 2011; Le Glatin, 2013). Solving medical problems is a mainstay in medicine and necessitates linking medical practice to scientific background. The COVID-19 crisis highlighted a misuse of the major principles of critical appraisal of medical literature, especially when practitioners promoted the use of some drugs without significant evidence. This finding puts emphasis on the necessity of achieving critical appraisal skills in medical practice. Critical appraisal of medical literature is integrated into the evidence-based medicine process. This process contains five steps that consist of facing a clinical problem, formulating an accurate and clear question related to the problem, critical appraisal of medical literature, assessment of the validity and applicability of the articles' conclusions, and global assessment of the process (Evidence-Based Medicine, Part 3, 1995; Evidence-Based Medicine, Part 5, 1995; Evidence-Based Medicine, Part 6, 1995; Evidence-Based Medicine, Part 2, 2007; Ilic, 2009). The third step of the EBM process consists of the critical appraisal of medical literature. It was the most assessed in the literature because it needs cognitive and technical abilities (Parkes et al., 2001; Ilic, 2009; Rohwer et al., 2013; Hryciw et al., 2017; Mink et al., 2019). The critical appraisal of medical manuscripts necessitates two different skills that are intimately linked: assessment of the evidence's hierarchy and critical decision making. A continuum exists between both competencies that are mandatory for medical students to achieve aptitude in medical practice by the Royal College of Physicians and Surgeons of Canada and the Accreditation Council for Graduate Medical Education (Edgar et al., 2003; Frank et al., 2015). Many articles dealing with critical appraisal competencies are available in the literature. They mainly assessed either the students' satisfaction or the cognitive skills. Also, despite the consensus about the necessity of teaching critical appraisal practice, few publications dealt

with the way of teaching undergraduates in the first cycle of medical education, who have no clinical background, the second cycle of medical studies, who are initiated to clinical clerkship and postgraduate students who have to achieve specialized skills. In these different levels, the students have different backgrounds in terms of scientific and preventive medicine knowledge (Godwin and Seguin, 2003a; Marusić and Marusić, 2003; Horsley et al., 2011; Zins et al., 2011; Hryciw et al., 2017). The competency of critical appraisal of medical literature concerns many domains including the knowledge and the know-how. This concept has been promoted in order to improve medical skills, management of the patients, and validity of the treatment process. In our Faculty of Medicine, EBM principals aren't included in the curriculum. Many courses dealing with methodology, statistical tests, and preventive medicine are dispensed during the first (theme 7) and second years (theme 15) of medical studies, but they aren't linked to the EBM practice. When dealing with this work, the authors' aims consisted of assessing the possibility of performing the same training for students with different backgrounds, highlighting the link between the critical appraisal practice and critical thinking, reflexive practice and the students' ability to integrate the principles of the critical appraisal practice independently from their learning levels.

## Materials and methods

In order to achieve these goals, the authors performed a longitudinal, prospective, descriptive, and qualitative study including students from the same Faculty of Medicine. This study occurred from September 2020 to June 2022. The quantitative approach assessed the cognitive, reflexive, conative students' profiles and attitudes. The qualitative study was performed in order to highlight other dimensions influencing the critical appraisal practice that have not been assessed in the quantitative study.

TABLE 1 The workshop's detailed program.

1. Self-assessment questionnaire
2. Lecture 1: why do we need to read?
3. Correction of the prerequisite test: we red for you
4. Lecture 2: Evidence-based medicine
5. Different types of manuscripts
6. Exercise 1: Situation-problem-based learning: clinical scenario + critical appraisal of the first manuscript: Hydroxychloroquine in Nonhospitalized Adults With Early COVID-19. A Randomized Trial.
7. Exercise 2: Situation-problem-based learning: clinical scenario + critical appraisal of the second manuscript: Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis.
8. Self-assessment questionnaire + satisfaction questionnaire
9. Final test

- **Population:** Students from the same Faculty of Medicine including undergraduate and postgraduate students.
- **Inclusion criteria:** Undergraduate and postgraduate students wishing to participate to the study were included.
- **Non-inclusion criteria:** Students that did not belong to the same Faculty of Medicine were not included.
- **Exclusion criteria:** Students who did not answer the different tests and questionnaires were excluded.
- **Learning activities:** The students were invited to participate in a 9-h workshop centered on the critical appraisal of medical literature. All of the workshops were tutored by the same tutor. Before the workshop, the students were invited to answer a self-assessment questionnaire and a prerequisite test. The authors chose the 2 following original articles dealing with COVID-19:

1. Hydroxychloroquine in non-hospitalized adults with early COVID-19. A randomized trial. Skipper CP, Pastick KA, Engen NW, et al. *Ann Intern Med.* doi:10.7326/M20-4207 (Skipper et al., 2020).
2. Diagnostic accuracy of serological tests for COVID-19: systematic review and meta-analysis. Bastos ML, Tavaziva G, Abidi SK, et al. *BMJ* 2020;370:m2516 (Lisboa Bastos et al., 2020).

