Document generated on 04/30/2024 7:46 a.m.

Canadian Medical Education Journal
Revue canadienne de l'éducation médicale

A systematic review of the effectiveness of journal clubs in undergraduate medicine

Revue systématique de l'efficacité des clubs de lecture dans la formation médicale de premier cycle

Jibril O Bello and Paul Grant

Volume 14, Number 4, 2023

URI: https://id.erudit.org/iderudit/1106723ar
DOI: https://doi.org/10.36834/cmej.72758

See table of contents

Publisher(s)

Canadian Medical Education Journal

ISSN

1923-1202 (digital)

Explore this journal

Cite this document

Bello, J. & Grant, P. (2023). A systematic review of the effectiveness of journal clubs in undergraduate medicine. *Canadian Medical Education Journal / Revue canadienne de l'éducation médicale*, *14*(4), 35–46. https://doi.org/10.36834/cmej.72758

Article abstract

Background: Training future doctors in the skills of evidence-based medicine (EBM) is clearly important. Journal club (JCs) are well-recognized educational interventions for teaching EBM. In contrast to postgraduate medical education, JCs use in undergraduate medical education (UME) has not been adequately explored. We conducted a realist review of the effectiveness of JCs in UME to unpack the underlying mechanisms by which the intervention works (or fails) in teaching EBM.

Methods: The scope of review was the evaluation of the effectiveness of JCs in UME settings. We searched major bibliographic databases - MEDLINE, Embase, ERIC, PSYCInfo, CINAHL, Scopus, and Web of Science and found fifteen articles eligible for inclusion. Data was extracted aided by a modified Kirkpatrick framework and presented in evidence tables. Themes and chains of inference were identified, and finally, we formulated new hypotheses on how and why JC intervention works.

Results: Mandatory vs. voluntary JC did not differentially impact attendance of JC in UME settings though JC duration beyond two hours decreased attendees' self-reported satisfaction. Coupling lectures to JCs positively impacts knowledge gain and retention. Coupled Mentorship or using critical appraisal worksheets helped the achievement of manuscript writing skills and a positive attitude towards EBM.

Conclusions: Journal clubs are effective interventions to teach EBM in UME settings and are well-received by learners. They improve specific learning outcomes of knowledge gain and retention, skills of manuscript writing and critical appraisal. However, we found no evidence that these translates to the practice of EBM nor impacts patient outcomes.

© Jibril O Bello, Paul Grant, 2023

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

Canadian Medical Education Journal

Reviews, Theoretical Papers, and Meta-Analyses

A systematic review of the effectiveness of journal clubs in undergraduate medicine

Revue systématique de l'efficacité des clubs de lecture dans la formation médicale de premier cycle

Jibril O Bello,[1,2] Paul Grant[2]

[1]Department of Surgery, College of Medicine, Nursing and Health Science, Fiji National University, Fiji; [2]University of South Wales, Pontypridd, South Wales

Correspondence to: Jibril O. Bello, Department of Surgery, College of Medicine, Nursing and Health Science, Fiji National University, Fiji; email: jabarng@gmail.com

Published ahead of issue: Feb 17, 2023; published: Sept 8, 2023. CMEJ 2023, 14(4) Available at https://doi.org/10.36834/cmej.72758

© 2023 Bello, Grant; licensee Synergies Partners. This is an Open Journal Systems article distributed under the terms of the Creative Commons Attribution License. (https://creativecommons.org/licenses/by-nc-nd/4.0) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited.

Abstract

Background: Training future doctors in the skills of evidence-based medicine (EBM) is clearly important. Journal club (JCs) are well-recognized educational interventions for teaching EBM. In contrast to postgraduate medical education, JCs use in undergraduate medical education (UME) has not been adequately explored. We conducted a realist review of the effectiveness of JCs in UME to unpack the underlying mechanisms by which the intervention works (or fails) in teaching EBM.

Methods: The scope of review was the evaluation of the effectiveness of JCs in UME settings. We searched major bibliographic databases - MEDLINE, Embase, ERIC, PSYCInfo, CINAHL, Scopus, and Web of Science and found fifteen articles eligible for inclusion. Data was extracted aided by a modified Kirkpatrick framework and presented in evidence tables. Themes and chains of inference were identified, and finally, we formulated new hypotheses on how and why JC intervention works.

Results: Mandatory vs. voluntary JC did not differentially impact attendance of JC in UME settings though JC duration beyond two hours decreased attendees' self-reported satisfaction. Coupling lectures to JCs positively impacts knowledge gain and retention. Coupled Mentorship or using critical appraisal worksheets helped the achievement of manuscript writing skills and a positive attitude towards EBM.

Conclusions: Journal clubs are effective interventions to teach EBM in UME settings and are well-received by learners. They improve specific learning outcomes of knowledge gain and retention, skills of manuscript writing and critical appraisal. However, we found no evidence that these translates to the practice of EBM nor impacts patient outcomes.

Résumé

Contexte : Il est important que les compétences enseignées aux futurs médecins soient conformes à l'approche de la médecine fondée sur les données probantes (MFDP) et les clubs de lecture (CL) sont reconnus comme une intervention pédagogique en ce sens. Le recours aux CL dans la formation prédoctorale a été peu étudié, contrairement à la place qu'ils occupent au postdoctorat. Nous avons effectué une revue réaliste de l'efficacité des CL dans la formation de premier cycle pour décortiquer les mécanismes sous-jacents qui déterminent la réussite ou l'échec de cette méthode pédagogique en ce qui a trait à l'enseignement de la MFDP.

Méthodes : La revue visait l'évaluation de l'efficacité des CL dans la formation médicale prédoctroale. Nous avons effectué des recherches dans les principales banques de données bibliographiques – MEDLINE, Embase, ERIC, PSYCInfo, CINAHL, Scopus et Web of Science – et avons trouvé quinze articles répondant à nos critères d'inclusion. Nous avons utilisé le modèle modifié de Kirkpatrick pour la collecte de données, que nous avons présentées sous forme de tableaux de preuves. Nous avons dégagé des thèmes et élaboré des chaînes d'inférences, puis formulé de nouvelles hypothèses sur le comment et le pourquoi de l'efficacité des CL.

