



About the efficacy of a serious game in critical appraisal learning: a cluster randomized controlled trial

A propos de l'efficacité d'un jeu sérieux dans l'apprentissage de la lecture critique d'articles médicaux : un essai contrôlé randomisé en grappes

Mona Mlika¹, Chadli Dziri², Faouzi Mezni¹

1. *Hopital Abderrahman Mami de l'Ariana / université Tunis el manar. Faculté de Médecine de Tunis*
2. *Faculté de Médecine de Tunis / Université Tunis el Manar*

ABSTRACT

Introduction: Serious games have been reported to be a valuable method of learning for over a decade.

Aim: The authors assessed the efficacy of using a serious games to teach critical appraisal practice to medical students in comparison to face-to-face learning methods.

Methods: This was a cluster randomised controlled trial including 3rd year medical students. The control group had to perform a critical appraisal of two articles guided by a checklist, while the intervention group performed a critical appraisal using a home-made serious game. Judgment criteria consisted of the students' scores attributed to their appraisal of a same article after the training period.

Results: Twenty-four students in the control and intervention group were enrolled in the study. There was no significant difference in scores between both groups. The satisfaction questionnaire highlighted more motivation and self-accomplishment feeling in the intervention group.

Conclusion: This study didn't show a significant difference between the scores of both groups but the motivation of the students included in the intervention group was improved and encourages the use of both methods in critical appraisal teaching.

Key-words: serious games, evidence-based-medicine, pedagogy, interactive learning.

RÉSUMÉ

Introduction: Les jeux sérieux ont été décrits comme étant une technique valable dans l'apprentissage depuis une décennie.

Objectif : Les auteurs ont évalué l'efficacité des jeux sérieux dans l'enseignement de la lecture critique d'articles médicaux en comparaison avec des techniques interactives d'enseignement présentiel.

Méthodes: les auteurs ont réalisé un essai contrôlé randomisé en grappes incluant les étudiants en troisième année médecine. Le groupe contrôle devait réaliser une lecture critique de 2 articles en utilisant des grilles publiées pré-établies. Le groupe interventionnel a utilisé un jeu sérieux disponible en ligne à partir de chez eux. La variable de jugement a été le score attribué aux étudiants après la lecture critique d'un même article. Par ailleurs, les étudiants ont rempli un questionnaire de satisfaction.

Résultats: 24 étudiants dans le groupe contrôle et 24 étudiants dans le groupe interventionnel ont été inclus. Il n'y avait pas de différence statistiquement significative dans les scores entre les 2 groupes. L'évaluation des questionnaires de satisfaction des étudiants a révélé plus de motivation et de sentiment d'accomplissement dans le groupe interventionnel.

Conclusion: cette étude n'a pas montré de différence significative entre les scores des étudiants des 2 groupes, cependant, la motivation des étudiants dans le groupe interventionnel encourage à l'utilisation des 2 techniques d'enseignement dans l'apprentissage de la lecture critique d'articles médicaux.

Mots-clés: jeux sérieux, médecine basée sur le niveau d'évidence, pédagogie, enseignement interactif.

Correspondance

Mona Mlika

Hopital Abderrahman Mami de l'Ariana / université Tunis el manar. Faculté de Médecine de Tunis

Email: mlika.zorgati.mona@hotmail.com

INTRODUCTION

A critical appraisal of current medical literature is a mandatory skill to deal with health problems. It is the third step of the evidence-based practice and the most challenging (1–3). Many universities promote this practice by introducing it into their curriculum. This learning is usually judged challenging by students because it is introduced early at a time students are still asking background questions and this practice seems easier during clerkship period when students are asking foreground questions (4, 5). Gamification concept has been introduced in many non medical fields in order to increase the motivation (4, 6). In order to make this practice more attractive to students, the authors developed a serious game and introduced it into the training of critical appraisal skills. Their aim was to compare the efficiency of the learning process and the motivation of the students between serious game-based learning and traditional lecture and demonstration-based learning.

METHODS

Population: Since 2013, 25 students in the third year of medical training are received in the Department of Pathology

every year. The students were divided into seven groups of 4 to 5 students with an interval period of 2 to 3 weeks between the different groups. The students were assigned randomly by the faculty of medicine. The training period lasts three weeks according to the students' portfolio. All the students that were assigned by the faculty and who gave their consent to participate in this study were included. The students that were performing their training period in other Departments were not included. The students that didn't want to be enrolled in this study were excluded.