The choice of manuscripts dealing with COVID-19 was due to the fact that every student, independently of the level, was aware of the different therapeutic and diagnostic challenges faced during this pandemic. Table 1 illustrates the workshop's program with the lecture details. The students were asked to perform a critical appraisal of both manuscripts in order to answer particular clinical situation problems that were presented, using checklists available online. Supplementary Table S1 represents both situation problem-based learning that were presented in order to introduce both original articles. For each manuscript, the students were asked to assess the validity of the study, based on the methods section, to analyze the results and determine whether or not to

apply the results to the situation problem. They were asked to fulfill 3 questionnaires for every manuscript.

## The prerequisite test

The prerequisite test was designed in order to assess the students' background in biostatistics. Biostatistics knowledge is necessary in order to appraise the methods' sections of the different manuscripts. According to the Faculty's curriculum, the students are taught biostatistics principles during the first 2 years of medical education. All of the students included in this study were supposed to have achieved the same objectives concerning biostatistics. The prerequisite test contained questions related to the methods' sections of published articles. It contained nine multiple-choice questions assessing the following concepts: the confidence interval, the significance of statistical tests, the judgment of causality, the comparability of groups, the number of participants needed to treat calculation in a clinical trial, the management strategy of the patients lost of view in a clinical trial, the sensitivity, specificity, positive predictive value and negative predictive value of a diagnostic test, and the judgment criteria in a study. Supplementary Table S2 illustrates the prerequisite test performed. The students were asked to answer the prerequisite test that was sent to them through Google Forms. At the beginning of the workshop, the questions were corrected and the concepts were clarified by the tutor.

## The questionnaires fulfilled in order to perform the critical appraisal of the manuscripts during the workshop

The questionnaires concerning the first manuscript, consisting of a clinical trial, were inspired from the CONSORT checklist (CONSORT, 2010). The students were asked to answer 3 questionnaires through Google Forms related to the validity of the methods' section, the results, and the applicability of the results to the situation problem. The three questionnaires are represented in Supplementary Tables S3–S5. The questionnaires concerning the second manuscript, consisting of systematic review of the literature with a meta-analysis, were inspired from the AMSTAR 2 checklist (Shea et al., 2017). They also dealt with assessing the validity of the study, the results section, and the applicability of the results to the situation problem. Supplementary Tables S6–S8 illustrate the 3 questionnaires fulfilled by the students.

## The self-assessment questionnaires

The students were asked to fulfill 2 self-assessment questionnaires before and after the workshop. They consisted of Likert scale questions and were performed by an expert committee with 3 full professors who were used to teach critical

appraisal principles. The different steps of validation of a questionnaire published by Tsang and co-workers were followed (Tsang et al., 2017). The expert committee met once a month for 3 months in order to assess the literature and prepare the questionnaires. They performed a review of the literature and tried to assess the different parameters or dimensions influencing the critical appraisal of medical literature practice. They established that the learning strategies or cognitive profile of the students, their attitudes, their reflexive attitude, and their conative profile may influence their behaviors. Learning strategies or cognitive strategies consist of the students' tendency to use either active pedagogy principles, making the students play a central role in their learning process, or traditional pedagogy, which puts the teacher in the center of the learning process. Reflexive attitude is defined as the students' capacity to identify proper learning needs when facing a problem or a task. Reflexive attitude is linked to self-directed learning and autonomy. Conative profile is defined as the propensity to apply the knowledge achieved. In order to assess these different determinants, the expert committee analyzed and used published questionnaires (Shehata et al., 2015; Maloney et al., 2019; Nieminen et al., 2020). After the identification of the construct's dimensionality, the expert committee identified and determined the format of the questionnaires and distributed them through Google Forms. They developed the items, determined the questionnaires' length, and reviewed and revised the initial items pool. A preliminary pilot study was conducted including 30 students in third year of medical education (TYME). The unclear items pointed out by the students were modified, and the expert committee validated the questionnaires. The internal consistency, which reflects the extent to which the questionnaire items are inter-correlated or whether they are consistent in the measurement of the same construct, was assessed using the coefficient alpha. A Cronbach's alpha of at least 0.7 indicates adequate internal consistency. Supplementary Tables S9, S10 illustrate the different questionnaires with the different coefficients. The final version of the questionnaire was assessed by 2 experts used to teach the critical appraisal of medical literature. They were asked to judge whether the questionnaire items were adequately measuring the intended construct to assess and whether the items were sufficient to measure the domains of interest. We choose the process of content validation, and the expert were asked whether the questions were clear and easy, whether they covered all determinants of critical appraisal practice, whether they would like to use the questionnaires for future assessment, and whether they lack important questions. The experts reviewed the questions and did not add modifications.

The final preworkshop questionnaire contained 13 Likert scale questions assessing the cognitive and learning profile of the students and their reflexive, conative profiles and attitudes. The final postworkshop questionnaire contained 27 Likert scale questions assessing the cognitive and conative profiles of the

students in addition to their satisfaction and attitudes. Cognitive domain assessed the self-perception of the students concerning their abilities to perform a critical appraisal of the medical literature after the workshop. Attitudes domain assessed the students' feelings and affect after the workshop, and conative domain assessed the students' wish to appraise medical literature in the future. Supplementary Tables S11, S12 illustrate the pre- and postworkshop questionnaires, respectively, with the rating details.