Résultats : Le caractère obligatoire ou facultatif de l'activité n'a pas eu d'effet différentiel sur la participation des étudiants de premier cycle aux CL, mais leur satisfaction autodéclarée diminuait lorsque la durée de la séance dépassait deux heures. En revanche, le fait de conjuguer les CL à des conférences a eu un effet positif sur l'acquisition et la rétention des connaissances et le fait de les conjuguer au mentorat ou à l'utilisation de feuilles d'évaluation critique a contribué à l'acquisition de compétences en rédaction et à l'adoption par les étudiants d'une attitude positive à l'égard de la MFDP.

Conclusions : Les clubs de lecture constituent un moyen efficace et bien accueilli par les apprenants pour enseigner la MFDP au prédoctorat. Ils améliorent les résultats d'apprentissage spécifiques sur le plan de l'acquisition et de la rétention des connaissances, ainsi que les compétences en matière de rédaction et d'évaluation critique. Toutefois, nous n'avons trouvé aucune indication que ces effets puissent se répercuter dans la pratique de la MFDP ou dans les résultats pour les patients

Introduction

Modern society expects up to date and evidence-based care from medical practitioners. One popular mechanism to facilitate this is the Journal Club (JC)—a structured gathering of medical trainees and professionals who review, discuss, and debate the contemporary medical literature. Journal clubs have been around since the second half of the 19th century and promote critical thinking amongst doctors, an essential skill for both academia and clinical practice and a major contributor to the development and implementation of evidence-based medicine (EBM).[1] The practice of EBM involves the practitioners combining their expert medical knowledge, analytic and appraisal skills to dissect relevant evidence from the scientific literature to support patient care because it enhances clinical decision making.[2] Knowledge gaps continue to exist on how best to teach and instill EBM. Journal clubs are believed to be an excellent instructional method to increase knowledge acquisition and improve the skills of critical appraisal of research. The JC process has many of the benefits identified above,[3] and are well-suited to the needs of adult learners because they are learner-centered, interactive with a community of practice; they encourage critical appraisal of new knowledge with critical reflection on long-held beliefs and spur the deliberate practice of EBM.

JCs are commonplace in postgraduate medical education worldwide, however, they are underutilised in undergraduate medical training globally.[4] Evidence of this can readily be deduced from the underrepresentation of undergraduate medical education in the systematic reviews of JCs that have been conducted which were mostly in postgraduate or continuing medical education settings. In the review by Harris *et al*.,[3] only one paper out of the 18 included was specifically on JCs in UME. Two other reviews completed by Deenadayalan et al., and Honey and Baker did not expressly state what proportion of included studies were of undergraduate studies, but it was clear that most if not all were studies of residents and practicing physicians.[5,6] A few other systematic reviews out rightly excluded studies on UME perhaps due to the observed underuse of JC instructional methods in medical schools;[7,8] a situation which may be due to constraints on curriculum time and available educational resource, and or the potential practical difficulty in applying this mode of instruction to very large groups of students. Edwards and colleagues demonstrated that JC instructional method is well received by medical students and can be effectively used to teach EBM.[9] The development of EBM related knowledge and skills could be potential benefits of the wider roll out of JCs to medical trainees at an early stage.

JCs appears to have remained largely in the format purportedly created by William Osler circa 1875, which can be described as the presentation and critical appraisal of a selected article followed by an enthusiastic discussion by participants.[1] It is not known if this traditional process is suitable for the UME learning environment as relatively little research has been undertaken in this area.[9] Arguably it is a good way for trainees to keep current with the contemporary literature and learn the important skills of critical analysis and debate as this will ultimately make them better practitioners for the future, rather than meekly accept the mass of medical information presented to them to ingest and regurgitate.

Some novel JCs, ingeniously modified from the traditional format, have been described and are in use in UME.[10,11] One example, the 'Jigsaw' JC described by Willett et al. was effectively used to instruct a large class of 125, five papers were discussed with the help of five preceptors, and the total duration of the JC was 2.5 hours. These innovative JCs are aligned to the identified learning needs of medical undergraduates, and this may help the novice medical student easily grasp new concepts and progressively develop and master the skill of critical appraisal and other desirable learning outcomes.

The effectiveness of JC educational intervention can be evaluated using the framework of the Kirkpatrick model;[12,13] first described by Donald Kirkpatrick in the 1950s, it provides a four-level classification of relevant and measurable outcomes intended for the evaluation of training in business organisations.[12] It is simple, easy to use and has well-defined terms in each level, and has been successfully adapted for the evaluation of the effectiveness of learning interventions.[13] For this review, it was important to focus on learning outcomes and effectiveness of our intervention, so we used the modified framework in our analysis. The model does have limitations which include its blindness to the motivation for learning, a program's contextual interaction with some resource utilisation and learners' variable entry levels with regards to knowledge and skills.[14] But for the purposes of our evaluation, it was believed that these does not have a major impact in our discrete area of interest.

JC instruction has been extensively and successfully used to teach critical appraisal and EBM skills in postgraduate

medical education, at the present time, there is limited understanding of whether this form of intervention works well or not in UME setting and if it does, whether such JCs would share similar characteristics and mechanisms of action as postgraduate JCs. Only a few systematic reviews of JCs in medical education exist and all were over a decade old and predominantly studies of postgraduate JCs.[3,5,6,7] Some of the prior reviews included studies from both postgraduate and undergraduate settings, and this introduces difficulty in interpreting outcomes as attendees were at very different Dreyfus levels of competency. Finally, an acknowledged limitation of systematic reviews is that their value diminishes over time, and in the last decade, there has been increased primary research in JCs in the undergraduate setting. A new systematic review specifically focused on JCs in UME is needed to incorporate evidence from these new studies.

We used realist review methodology to synthesise evidence about the effectiveness of JC educational intervention in UME settings. Realist synthesis is particularly well-suited to evaluating complex interventions such as a JC.[15] We interrogated the available evidence in selected articles aided by the modified Kirkpatrick framework. Therefore, our aim was to describe the evidence of effectiveness of JCs in UME and the underlying mechanisms and contexts in which it works or not. We also hope to provide supportive evidence for its introduction into undergraduate medical curricula and to inform future research in this form of instruction.

Methods

We conducted a realist review which focused on unpacking the mechanisms by which JC educational intervention works (or fails to work) in UME settings. This review was guided by three steps adapted from the description of Realist methodology by Parson et al.[15] In the first step, we defined the review scope and articulated the initial programme theory. The scope of this review encompasses how and why a JC works (or not) in teaching EBM in UME settings. We developed an initial programme theory–that JC educational intervention can engender critical thinking and appraisal skills and would translate to EBM skills in medical students - based on empirical evidence on the well-established postgraduate JCs and reviewed it in a comparative setting (UME).