Type of study: This was a cluster randomised controlled trial. This kind of randomization was used because of the university board's assignment.

Intervention: This interventional and longitudinal study was performed during a 2-year-period (2017/2018 and 2018/2019). The students' groups were randomly assigned to two groups by using a computerised random number allocation and were given the same pre and post-tests to answer. The test consisted of an adapted version of the Fresno Test that was represented in Figure 1. seven multiple-choice-questions in addition to 5 multi-choice questions relating to an article. The students fulfilled a Likert-type-scale satisfaction questionnaire to report their self-assessed competences and their motivation. Every suggestion was coded between 1 to 4.

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:

Methods

To assess the similarity between groups of participants, Wilcoxon non-parametric test was used for medians. Participants were compared using the Fisher's exact test. Sensitivity, specificity and their 95% confidence interval (CI) were calculated for each test using the confirmed cases as the reference standard and all cases (performed and not) as the second reference standard, according to the formula for reporting of diagnostic accuracy studies (Fleiss) (10).

Test	n	True positive results (TP)	Sensitivity (95% CI)	True negative results (TN)	Specificity (95% CI)
AFB9 assay ^a	53	1	28 (26, 30)	—	—
AF culture ^b	61	14	245 (240, 250)	—	—
Phlebotomy histopathological examination ^c	42	28	435 (405, 465)	—	—
AF AFB ^d	42	23	105 (75, 135)	22	876 (871, 881)
AF PCR ^e	40	1	28 (26, 30)	28	876 (871, 881)
CMR ^f	84	9	145 (135, 155)	24	888 (883, 893)
AF test ^g	85	2	25 (23, 27)	28	1020 (1015, 1025)

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 progesterone but you read recently that this association could interfere negatively with the milk production more than progesterone 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 adapted version of the Fresno test used to assess the students' skills in critical appraisal of medical literature.

Learning activities: Training objectives were listed in the students' university portfolio. These objectives were performed by the faculty staff. The objectives were divided into items related to technical skills, solving health problems, ethical attitude and the critical appraisal of a scientific publication.

A diary about the daily activities of the students was available in the e-portfolio of the students (7). It contained their different tasks in the Department and the different learning activities. Different methods of learning were displayed in the Department of Pathology including traditional methods which consisted of lectures dealing with the pathologies mentioned in the training board, integrative and participative methods consisting in case-based and problem-based learning (8, 9). Lectures were necessary because the students were not familiar with the concept of evidence-based medicine.

The five steps of the evidence-based-practice (EBP) were planned during the third week. In the four initial days, the tutor explained the elementary principles of evidence-based-medicine practice. Teaching techniques consisted

in lectures and problem-based learning. The details of the face-to-face learning was detailed by the authors in a previous publication (10). On the fifth day, the groups that were randomized to the control group participated to a workshop, which consisted of a collective critical appraisal session. The students were asked to critically appraise a case report using the guidelines entitled: "Spontaneous regression of locally advanced non small cell lung cancer" and "Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up" (11, 12, 13). The choice of case report and recommendations was motivated by the fact that the students weren't used to reading manuscripts and dealing with original manuscripts seems too challenging. The students were given a checklist of a PubMed-indexed journal in order to facilitate this practice.

The intervention group critically appraised the same manuscripts in conjunction with using the serious game developed by the tutor. The different steps of the randomization are represented in figure 2.