## The final test assessing the new knowledge

The final test was inspired from the Fresno test (Halm, 2018). The Fresno test is a consensual published test assessing medical residents' knowledge of basic evidence-based medicine principles, including how to frame a research question, how to search for evidence to answer the question, understanding the hierarchy of evidence, being able to interpret its magnitude, internal and external validity of the evidence, and basic and statistical concepts. It contains 7 short-answer questions, 2 questions that require a series of mathematical calculations, and three fill-in-the-blank questions. All of the questions are rated in details. We modified the Fresno test because we were not assessing all of the evidence-based medicine practice steps. The workshop focused on the critical appraisal of medical literature. For that reason, we included 5 short-answer questions from the Fresno test that were related to the critical appraisal of medical literature, and we added 6 short-answer questions related to an original manuscript part dealing with a diagnostic test. The final tests were rated by the same tutor. The final test is presented in Figure 1.

Figure 2 illustrates the scenario of the 9-h workshop.

## The interview following the workshop

Three months after the workshop, the participants were invited to answer a semi-structured interview. The interviewer asked only a few predetermined questions, while the rest of the questions were not planned in advance. The predetermined questions concerned the workshop's impact felt by the students, their feelings about the training, their self-confidence about their competencies in that matter, and the improvement in the management of their patient secondary to this workshop. All of the interviews were recorded with the consent of the participants. We proceeded to a content analysis following 3 steps: pre-analysis, treatment of the results, and interpretation.

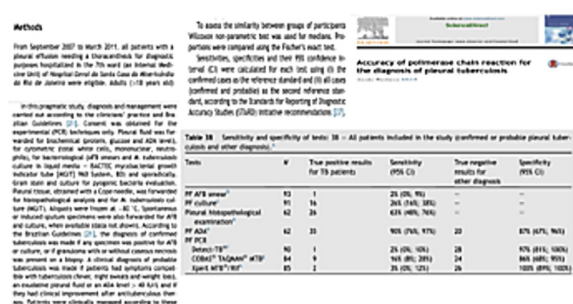
## Statistical tests

Quantitative data related to the prerequisite and final test scores were represented as means. Qualitative data related to the self-assessment questionnaires and post-workshop



**Name and surname:**  
**E-mail:**

1- You are intern in a Pulmonology department and you're taking care of a 35-year-old patient presenting a pleural effusion. You suspected a pleural tuberculosis. A pleural puncture was performed and addressed to the Microbiology Department. A kit (detect-TB) PCR was performed and was negative. Your chief of Department asked you if you ruled out the diagnosis of tuberculosis based on this test. In order to answer to his question, you performed a literature review and found an evidence entitled: Accuracy of polymerase chain reaction for the diagnosis of pleural tuberculosis. Based on the methods section of this article, answer to the following questions:



1-1 Is the test compared to a standard test?  - Yes  - No

1-2 What is the gold standard test?

1-3 Did the patients represent the patients you have to manage in your routine practice and to which you're going to use the test?  - Yes  - No

1-4 Test PCR (detect-TB) sensitivity :

1-5 Test PCR (detect-TB) specificity :

1-6 According to these data, are you going to perform a pleural biopsy or is the PCR test sufficient?

- Pleural biopsy  
 - PCR (detect-TB) sufficient

2- Miss MM presented to your consultation after a delivery. She is breast feeding her child but wants to start an oral contraception. You used to prescribe a combination of contraception associating oestrogens and progesteron but you red recently that this association could interfere negatively with the milk production more than progesteron alone

2-1 You searched for the answer in the literature. What are the types of articles that could answer your research question?

.....

2-2 After selecting your articles, what are the criteria you're going to consider in order to assess their validity?

.....

3- The study HERS compared women under oestrogen to women under placebo and revealed a relative risk of thrombo-embolic complications reaching 2.89 for women under oestrogen. This result suggested that oestrogens exposed the patients to a risk of thrombo-embolic complications. We wonder if there is a significant difference between them. Give an example of confidence interval sustaining the hypothesis that the proportion of thrombo-embolic complications are different between both groups

.....

4- What kind of study is useful to assess a diagnostic test?

.....

5- What type of study is used to perform a prognostic research?

.....

FIGURE 1 The final test adopted from the Fresno test.

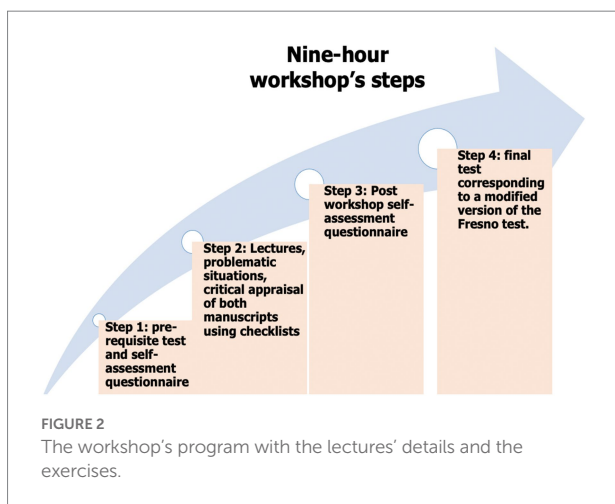


FIGURE 2 The workshop's program with the lectures' details and the exercises.

questionnaires were rated according to the Likert scale questions' answers. Non-parametric tests were used to compare the mean scores attributed to the affective, cognitive, and conative domains before and after the workshop for every student. Statistical differences in mean scores were assessed using an ANOVA test. The Pearson correlation coefficient ( $r$ )

was used to assess the strength and direction of the linear association between two paired outcomes, including associations between the students' final scores and prerequisite tests with graduation level, satisfaction, conative, cognitive, and affective profiles, and gender. The Pearson test was also used to assess the correlation between the learning, conative, cognitive, and affective profiles and the students' levels. A value of  $p < 0.05$  was considered significant. SPSS software 16.0 was used for statistical analysis.