In the second step, we conducted an electronic search of relevant literature in the following databases: MEDLINE, Embase, ERIC, PSYCInfo, CINAHL, Scopus, Web of Science and Google Scholar for articles published between 1st January 2010 and 31st July 2020. MeSH terms were: "Education, Medical, Undergraduate" AND "Periodicals as Topic" OR "Journal club" and free text: evidence-based medicine, undergraduate medical education and journal club. The 10-year timeframe was considered appropriate due to the dearth of reviews of JCs in medical education during that period. The following inclusion criteria were used: 1) Study population included undergraduate medical students 2) The intervention was Journal club instruction in undergraduate medical education setting 3) Reported outcomes of students' reactions, change in attitudes, knowledge and or skills, behavioural change, and impact on patients care. Only articles published in English language were included from this review. Articles on JCs in postgraduate medical training and for CME of physicians were excluded except in situations where such articles also report on JCs in UME.

Articles retrieved from the search were initially screened by title and abstract, papers deemed potentially relevant based on this preliminary screen were then fully reviewed by one reviewer (JB) for consideration. Where it was not clear if an article deserved inclusion or not, it was resolved through discussion with the second reviewer (PG).

In the third and final step, we extracted relevant data and conducted a synthesis of the data. We used a modified framework of the Kirkpatrick model to interrogate selected papers and guide the data extraction.[12,13] The modified framework described specific outcomes and assessments in each of four levels and was well-suited to evaluating the effectiveness of JC educational interventions. These learning outcomes include: 1) Reactions (participation in JC educational intervention, engagement at intervention, satisfaction with the intervention), 2) Learning (knowledge: knowledge of EBM - knowledge gain on topics discussed, statistics; Skill: skill of literature search, critical appraisal, manuscript writing and achievement of publication; Attitudes: attitudes towards JC or EBM), 3) Behavior (adoption of evidence-based practice), and 4) Results (improved patient outcomes).

We also used the Medical Education Research Study Quality Instrument (MERSQI) to evaluate the quality and methodological rigor of the included studies. The instrument scores ten items in six domains: study design, sampling, type of data, validity evidence, data analysis and outcomes.[16] This instrument, with a maximum score of 3 in each domain and total possible score of 18 (highest quality), is particularly well-suited for this review. Its

strengths are the substantial body of validity evidence including excellent interrater reliability, its broad and objective items, and its adoption of Kirkpatrick's four-level model construct in the 'outcomes' domain (item 10).[16]

Data were organised into evidence tables, themes identified, and chains of inference linked. The process of synthesis helped the refinement of the initial programme theory and the formulation of hypotheses. Significant variability of quantitative data types and formats in the included studies precluded a meta-analysis. The review protocol was registered and published in the International Prospective Register of Systematic Reviews (PROSPERO) with the ID PROSPERO 2020 CRD42020216180. The Faculty of Life Science and Education, University of South Wales granted ethical approval for the study.

Results

From the search of databases conducted (Figure 1), 457 papers were identified; on subsequent screening of the abstracts or full text of these articles, 15 met the minimum inclusion criteria described and were included in this review.[10,11,17-29]

Of the 15 studies included, 10 were conducted in the United States of America, two in the United Kingdom, and one each in Croatia, Ireland and the United Arab Emirates.

The size of JCs varied greatly and ranged from three to 250 members. Eight studies described ≤ 30 attendees at the JCs, four studies had > 30 but ≤ 55 attendees, and two studies gave no information on the number of attendees at the JCs. One study which described a novel interactive JC had 250 attendees. The MERSQI mean score for 14 studies (one study with qualitative design could not be assessed) was 8.32 (SD 2.93; range: 5.5–15.0). Most of the included papers had non-experimental design with the majority having single group cross-sectional or single group post-test only study design (10, 71.4%) and only one study (7.1%) was a RCT. Mean domain scores were lowest for validity (0.93) and highest for data analysis (1.68) (Table 1).

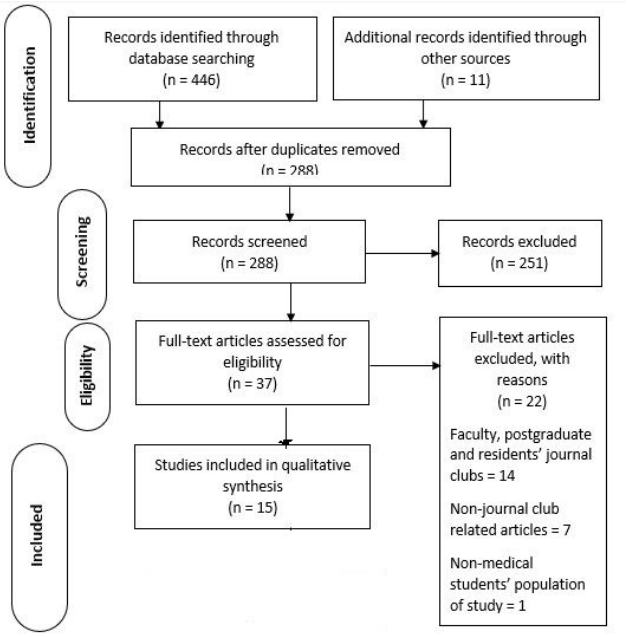

Figure 1. Flowchart of database search and the article selection process

Themes

The study designs, characteristics of participants studied, descriptions of JC instructional method and Kirkpatrick's outcome level in the 15 included studies are summarized in Table 2. Table 3 also summarises characteristics of undergraduate JCs—the type of JC, the underpinning learning theories, regularity and length of sessions and the coupled activity to JC.

Of the 15 studies included, 10 cited the learning theories underpinning their JCs. Four studies reported adult learning theory (Willett; Berman; Marusic; Rosenthal),[11,17,18,19] two cited constructivism (Banerjee; Williams),[10,20] and each of the remaining four studies cited experiential learning theory (Bahner),[21] team-based learning (Chakraborti),[22] peer-based learning (Gokani),[23] and social learning theory/communities of practice (Quinn).[24]

Table 1. Summary of Medical Education Research Study Quality Instrument (MERSQI) tool by domain and item scores for 14 included studies of undergraduate journal clubs

Domain	MERSQI Item	Item Score	Max. Domain Score	Studies, n (%)	Mean score (SD) Item	Domain
Study Design			3		1.36(0.72)	1.36(0.72)
	Study Design					
	a. Single group cross-sectional or single group post-test only	1		10 (71.4)		
	b. Single group pre and post-test	1.5		3 (21.4)		
	c. Non-randomized, 2 group	2		0 (0)		
	d. Randomized controlled experiment	3		1 (7.1)		
Sampling			3			0.83(0.46)
	Institutions				0.61(0.29)	
	a. Single institution	0.5		12 (85.7)		
	b. Two institutions	1		1 (7.1)		
	c. More than 2 institutions	1.5		1 (7.1)		
	Response Rate				1.13(0.39)	
	a. Not applicable	n/a		7 (50.0)		
	b. Response rate <50% or not reported	0.5		1 (7.1)		
	c. Response rate 50-74%	1		1 (7.1)		
	d. Response rate ≥75%	1.5		5 (35.7)		
Type of Data			3		1.29(0.73)	1.29(0.73)
	Type of Data					
	a. Assessment by study subject	1		12 (85.7)		
	b. Objective measurement	3		2 (14.3)		
Validity of Evaluation Instruments' Scores			3			0.93(1.07)
	Not applicable	n/a				
Internal Structure						
a.	Not reported	0		13 (92.9)		
b.	Reported	1		1 (7.1)		
Content						
a.	Not reported	0		8 (57.1)		
b.	Reported	1		6 (42.9)		
Relationships to other variables						
a.	Not reported	0		9 (64.3)		
b.	Reported	1		5 (35.7)		
Data Analysis			3			1.68(0.77)
	Appropriateness of analysis				1.00 (0)	
	a. Data analysis inappropriate for study design or type of data	0		0 (0)		
	b. Data analysis appropriate for study design and type of data	1		14 (100)		
	Sophistication of analysis				2.36(0.50)	
	a. Descriptive analysis only	1		9 (64.3)		
	b. Beyond descriptive analysis	2		5 (35.7)		
Outcome			3			1.32(0.25)
	Outcome					
	a. Satisfaction, attitudes, perceptions, opinions, general facts	1		5 (35.7)		
	b. Knowledge, skills	1.5		9 (64.3)		
	c. Behaviours	2		0 (0)		
	d. Patient/health care outcome	3		0 (0)		
Total			18			8.32(2.93)

Table 2. Study designs, participants/learners studied, descriptions of journal club instructional methods and outcomes in the 15 included studies of undergraduate medical education journal clubs

Source	Study design	Participants	Description of JCs	Outcomes (Kirkpatrick levels)
Bahner and Royall,[21] 2013	Descriptive-evaluative	14->40; 4th year, Clinical	Part of a program curriculum with multimodal instructional method including JC. An online quiz based on selected paper prior to JC prepares students for JC session. JC session is the traditional format where student presents article after which it is discussed by participants facilitated by a leader (faculty).	1, 2
Banerjee et al.,[10] 2018	Descriptive-evaluative	54; 1st year, preclinical	An initial short didactic lecture prior to JC sessions prepares the students. Article is shared over LMS. Students in groups of 3-4 present article following a novel 6D approach (didactic lecture, designate presenters, distribute paper, design presentation, deliver presentation and discussion). First 4 steps are facilitated by a mentor/faculty and the other 2 are student driven.	1, 2
Berman et al.,[17] 2019	One-group posttest-only	15-30; 1st, 2nd &3rd year, preclinical & clinical	Traditional JC format	1, 2
Bertelsen et al.,[26] 2015	Descriptive-evaluative	9-12; mostly 3rd year, clinical	Initial didactic lecture prior to a traditional JC session	1
Chakraborti,[22] 2011	Descriptive-evaluative	NA; 2nd year, preclinical	An initial online quiz, use of online CAW by students to work through article appraisal in class. Faculty uses the concluding part of JC session to reinforce concepts planned in syllabus e.g Odds ratio, risk, NNT, PICO.	1, 2
Currier et al.,[28] 2013	One-group pretest-posttest	17; preclinical	Traditional JC format	1, 2
Curtis et al.,[27] 2016	Descriptive-evaluative	>10; across year groups, preclinical & clinical	Traditional JC format	1, 2
Gokani et al.,[23] 2019	One-group pretest-posttest	3-6 (98); across year groups, preclinical and clinical	A novel National JC program. Small groups of attendees use CAW to work through critical appraisal of selected paper. Discussion then follows facilitated by tutor(s)	1, 2
Hultman et al.,[29] 2012	One-group pretest-posttest	26; 4th year, clinical	Part of a program curriculum with multimodal instructional method including JC. Minimal description JC provided, traditional format	1, 2
Lucia and Swanberg,[25] 2018	Descriptive-evaluative	4-6; 1st year, preclinical	Prior graded written assignment prior to a traditional JC session to encourage student to read the selected article. Post JC feedback from faculty and student is provided to presenters.	1
Marusic et al.,[18] 2014	Descriptive-evaluative	NA; 3rd year clinical	Traditional JC with each session focused on one type of study design. JC sessions are graded	1
Quinn et al.,[24] 2014	Qualitative study (thematic content analysis)	41; clinical	Traditional JC format	1
Rosenthal and Rosenthal,[19] 2017	Descriptive-evaluative	70-250; clinical	A novel interactive JC with a defined structured approach intended to make session a more active learning experience. Slides are presented in a defined sequential format to aid critical appraisal of the selected paper. Only the presenter (designated leader) reads the article in advance. Slides are composed of the following: article title slide; slides of article figures and tables / data; slides of key appraisal process steps/questions that guides JC session; summary slide; revisit title slide, final analysis & discussion.	1
Willett et al.,[11] 2013	Descriptive-evaluative	25 (125); 2nd year preclinical	A novel JC with 2-stage process and based on cooperative learning technique. The authors described the use of 5 classrooms each with 25 students (in groups of five) and a preceptor. Each class is assigned one of five articles (all five covering different aspects of one condition). Stage1: each small group in each class use CAW to work through the designated article and proceed to have a class discussion facilitated by the preceptor. Stage 2: students are moved to different classrooms to make up new groups of five (each member a expert of one of the five articles). Stage 2 completes with students with cooperative appraisal of all 5 articles.	1
Williams and Mann,[20] 2017	Randomized controlled trial	50; 3rd year, clinical	Initial didactic lecture prior to traditional JC session.	1, 2