## Research approval

This study was approved by the research committee of a University Hospital (Ref 01/2022).

## Ethics

The present study has been conducted according to the principles of the Declaration of Helsinki. Ethical approval for the study was obtained from the ethics committee of a University Hospital (Ref 07/2022). Also, participants were made aware of the

purpose of the study, the anonymous nature of the purpose, the anonymous nature of the dataset generated, and the option to not respond if they so wished. This information served as the basis for an informed consent from each respondent.

## Results

### Descriptive study

Eight workshops were organized. According to our inclusion, non-inclusion, and exclusion criteria, 23 participants were excluded and 95 participants were included (76 women and 19 men). Six participants (6.3%) were in the second year of medical education (SYME), 54 in the third year of medical education (TYME; 56.8%), 16 in continuing medical education (CME; 16.8%), and 19 participants were family doctors (FD) with a minimum of 10 years of expertise (20%).

### The different mean scores

#### The prerequisite and final tests' means

The prerequisite test mean reached 0.874 over 10. The mean scores reached 1.16 in the SYME, 0.593 in the TYME, 1.563 in CME, and 1 in FD. The prerequisite mean scores reached 0.729 in women and 1.235 in men. The final test mean score reached 4.529 over 10. The mean scores reached 5.83 in SYME, 4.287 in TYME, 3.813 in CME, and 5.41 in FD. The final test mean score reached 4.33 in women and 4.235 in men.

#### The preworkshop self-assessment questionnaires scores

##### The cognitive score

Concerning cognitive scores, scores between 0 and 9 correlated with a passive learning strategy, scores between 9 and 12 correlated with a moderate active learning strategy, and scores between 12 and 16 correlated with an active learning strategy.

The cognitive preworkshop mean score reached 10.13, revealing a general intermediate active learning strategy. The mean scores reached 10.83 in SYME, 9.72 in TYME, 11.19 in CME, and 10.16 in FD, revealing a moderate active learning strategy in all levels. The mean score reached 10.41 in women and 9.12 in men.

##### The attitude score

In the attitude domain, scores between 6 and 8 correlated with a positive affective potential to the critical appraisal practice, scores between 0 and 4 correlated with a negative affective potential, and scores between 4 and 6 were correlated with an intermediate affective potential.

The attitude preworkshop mean score reached 5.96, highlighting an intermediate affective potential to critical appraisal

principles of learning. The mean score reached 7 in SYME, 6.04 in TYME, 3.88 in CME, and 7.16 in FD, revealing a positive affective potential in SYME, TYME, and FD and a negative affective potential in CME. The mean score reached 5.92 in women and 4.76 in men, revealing an intermediate affective potential in women and men.

##### The conative score

In the conative domain, we considered scores between 9 and 12 as correlated with a high potential to use the principals of critical appraisal practice, scores between 7 and 9 as correlated with an intermediate potential to reproduce the experience, and scores between 0 and 7 as correlated with a negative potential to reproduce the experience of critical appraisal during the medical practice.

The conative preworkshop mean score reached 8.08, pointing out an intermediate potential to use critical appraisal principles by the respondents. The mean score reached 9 in SYME, 8.06 in TYME, 7.19 in CME, and 8.63 in FD, revealing a high potential in SYME and an intermediate conative potential in the other levels. The mean score reached 8.14 in women and 7.29 in men, revealing an intermediate conative potential.

##### The reflexive score

In the reflexive domain, we considered scores between 0 and 9 as correlated with a passive attitude, scores between 9 and 11 as correlated with an intermediate reflexive attitude, and scores between 11 and 16 as correlated with a high reflexive attitude. The reflexive mean score reached 10.98, reflecting a moderate reflexive attitude among the participants. The mean score reached 13 in SYME, 10.76 in TYME, 9.63 in CME, and 12.11 in FD, revealing a high reflexive potential in SYME and FD and an intermediate potential in TYME and CME. The mean score reached 10.59 in women and 11.06 in men, highlighting a moderate reflexive attitude.

#### The postworkshop self-questionnaires scores.

##### The cognitive score

We considered scores between 41 and 61 as correlated with a good perceived learning achievement, scores between 21 and 41 correlated with an intermediate perceived learning achievement, and scores lower than 21 correlated with a bad perceived achievement. The cognitive mean score reached 38.92, revealing a general intermediate self-assessment of abilities to perform a critical appraisal of medical literature. The mean score reached 55.17 in SYME, 30 in TYME, 54.56 in CME, and 45.95 in FD, revealing a highly positive self-assessment in SYME, CME, and FD and an intermediate self-assessment in TYME. The mean score reached 35.88 in women and 41.59 in men, highlighting a moderate self-assessment in women and a highly positive self-assessment in men.

### The attitude score

We considered scores ranging from 3 to 4 as correlated with good behaviorism, scores ranging from 2 to 3 as correlated with mild behaviorism, and scores <2 correlated with bad behaviorism. The affective mean score reached 2.09, revealing a mild affect concerning the practice of critical appraisal of medical literature. The mean score reached 3.83 in SYME, 1.5 in TYME, 2.81 in CME, and 2.63 in FD, revealing a good behaviorism in SYME, an intermediate behaviorism in CME and FD, and a bad behaviorism in TYME. The mean score reached 1.93 in women and 2.06 in men, revealing an intermediate behaviorism in men and a bad behaviorism in women.

### The conative score

We considered scores between 7 and 10 as correlated with good will to pursue the practice of critical appraisal of medical literature, scores between 5 and 7 as correlated with a moderate will, and scores under 5 as correlated with a weak will to pursue the practice. The conative mean score reached 6.99, highlighting a moderate potential to use the principles of critical appraisal of medical literature. The mean score reached 10.33 in SYME, 5.61 in TYME, 7.56 in CME, and 9.37 in FD, highlighting a high conative potential in SYME, CME, and FD and an intermediate conative potential in TYME. The mean score reached 6.29 in women and 6.76 in men, highlighting a moderate conative potential.

### Satisfaction mean score

We considered scores between 18 and 28 as correlated with high satisfaction, scores between 9 and 18 as correlated with

moderate satisfaction, and scores under 9 as correlated with unsatisfaction.

The satisfaction mean score reached 16.842, revealing a general moderate satisfaction toward the workshop performed. The mean score reached 26 in SYME, 12.98 in TYME, 18.37 in CME, and 23.63 in FD, highlighting a high satisfaction in SYME, CME, and FD and a moderate satisfaction in TYME. The mean score reached 14.797 in women and 16.353 in men, revealing a moderate satisfaction.

Tables 2, 3 represent the detailed results according to the level and gender.

### Variation of the mean scores according to the students' level

The prerequisite test scores varied according to the level ( $p = 0.029$ ). That was also the case for the final test scores ( $p = 0.02$ ). The pre and postworkshop attitudes, the postworkshop cognitive, the conative and satisfaction profiles varied also according to the students' level ( $p = 0.00$ ).

### Variation of the mean scores according to the students' gender

The variation of the mean scores according to the students' gender revealed a significant difference in the preworkshop attitude ( $p = 0.019$ ), postworkshop conative ( $p = 0.008$ ), and satisfaction scores ( $p = 0.00$ ) according to the gender.

TABLE 2 Detailed results of the general means and the means according to the students' levels.

	General mean			Means according to level											
	Mean	Extermes	SD	SYME			TYME			CME			FD		
				Mean	95%CI	SD	Mean	95%CI	SD	Mean	95%CI	SD	Mean	95% CI	SD
Prerequisite test	0.874	[0, 6]	1.18	1.17	[0.14, 2.2]	0.4	0.6	[0.3, 0.85]	0.13	1.6	[0.6, 2.4]	0.42	1	[0.42, 1.57]	1.2
Final test	4.529	[0, 8.5]	1.99	5.8	[4.6, 7]	0.48	4.3	[3.7, 4.8]	0.27	3.8	[2.4, 5.2]	0.65	5.41	[4.9, 5.8]	0.96
Precognitive score	10.13	[2, 17]	2.44	10.8	[10.04, 11.62]	0.3	9.7	[9.2, 10.3]	0.27	11.19	[8.9, 13.42]	1.05	10.16	[9.2, 11.03]	1.8
Preaffective score	5.96	[0, 8]	2.12	7	[5.24, 8.74]	0.7	6.04	[5.5, 6.57]	0.26	3.9	[2.56, 5.2]	0.62	7.16	[6.7, 7.5]	0.89
preconative score	8.08	[0, 12]	2.31	9	[7.7, 10.33]	0.52	8.06	[7.51, 8.61]	0.27	7.19	[5.3, 9]	0.9	8.6	[7.6, 9.5]	1.95
Reflexive score	10.98	[2, 16]	3.6	13	[10.35, 15.65]	1.033	10.76	[9.75, 11.7]	0.5	9.63	[7.18, 12.07]	1.14	13	[10.35, 15.6]	2.5
Postcognitive score	38.92	[0, 61]	20.146	55.2	[50.27, 60.06]	1.9	30	[23.9, 36.03]	3	54.56	[51.35, 57.78]	1.5	45.95	[42.4, 49.4]	7.18
Postaffective score	2.09	[0, 4]	1.55	3.8	[3.4, 4.26]	0.16	1.5	[1.09, 1.91]	0.2	2.8	[2.25, 3.37]	0.2	2.6	[1.9, 3.3]	1.46
Postconative score	6.99	[0, 12]	3.77	10.33	[2.25, 3.37]	0.33	5.61	[4.45, 6.77]	0.6	7.56	[6.57, 8.55]	0.46	9.37	[8.6, 10.11]	1.5
Satisfaction	16.84	[0, 27]	9.1	26	[25, 26.93]	0.36	12.9	[10.4, 15.5]	1.2	18.4	[14.5, 22.2]	1.8	23.63	[22.6, 24.5]	1.94