Description of medical undergraduate JCs

Table 3 summarises the characteristic of JCs in UME in the included articles, such as whether they were compulsory, how frequent they were, whether they had an evaluation process built in and what the underlying learning theory was. Thirteen studies reported on how articles were selected for appraisal and discussion. Faculty or tutors selected articles for JCs in eight studies (Bahner; Chakraborti; Gokani; Quinn; Rosenthal; Willett; Williams; Lucia),[11,19,20,21,22,23,24,25] students selected the articles in three studies (Banerjee; Bertelsen; Curtis),[10,26,27] and in the remaining two studies (Berman, Currier),[17,28] the articles were selected via a collaboration between the students and the faculty or tutor. Sessions in many of the included studies were similar to a traditional journal club format. However, some elements like didactic lectures, online quizzes, graded written assignment, critical appraisal worksheet (CAW) were sometimes coupled to the JC and used to ensure the students engaged with and read the selected article(s) and were well-prepared for the journal club sessions, (Tables 2 & 3).

Table 3. Characteristics of undergraduate medical education journal clubs in the 15 included articles

Source	Type; voluntary Vs Mandatory JC	Topic/focus; Underpinning learning theory	Regularity; length of session	Coupled activity to JC session
Bahner and Royall,[21] 2013	Student-led JC; mandatory	Ultrasonography; experiential LT	Monthly; none provided	Quiz
Banerjee et al.,[10] 2018	Mentored JC; mandatory	Basic science; Constructivism	Fortnightly; 40 minutes	Didactic lecture
Berman et al.,[17] 2019	Student-led JC; voluntary	Surgery; adult LT	Monthly; 60 minutes	None
Bertelsen et al.,[26] 2015	Critical appraisal JC; mandatory	Global health; none provided	Weekly, 90 minutes	Didactic lecture
Chakraborti,[22] 2011	Critical appraisal JC; mandatory	Evidence-based medicine; team-based learning	None provided; 60 minutes	Quiz, online CAW
Currier et al.,[28] 2013	Mentored JC; voluntary	Evidence-based medicine; none provided	Weekly, 60 minutes	None
Curtis et al.,[27] 2016	Student-led JC; voluntary	Any topic relevant to curriculum; none provided	Monthly; 60-120 minutes	Writing letter to editor
Gokani et al.,[23] 2019	Critical appraisal-based JC; voluntary	Evidence-based medicine; peer-based learning	Annually, 90 minutes	Didactic lecture, CAW, Writing letter to editor
Hultman et al.,[29] 2012	No description of type; voluntary	Professionalism; none provided	None provided; 60 minutes	None
Lucia and Swanberg,[25] 2018	Mentored JC; mandatory	Evidence-based medicine; none provided	bi-annually; 45 minutes	Graded written assignment, feedbacks from faculty and students
Marusic et al.,[18] 2014	Mentored JC; mandatory	Science/research education; adult LT	Weekly; 180 minutes	Grading of JC session
Quinn et al.,[24] 2014	Traditional JC; unknown	Surgery; Social LT	Weekly; none provided	None
Rosenthal and Rosenthal,[19] 2017	Critical appraisal-based JC; unknown	Evidence-based medicine; adult LT	None provided; 50 minutes	None (novel JC described)
Willett et al.,[11] 2013	Critical appraisal-based JC; unknown	Evidence-based medicine; active learning principles	Bi-annually; 150 minutes	None (novel JC described)
Williams and Mann,[20] 2017	Traditional JC; mandatory	Surgery; constructivism	Monthly; 60 minutes	Didactic lecture

Didactic lectures coupled to JCs in UME were used in many interesting ways. For instance, Banerjee and colleagues described lectures on key concepts in their basic science syllabus which were aligned to the subject of the paper(s) selected for the JCs. Wiliams and Mann similarly used a 30-minute lecture to review basic science concepts related to the clinical content of the selected paper.[20] Bertelsen and colleagues described the use of the didactic exercise to inform JC participants of the clinical background of the global health topic of the selected journal article while Gokani and colleagues used an initial 20 minutes for a lecture on critical appraisal.[23,26]

In two studies (Chakraborti, Willett),[11,22] the use of critical appraisal worksheets was described, and they were used to guide the article appraisal process. Two other studies (Curtis, Gokani) described mentorship of students in the skill of writing letters to editor based on the JC discussions and guiding the subsequent submission for publication.[23,27]

Novel and innovative JCs were described in three studies.[10,11,19] Rosenthal and Rosenthal described what the authors termed 'an interactive JC' in which a designated leader uses a structured appraisal process with slides and key questions to guide the article appraisal process.[19] The participants are engaged and active throughout the process and have opportunities to develop critical appraisal skills, communication skills, and confidence. Willett et al. described a two-stage 'jigsaw' cooperative learning approach which required five classrooms each with a preceptor and 25 students divided into groups of five.[11] Each student involvement and input are important in completing the appraisal of selected papers. In the first of the two stages, students in a group / classroom appraise a paper to become an expert; in the second stage, each expert student in one selected paper engages other expert students in other selected papers and they help each other to understand the appraisal of all the selected papers. A third innovative JC was described by Banerjee et al.[10] They described a novel and self-explanatory 6D approach (**D**idactic lecture, **D**esignate presenters, **D**istribute paper, **D**esign presentation, **D**eliver presentation and **D**iscussion) used to teach basic science concepts in preclinical medical education.[10]

Effectiveness of JC educational strategies

Nine studies (60%) reported Kirkpatrick's level 2 outcomes and the remaining studies had level 1 outcomes, no study demonstrated level 3 or 4 impact (Table 4). Most of the reported outcome variables were self-reported measures (16/24 overall outcomes in this review), observation was recorded in 7/24 and only one knowledge test conducted and reported.

Table 4. Realist synthesis: Journal club characteristics by outcomes based on Kirkpatrick evaluation framework

Source	Characteristics	Attendance, Engagement	Satisfaction	Knowledge	Skills, Attitudes, Confidence
Bahner and Royall,[21] 2013	Man, Q, P^f, M^m	+(O)		+(O)	
Banerjee et al.,[10] 2018	Man, D, P^s, M, N, T^1		+(SR)	+(O)	
Berman et al.,[17] 2019	Vol, T^1		+(SR)	+(SR)	
Bertelsen et al.,26 2015	Man, D, M^m, T^{1-2}		+(SR)		
Chakraborti,[22] 2011	Man, Q, CAW, T^1				+(SR)
Currier et al.,[28] 2013	Vol, T^1	+(O)	+(SR)	+(SR)	
Curtis et al.,[27] 2016	Vol, M, P^s, T^{1-2}	+(O)			+(O)
Gokani et al.,[23] 2019	Vol, D, CAW, M, T^{1-2}		+(SR)	+(SR)	+(SR)
Hultman et al.,[29] 2012	Vol, M^m, T^1		-(SR)	+(SR)	
Lucia and Swanberg,[25] 2018	Man, WA, T^1		-(SR)		
Marusic et al.,[18] 2014	Man, G, $T^{>2}$		-(SR)		
Quinn et al.,[24] 2014		-(O)			
Rosenthal and Rosenthal,[19] 2017	N, T^1		+(SR)		
Willett et al.,[11] 2013	N, CAW, $T^{>2}$		-(SR)		
Williams and Mann,[20] 2017	Man, D, T^1			+(Te)	

CAW: Critical appraisal tool; D: didactic lecture; G: grading JC; M: mentoring; Man: mandatory JC; M^m: multimodal instruction; N: novel journal club; P^f: paper assigned by faculty; P^s: paper assigned by student; T^1: duration of journal club ≤ 1 hour; T^{1-2}: duration of journal club > 1 hour < 2 hours; $T^{>2}$: duration of journal club > 2 hours; Vol: Voluntary JC; WA: written assessment; Q: quiz. Variable measurement: O, Observed; SR, Self-report; Te, Test.

Level 1 outcomes (Reactions)

Participation. Only four studies (Bahner, Quinn, Curtis, Currier) reported on students' participation or engagement at the JC and they all reported good or improved attendance. [21,24,27,28] One study (Bahner) reported an increase in attendance from 14 to 40 over five years,[21] a second study (Currier) described an increase in attendance at JC of 81% from the previous year (from 9 to 26 students out of a total of 32 students),[28] and the remaining two studies (Curtis, Quinn) similarly reported good attendance and enthusiasm.[24,27] Realist synthesis revealed that, of the four studies reporting on attendance, two were voluntary JCs, one was a mandatory JC, and one (Quinn) did not have sufficient information to determine this status (Table 4). All three with sufficient information reported improved attendance with no difference between mandatory vs voluntary JC.

Satisfaction. Ten studies reported on this outcome with six reporting students' satisfaction with the educational intervention (Table 4). In one study (Currier),[28] students were so enthused with the JC intervention that the sessions typically exceeded the scheduled one-hour duration, the Gokani study reported that students wanted more time to be added to the JC sessions and in the Rosenthal study, students considered the JC as fun.[19,23] Two studies (Banerjee, Bertelsen) reported that 89% and 86% of students were extremely satisfied with the JC respectively,[10,26] and participants in another study (Berman) reported a satisfying experience with rating of 4 out of 5 on Likert scale.[17] Some studies however reported dissatisfaction with the intervention. In one study (Hultman),[29] students ranked the intervention a rating of 2.25 out of 5, the Lucia and Swanberg study found that their students disliked the coupled written activity to the JC educational intervention,[25] and in two studies (Marusic, Willett),[11,18] students were dissatisfied with the long duration of the JC. Realist synthesis revealed several interesting information. The duration of JC was related to participants' self-reported satisfaction with the educational intervention, students reported dissatisfaction with two JC interventions with duration of over 2 hours (Marusic and Willett).[11,18] Six JCs with one-hour duration or less gave information on self-reported satisfaction,[10,17,19,25,28,29] and 4/6 (Rosenthal, Currier, Berman and Banerjee) gave positive satisfaction reports while 2/6 (Lucia, Hultman) reported participants were dissatisfied. Two JCs had durations between one and two hours and both reported participants were satisfied with the intervention. The coupling of didactic lectures with JC appeared well received by students. Three studies (Banerjee, Bertelsen, Gokani) with coupled didactic instruction reported that students were satisfied with the JC.[10,23,26] Students expressed satisfaction with half (2/4) of the mandatory JCs and were dissatisfied with the other half. Of four voluntary JCs with information on self-reported satisfaction, participants in three studies (3/4) expressed satisfaction with the intervention.

Level 2 (Knowledge, skills, and attitude)

Knowledge. Seven studies assessed knowledge and four of these were self-reported change in knowledge (5-point Likert scale), two were observations of test scores and one study used an unvalidated knowledge test (Table 4). Two studies (Bahner, Banerjee) reported a pass rate of > 85% and 96% following JC instructional intervention.[10,21] A randomized controlled trial (Williams and Mann) coupled a didactic lecture on the basic science background of a selected article to the JC discussions (intervention arm) and compared it to JC discussion sessions only (control arm), found that knowledge was gained with regards to basic science aspect of the paper and this knowledge was retained at 3 months, no similar retain of knowledge was found with clinical knowledge based on the discussions at the JC sessions.[20] A 2019 study (Berman) reported that participants at a JC mostly agree or strongly agree that the educational intervention helped them to learn to detect research bias and improve their knowledge of biostatistics with a mean of 4.5 and 4.3 out of 5 respectively.[17] In another study (Currier),[28] students reported improvement in their ability to distinguishing translational research from basic and clinical research, understanding the tenets of study design and distinguishing between statistical, biological and clinical significance. The Gokani study reported an increase in students' knowledge on how to assess methodology of articles (median 3 Vs 4, $p < 0.01$); they also reported an increased confidence with critical appraisal of articles (median 2 Vs 4, $p<0.01$) and an increased confidence in writing letters to journal editors (median 2 Vs 4, $p<0.01$).[23] The Hultman study also reported a significant improvement in knowledge of defining professionalism (3.46 Vs 4.29), understanding the attributes of professionalism (3.75 Vs 4.5) and understanding the importance of professionalism (4.17 Vs 4.67).[29] Realist synthesis revealed that all the seven studies assessing knowledge reported positive impact or changes. Three out of the seven studies (3/7) used didactic support to the JC, one of these (Williams and Mann) reported an objective change in test score using a rigorous RCT study design (Table 4).