**TABLE 3** Detailed results of the means according to the students' gender.

	Women			Men		
	Mean	95% CI	SD	Mean	95%CI	SD
Prerequisite test	0.7	[0.4, 1]	0.14	1.2	[0.5, 1.9]	0.33
final test	4.3	[3.8, 4.8]	0.26	4.2	[2.9, 5.5]	0.6
precognitive score	10.4	[9.7, 11.2]	0.3	9.12	[7.5, 10.7]	0.76
preaffective score	5.9	[5.38, 6.45]	0.26	4.7	[3.4, 6.13]	0.64
preconative score	8.14	[7.5, 8.7]	0.3	7.29	[5.9, 8.7]	0.65
reflexive score	10.59	[9.5, 11.65]	0.5	11.6	[9.4, 12.7]	0.7
postcognitive score	35.8	[30.09, 41.06]	2.9	41.6	[30.8, 52.35]	5.07
postaffective score	1.9	[1.5, 2.33]	0.2	2.06	[3.4, 4.26]	0.4
postconative score	6.29	[5.22, 0.35]	0.5	6.7	[4.9, 8.55]	0.84
satisfaction score	14.8	[12.3, 17.3]	1.23	16.3	[11.5, 21.2]	2.27

**TABLE 4** variation of the different scores according to the students' level and gender.

	Variation according to the students' levels		Variation according to the students' gender	
	F	Significance	F	Significance
Prerequisite test	4.8	0.029	2.4	0.12
Final test	2.04	0.02	0.02	0.8
Precognitive score	2.3	0.1	3.3	0.07
Preaffective score	8.2	0.00	3.5	0.019
Preconative score	1.5	0.24	1.6	0.2
Reflexive score	1.7	0.18	0.19	0.6
Postcognitive score	13.17	0.00	0.8	0.34
Postaffective score	11.7	0.00	0.08	0.7
Postconative score	5.3	0.00	0.2	0.008
Satisfaction	7.4	0.00	0.35	0.00

The detailed results of the mean scores' variations according to the gender and level are represented in [Table 4](#).

## Correlations between the different scores

Pearson's test reached 0.065 ( $p = 0.531$ ) with a rho coefficient reaching 0.127 ( $p = 0.219$ ), revealing no association between the prerequisite and the final test scores. This result highlights the fact that the final test scores of the students aren't dependent of their background in biostatistics. Pearson's test reached  $-0.03$  ( $p = 0.77$ ) with a rho coefficient reaching 0.023 ( $p = 0.828$ ), revealing the absence of association between the final test scores and the

reflexive scores. Pearson's test reached 0.059 ( $p = 0.56$ ) with a rho coefficient reaching 0.107 ( $p = 0.3$ ), revealing the absence of association between the pre- and postworkshop attitude scores. Pearson's test reached 0.212 ( $p = 0.03$ ) with a rho coefficient reaching 0.17 ( $p = 0.08$ ), revealing a minor correlation between the pre- and postworkshop cognitive scores. This result highlights the fact that students that are active learners tend to be more knowledgeable after the training. Pearson's test reached 0.043 ( $p = 0.68$ ) with a rho coefficient reaching 0.039 ( $p = 0.71$ ), revealing the absence of correlation between the pre- and postworkshop conative scores.

## Qualitative analysis

Five students from the TYME, 2 students from the SYME, 2 students from CME, and 2 FD were interviewed. The content analysis highlighted 4 themes, including the teaching organization, the institutional assessment, the impact on the research practice, and the impact on the patients' management. Comments concerning the teaching organization varied according to the students' levels. Students from the SYME, CME, and FD agreed that the workshop structure and length were available and motivating. TYME students felt that the workshop was too long and expressed the need for shorter but multiple learning sessions. The institutional assessment theme was expressed differently according to the students' levels. Students from the TYME expressed the need for institutional assessment: "there is no institutional exam allowing us to assess our competencies. All of us are centered on the exams and felt that we were wasting our time because we were dealing with notions that aren't included in our curriculum." Students from the SYME expressed their wish to attend more certified courses about EBM. The impact on the research practice was mentioned differently according to the students' levels. TYME students did not understand the impact of such a teaching on the research practice: "we need all of the principles of critical appraisal when we'll be enrolled in research projects, and this will happen in the future years. At that time, we'll all forget what we have learned about critical appraisal of medical literature." The other students agreed that improving their competencies in EBM will motivate them to be enrolled or to conduct research projects in the future. Concerning the impact on the patients' management, SYME and CME students and FDs agreed that this learning will improve their patient management. The TYME students expressed that they were not aware of the impact of this learning on patient management: "we are more interested in passing the exams than in thinking about patient management." [Figure 3](#) illustrates the word cloud of the students' interviews.