Skills. One study (Chakraborti) reported a 17.5% (95%CI 9.3 – 25.3) increase in students' skill of literature search.[22] Another study (Curtis) demonstrated that students participating in a JC acquired the skill of writing (letters to editors) and successfully published one such letter.[27] A third study (Gokani) included mentorship of students on writing letters to journal editors as a post JC activity.[23] Realist synthesis revealed that of the three studies reporting positive impact of JC on skills (literature search skills and writing letters to editors), two studies (2/3) used coupled mentoring and structured CAW (Table 4).

Attitudes. In the Chakraborti study,[22] 70% of the students reported that they agreed or strongly agreed (using a Likert scale of 1-5 where 5 is strongly agree) that EBM skill gained from JC will be helpful in their future practice (mean score

3.86, 95%CI 3.65 – 4.06). The Gokani study also reported that students agreed that the JC was useful to their careers (mean score 5, IQR 4-5).[23] Both studies reporting change in attitudes employed the use of structured CAW.

Discussion

Although JCs have a long history and traditional role in the ongoing medical education of practicing doctors, there is limited work and evidence on their role and effectiveness in undergraduates. We have critically analysed the limited literature on this subject with most of the available published papers mainly from the USA and the UK; this distribution follows similar patterns from earlier reviews.[3,7]

Overall, there is conflicting evidence on the value of JCs in medical education. A previous review found that it significantly improved knowledge and critical appraisal skills amongst all journal club attendees;[5] however, in another review, no clear evidence to support the effectiveness of JCs was found.[3] This may be the consequence of methodological problems as there were inconsistent definitions of the effectiveness outcome measures in included studies such that it was difficult to make valid comparisons between them.[3] Though we also found varied and non-standardized outcomes measured with non-validated tools in the current review, the realist synthesis we conducted and our use of Kirkpatrick's framework to guide data extraction and evaluate the effectiveness of the intervention, proved very useful in unpacking the underlying mechanisms by which the JCs worked and under what contexts they do (or fails to) work. For example, this review found that medical students expressed great enthusiasm for the JC instructional method, and this was seen as increased attendance in three out of four studies that reported this outcome measure. This was mapped to Level 1 of the modified Kirkpatrick's framework. Additionally, in six out of 10 studies, the participating medical students expressed satisfaction with JCs. When analysed across identified themes like mandatory and voluntary JCs, we found no difference in attendance of mandatory JCs compared to voluntary JCs. This may be due to the research focus and the interactive nature of JCs, learners were interested enough to attend sessions and satisfaction with the mode of instruction probably encouraged continued attendance. Realist synthesis revealed that the duration of a JC does appear to impact self-reported satisfaction of attendees with JC duration more than two hours being viewed particularly negatively. This may be due to the impact of longer durations on attention spans and the participants get tired and tune out. Students may also resent the time taken away from other learning activities.[18] This current review is the first to evaluate the relationships between JC duration and participants satisfaction, previous reviews have been silent on this.

The coupling of didactic lectures to JCs appear to be well received by students with students reporting satisfaction in all studies that included this additional support. We also found that lectures added to JC's impacted on knowledge gain and retention, a Level 2 feature on the modified Kirkpatrick framework, and this finding is supported by conclusions of a previous review which reported that didactic support improved reading habits, knowledge and skills.[3] Ebbert and colleagues also found that JC's could improve knowledge of reading habits, clinical epidemiology and biostatistics.[7] In formulating chains of inference from identified themes, we synthesised that all the seven studies assessing knowledge reported positive changes. Three of these seven used didactic support and one of the three (Williams and Mann) used a rigorous RCT study design and reported an objective change in test scores. We believe didactic support is helpful in UME settings because the medical students are at lower Dreyfus levels of competence and there is ample room for improvement in knowledge of critical appraisal and EBM.

This review found only limited evidence of improvement in skills of critical appraisal with JC instruction in UME setting, also a level 2 feature in Kirkpatrick's framework. This finding contrasts with that of Honey and Baker in their review of work-based JCs which found an improvement in critical appraisal skills among participants in 12 out of the 16 included studies.[6] The difference could be explained by the obvious difference in the Dreyfus levels of participants included in the compared reviews, the undergraduate (novice, advanced beginner) vs work based participants (competent, proficient, expert).[30] Postgraduates and practicing physicians are more likely to have greater content mastery and situational perception than undergraduates and this will enable them to grasp, much easier, the important skill of critical appraisal of research literature.[3]

We found that the themes mentorship and the use of CAW were associated with desirable skills and attitude outcomes of JCs. These include improved skills of literature search, manuscript writing and a positive attitude towards JC instructional method. Our findings corroborate the work of others; mentorship and use of a critical appraisal

framework is associated with improved knowledge of EBM, enhanced skills of critical appraisal and the successful publishing of manuscripts.[3,9] We believe that mentorship provided structured and continuing support to JC intervention which is critical to the skill development.

The limitations of this systematic review include the low number of eligible studies included and their relatively low quality, the limited spread of published studies which probably reflects the predominant representation of the western medical education cultural model in the literature rather than the true state of real-world practice. We need greater numbers of methodologically rigorous studies to be published and primary research that are multi-institutional and more representative. Researchers should employ assessment instruments with reported validity and reliability and focus on clinically relevant outcomes like those of Kirkpatrick's levels 3 or 4.

Conclusion

This review provides a timely update of the pedagogical effectiveness of JCs in undergraduate medical education not captured by previous reviews. Overall, our findings suggests that the JC may be an effective instructional tool to teach EBM in UME particularly when coupled to lectures, a critical appraisal framework, or mentoring. We demonstrated that JCs are well-received by students though satisfaction fell with JCs longer than two hours. Therefore, we recommend that 1) JCs in UME should be less than two hours long, 2) coupled with lectures and mentorship to JCs and 3) include a critical appraisal tool, framework or process for JC instruction.

Conflicts of Interest: I declare that no conflicts of interest exist. This research was completed in partial fulfilment of an MSc in Medical Education degree from the University of South Wales, UK.
Acknowledgments: Many thanks to Paul Grant, my dissertation supervisor at the University of South Wales
Funding: No funding was received for this research.

References

1. Aronson JK. Journal Clubs: 1. Origins. *BMJ Evidence-Based Med.* 2017;22:231. https://doi.org/10.1136/ebmed-2017-110860
2. Straus S, Glasziou P, Richardson W, Haynes R. *Evidence-based medicine: how to practice and teach EBM*. Edinburgh: Elsevier; 2019.
3. Harris J, Kearley K, Heneghan C, et al. Are journal clubs effective in supporting evidence-based decision making? A systematic review BEME Guide No. 16. *Med Teach*. 2011;33(1):9-23. https://doi.org/10.3109/0142159X.2011.530321
4. Biswas T. Role of journal clubs in undergraduate medical education. *Indian J Community Med*. 2011;36(4):309-310. https://doi.org/10.4103/0970-0218.91422
5. Deenadayalan Y, Grimmer-Somers K, Prior M, Kumar S. How to run an effective journal club: a systematic review. *J Eval Clin Pract.* 2008;14:898-911. https://doi.org/10.1111/j.1365-2753.2008.01050.x
6. Honey CP, Baker JA. Exploring the impact of journal clubs: a systematic review. *Nurse Educ Today*. 2011;31(8):825-831. https://doi.org/10.1016/j.nedt.2010.12.020
7. Ebbert JO, Montori VM, Schultz HJ. The journal club in postgraduate medical education: a systematic review. *Med Teach*. 2001;23(5):455-461. https://doi.org/10.1080/01421590120075670
8. Coomarasamy A, Khan KS. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. *BMJ*. 2004;329(7473):1017. https://doi.org/10.1136/bmj.329.7473.1017
9. Edwards R, White M, Gray J, Fischbacher C. Use of a journal club and letter-writing exercise to teach critical appraisal to medical undergraduates. *Med Educ*. 2001;35:691-694. https://doi.org/10.1046/j.1365-2923.2001.00972.x
10. Banerjee Y, Azar AJ, Tuffnell C, Lansberg PJ, Bayoumi R, Davis D. A novel 6D-approach to radically transform undergraduate medical education: preliminary reflections from MBRU. *BMC Med Educ*. 2018; 18(1):304. https://doi.org/10.1186/s12909-018-1402-0
11. Willett LR, Kim S, Gochfeld M. Enlivening journal clubs using a modified 'jigsaw' method'. *Med Educ.* 2013; 47(11):1127-1128. https://doi.org/10.1111/medu.12342
12. Kirkpatrick DL, Kirkpatrick JD. *Kirkpatrick's four levels of training evaluation.* Alexandria, VA: ATD Press; 2016.
13. Praslova L. Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education. *Educ. Assessment, Eval. Account.* 2010;22:215–225. https://doi.org/10.1007/s11092-010-9098-7
14. Frye AW, Hemmer PA. Program evaluation models and related theories: AMEE guide no. 67. *Med Teach.* 2012;34(5):e288-e299. https://doi.org/10.3109/0142159X.2012.668637
15. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complx policy interventions. J Health Serv Res Policy. 2005;10(Suppl 1):21–34. https://doi.org/10.1258/1355819054308530
16. Reed DA, Beckman TJ, Wright SM, Levine RB, Kern DE, Cook DA. Predictive validity evidence for medical education research study quality instrument scores: quality of submissions to JGIM's Medical Education Special Issue. *J Gen Intern Med*. 2008;23(7):903-907. https://doi.org/10.1007/s11606-008-0664-3
17. Berman D, Braig Z, Simms B, et al. Efficacy of medical student surgery journal club. *J Surg Educ.* 2019;76(1):83-88. https://doi.org/10.1016/j.jsurg.2018.06.006
18. Marušić A, Malički M, Sambunjak D, Jerončić A, Marušić M. Teaching science throughout the six-year medical curriculum: two-year experience from the University of Split School of Medicine, Split, Croatia. *Acta Med Acad*. 2014;43(1):50-62. https://doi.org/0.5644/ama2006-124.100

19. Rosenthal J, Rosenthal KS. interactive journal club: teaching an old dog new tricks. *J Med Educ Curric Dev.* 2017;4:2382120517719710. https://doi.org/10.1177/2382120517719710
20. Williams AD, Mann BD. Improved knowledge gain and retention for third-year medical students during surgical journal club using basic science review: a pilot study. *Am J Surg.* 2017;213(2):238-243. http://dx.doi.org/10.1016/j.amjsurg.2016.09.055
21. Bahner DP, Royall NA. Advanced ultrasound training for fourth-year medical students: a novel training program at The Ohio State University College of Medicine. *Acad Med.* 2013;88(2):206-213. https://doi.org/10.1097/ACM.0b013e31827c562d
22. Chakraborti C. Teaching evidence-based medicine using team-based learning in journal clubs. *Med Educ.* 2011;45(5):516-517. https://doi.org/10.1111/j.1365-2923.2011.03972.x
23. Gokani SA, Sharma E, Sharma T, Moudhgalya, SV, Selvendran SS, Aggarwal N. Impact of a national journal club and letter writing session on improving medical students' confidence with critical appraisal. *Adv Med Educ Pract.* 2019;10:1081-1087. https://doi.org/10.2147/AMEP.S235260
24. Quinn EM, Cantillon P, Redmond HP, Bennett D. Surgical journal club as a community of practice: a case study. *J Surg Educ.* 2014;71(4):606-612. https://doi.org/10.1016/j.jsurg.2013.12.009
25. Lucia VC, Swanberg, S.M. Utilizing journal club to facilitate critical thinking in pre-clinical medical students. *Int J Med Educ*. 2018;9:7-8. https://doi.org/10.5116/ijme.5a46.2214
26. Bertelsen NS, DallaPiazza M, Hopkins MA, Ogedegbe G. Teaching global health with simulations and case discussions in a medical student selective. *Global Health.* 2015;11:28. https://doi.org/10.1186/s12992-015-0111-2
27. Curtis A, Viyasar T, Ahluwalia V, Lazarus K. Educating medical students: introducing a journal club. *Clin Teach*. 2016;13(3):233-234. https://doi.org/10.1111/tct.12356
28. Currier RL, Schneider MR, Heubi JE. Taking journal clubs off autopilot: a case study of teaching literature evaluation skills to preclinical md/phd students. *Med Sci Educ*. 2013;23(4):572-577. https://doi.org/10.1007/BF03341682
29. Hultman CS, Connolly A, Halvorson EG, et al. Get on your boots: preparing fourth-year medical students for a career in surgery, using a focused curriculum to teach the competency of professionalism. *J Surg Res. 2012;*177(2):217-223. https://doi.org/10.1016/j.jss.2012.06.019
30. Eraut M, Hirsh W. The significance of workplace learning for individuals, groups and organisations; 2007 Available from: http://www.skope.ox.ac.uk/wp-content/uploads/2014/12/Monogrpah-09.pdf