## Discussion

In this study, we aimed to assess the possibility of performing the same training for students with different backgrounds. The students included were from different levels and have different backgrounds





participants (Stern, 2005). In a questionnaire-based study, the authors reported a response rate of 30% to the questionnaire, which can reflect the lack of motivation of the participants to respond (Godwin and Seguin, 2003b). Moreover, the majority of the studies reporting improvements in critical appraisal skills are based on perception or pre- and post-test scores or self-assessment rather than randomized control trials (34–36). In this study, a modified version of the Fresno test was used to assess the cognitive skills of the learners. Other assessment tools have been reported in the literature, such as the Objective Structured Clinical Exam, which has also been demonstrated as a reliable tool to assess communication skills (Ilic, 2009). Some authors reported the need of implementing activities diaries or e-portfolios that can highlight any changes related to the EBM practice (Ilic, 2009). The major limitations of this study concern the absence of randomization and the lack of assessment of the impact of this training on the process of care. In a review article dealing with the assessment of the effects of the practice of the EBM on the process of care, Horsley and co-workers reported that among 11,057 abstracts, no study evaluated the process of care or patient outcomes in relation to the practice of EBM (Horsley et al., 2011). In an intervention review, Parkes and colleagues performed a review of the literature in order to assess the effects of critical appraisal on knowledge, patient outcomes, and the process of care. According to their inclusion criteria, they included only one randomized controlled trial performed by Linzer and colleagues (Parkes et al., 2001) about 44 doctors, in which the authors reported a 25% improvement in critical appraisal knowledge in the intervention group in comparison to 6% improvement in the control group. The relative absence of evidence concerning the improvement of the attitudes and the practice thanks to critical appraisal practice cannot be considered as evidence of absence of efficiency. In fact, this can be explained by the scale of the critical appraisal programs that aren't very large and aren't followed up over many years. Also, many authors reported different effects on the participants according to their level of knowledge (Parkes et al., 2001; Horsley et al., 2011).

This study highlighted the necessity of an early introduction of the critical appraisal practice in the medical education and the necessity of improving the teaching of biostatistics principles. The fact that the different scores recorded during the TYME were inferior to scores recorded in the other levels may be explained by the major worries of the students concerning the beginning of the clerkship period. The qualitative analysis highlighted the necessity of introducing an assessment of the critical appraisal competencies in the medical education in order to make students aware of the necessity of achieving these competencies to improve their clinical practice. The major limitation of this study was the lack of the impact's assessment. Also, the different groups of students were not equal.

## Conclusion

Our results put emphasis on the necessity of introducing the teaching of critical appraisal of medical literature early in the

curriculum in order to avoid negative behaviorism during the clerkship period.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving human participants were reviewed and approved by Ethic Committee of Abderrahman Mami Hospital. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

MM had the idea, made the literature review, tutored the workshops. AN participated to the literature review and supervised the qualitative analysis, participated to the process. CD analyzed the results and reviewed the manuscript, FM, LZ, IL, and MJ contributed to the literature research, red the manuscript and reviewed it. All authors contributed to the article and approved the submitted version.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2022.1036627/full#supplementary-material>

## References

- CONSORT (2010). Checklist of information to include when reporting a randomised trial\* section/topic item no checklist item reported on page no [internet]. Available at: [www.consort-statement.org](http://www.consort-statement.org) (Accessed September 2011).
- Edgar, L., Sydney McLean, C., Sean Hogan, M. O., Hamstra, S., and Holmboe, E. S. (2003). *The Milestones Guidebook [eBook]*. Edinburgh, GB: Bailliere Tindall. Available at: <https://www.acgme.org/What-We-Do/Accreditation/Milestones/Resources> (Accessed September 2003).
- Evidence-Based Medicine, Part 2 (2007). An introduction to critical appraisal of articles on therapy [internet]. Available at: <http://www.cancer.gov/can>
- Evidence-Based Medicine, Part 3 (1995). An introduction to critical appraisal of articles on diagnosis [internet]. Available at: <http://jaao.org/> (Accessed September 1996).
- Evidence-Based Medicine, Part 5 (1995). An introduction to critical appraisal of articles on prognosis [internet]. Available at: <http://jaao.org/> Sep 2000 (Accessed September 1996).
- Evidence-Based Medicine, Part 6 (1995). An introduction to critical appraisal of clinical practice guidelines [internet]. Available at: <http://jaao.org/> (Accessed September 1996).
- Frank, J. R., Snell, L., and Sherbino, J. (2015). "Royal College of Physicians and surgeons of Canada," in *CanMEDS Physician Competency Framework. Vol. 1*. 35. eds J. R. Frank, L. Snell and J. Sherbino (Ottawa: Royal College of Physicians and Surgeons of Canada).
- Godwin, M., and Seguin, R. (2003a). *BMC Medical Education Critical Appraisal Skills of Family Physicians in Ontario [internet]*. Canada: Biomedcentral; Available at: <http://www.biomedcentral.com/1472-6920/3/10> (Accessed 2003 September).
- Godwin, M., and Seguin, R. (2003b). Critical appraisal skills of family physicians in Ontario, Canada. *BMC Med Educ.* 3:10. doi: 10.1186/1472-6920-3-10
- Goichot, B., and Meyer, N. (2011). *Lecture critique d'un article medical original (LCA)*. Strasbourg: ulpméd. Available at: <http://www-ulpméd.u-strasbg.fr> (Accessed August 01, 2012).
- Halm, M. A. (2018). Evaluating the impact of EBP education: development of a modified Fresno test for acute care nursing. *Worldviews Evid.-Based Nurs.* 15, 272–280. doi: 10.1111/wvn.12291
- Horsley, T., Hyde, C., Santesso, N., Parkes, J., Milne, R., and Stewart, R. (2011). Teaching critical appraisal skills in healthcare settings. *Cochrane Database Syst. Rev.* Art. No.: CD001270. doi: 10.1002/14651858.CD001270.pub2
- Hryciw, N., Knox, A., and Arneja, J. S. (2017). How well are we doing at teaching critical appraisal skills to our residents? A needs assessment of plastic surgery journal club. *Plastic Surgery* 25, 261–267. doi: 10.1177/2292550317731760
- Ilic, D. (2009). Assessing competency in evidence based practice: strengths and limitations of current tools in practice. *BMC Med. Educ.* 9, 213–223. doi: 10.1186/1472-6920-9-53
- Jouquan, J. (2009). La lecture critique d'article scientifique au-delà du contexte franco-français. pour éviter que l'arbre ne cache la forêt ! *Pédagogie Médicale*. 10, 77–82. doi: 10.1051/pmed/20099999
- Le Glatin, L. (2013). *Méthodologie et Lecture critique d'article médical [eBook]*. France: Faculté Humanis. Available at: [https://sofia.medicalistes.fr/spip/IMG/pdf/Methodologie\\_et\\_lecture\\_critique\\_d\\_article\\_medical.pdf](https://sofia.medicalistes.fr/spip/IMG/pdf/Methodologie_et_lecture_critique_d_article_medical.pdf) (Accessed August 26, 2013).
- Lisboa Bastos, M., Tavaziva, G., Abidi, S. K., Campbell, J. R., Haraoui, L. P., Johnston, J. C., et al. (2020). Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ*. 370:m2516. doi: 10.1136/bmj.m2516
- Ma, K. S. K., Chang, H. C., and Krupat, E. (2021). Teaching evidence-based medicine with electronic databases for preclinical education. *Adv. Physiol. Educ.* 45, 849–855. doi: 10.1152/advan.00057.2021
- Maloney, L. M., Marshall, R. T., Werfel, P. A., and Johnson, S. E. (2019). Using a journal Club series to introduce paramedic students to research fundamentals and critical appraisal of medical literature. *Prehosp. Disaster Med.* 34, 449–453. doi: 10.1017/S1049023X19004618
- Marusić, A., and Marusić, M. (2003). Teaching students how to read and write science: a mandatory course on scientific research and communication in medicine. *Acad. Med.* 78, 10–23. doi: 10.1097/00001888-200312000-00007
- Mink, J., Mitzkat, A., Mihajjevic, A. L., Trierweiler-Hauke, B., Götsch, B., Schmidt, J., et al. (2019). The impact of an interprofessional training ward on the development of interprofessional competencies: study protocol of a longitudinal mixed-methods study. *BMC Med. Educ.* 19, 751–759. doi: 10.1186/s12909-019-1478-1
- Nieminen, P., Uma, E., Pal, S., Laitala, M. L., Lappalainen, O. P., and Varghese, E. (2020). Information retrieval and awareness about evidence-based dentistry among dental undergraduate students—a comparative study between students from Malaysia and Finland. *Dent J (Basel)*. 8:103. doi: 10.3390/dj8030103
- Parkes, J., Hyde, C., Deeks, J. J., and Milne, R. (2001). "Teaching critical appraisal skills in health care settings" in *Cochrane Database of Systematic Reviews*. ed. John Wiley & Sons. 2nd ed (West Sussex: Blackwell Publishing), 123–140.
- Rohwer, A., Young, T., and van Schalkwyk, S. (2013). Effective or just practical? An evaluation of an online postgraduate module on evidence-based medicine (EBM). *BMC Med. Educ.* 13:77. doi: 10.1186/1472-6920-13-77
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., and Moran, J. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 21:j4008. doi: 10.1136/bmj.j4008
- Shehata, G. M., Zaki, A., Dowidar, N. L., and El, S. I. (2015). Critical thinking and attitude of physicians toward evidence-based medicine in Alexandria, Egypt. *J. Egypt. Public Health Assoc.* 90, 115–120. doi: 10.1097/01.EPX.0000470848.07476.f9
- Skipper, C. P., Pastick, K. A., Engen, N. W., Bangdiwala, A. S., Abassi, M., Lofgren, S. M., et al. (2020). Hydroxychloroquine in nonhospitalized adults with early covid-19: a randomized trial. *Ann. Intern. Med.* 173, 623–631. doi: 10.7326/M20-4207
- Stern, P. (2005). A holistic approach to teaching evidence-based practice. *Am. J. Occup. Ther.* 59, 157–164. doi: 10.5014/ajot.59.2.157
- St-Hilaire, S., Nekouei, O., Parkes, R. S. V., and Rosanowski, S. M. (2022). Active learning for an evidence-based veterinary medicine course during COVID-19. *Front. Veterinary Sci.* 1, 1–7. doi: 10.3389/fvets.2022.953687
- Tsang, S., Roysse, C. F., and Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J. Anaesthesia* 17, S80–S89. doi: 10.4103/sja.SJA\_203\_17
- Zins, M., Aegerter, P., Carton, M., Czernichow, S., and Ankri, J. (2011). *Lecture commentée d'article scientifique Alexis Descatha [Internet]*. France:UVSG. Available at: <http://e-campus2.uvsq.fr/cours/alexdesc/Cours.alexdesc.2011-08-31.1809/view> (Accessed October 2011).