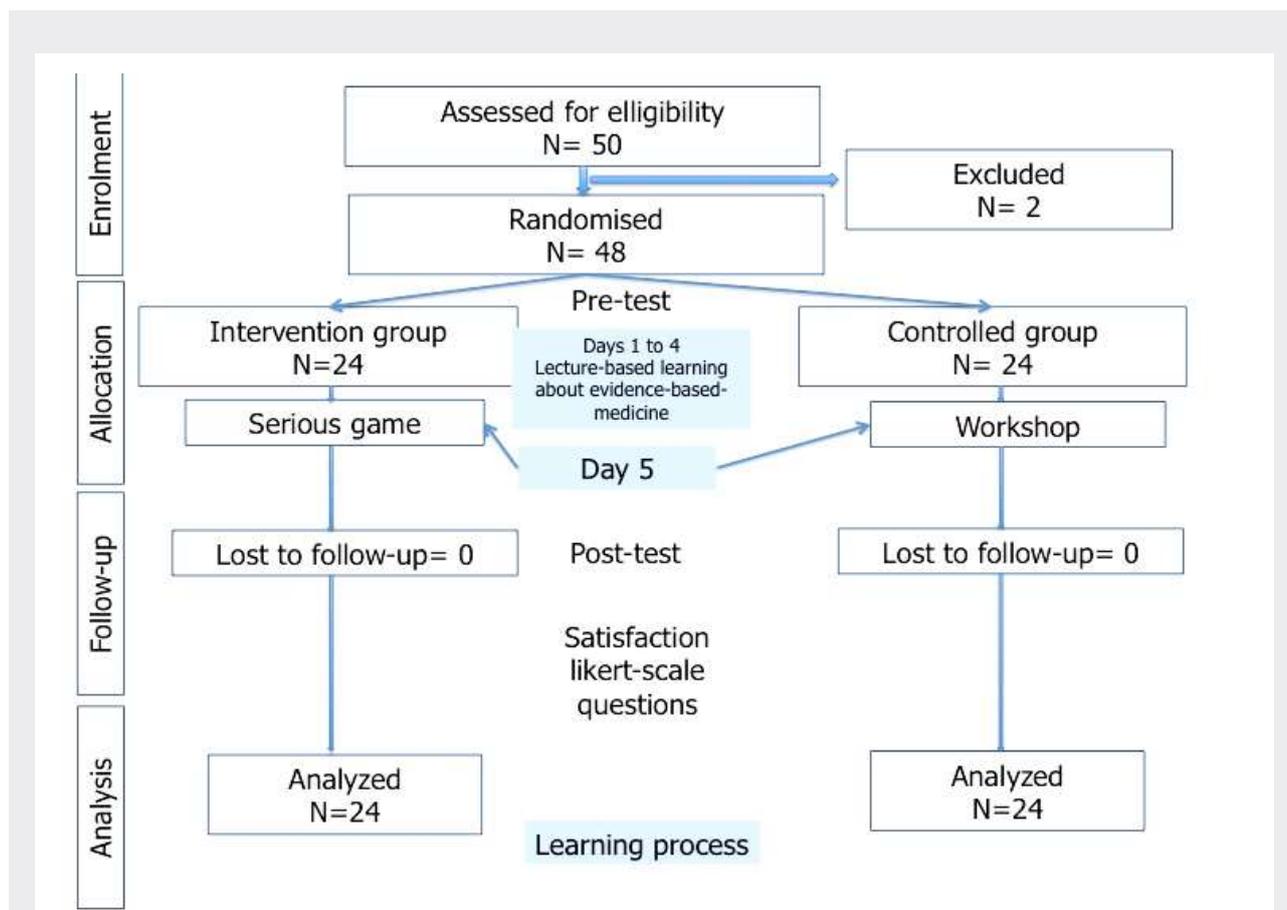


Figure 2. Study flowchart based on the Consolidated Standards of Reporting Trial (CONSORT) guidelines

Serious game: The authors developed a simulation-based serious game with design elements consisting of avatars, scores and explanations. This game was prepared using itystudio software. The different scenes took place in a medical consultation. The students had to choose the doctor's avatar. A scenario was performed in order to include the approach of critical appraisal into a realistic health problem. The scenario focused on a 50-year-old-patient visiting his family doctor to explore an episode of haemoptysis. This episode was explored and the medical doctor had to inform the patient of a diagnosis of lung cancer. In order to explain the disease to the patient and to prepare him to future investigations and treatment modalities. The doctor had to deal with a case report about a spontaneous regression of a

lung cancer then recommendations of the ESMO. With both articles downloaded by the students, a series of questions related to decision-making were asked and scored. When the students gave good answers, the game continued and if they failed the game was over. Every student had the opportunity to repeat the game. Three competencies were assessed: announcing bad news, the critical appraisal of a case report and the critical appraisal of recommendations. Every competency was scored and the player received the final scores represented into three axes in the end of the game. The scenario established in the game is represented in figure 3. Some illustrations of the game are represented in figure 4. The game was also made available through the Department site link.

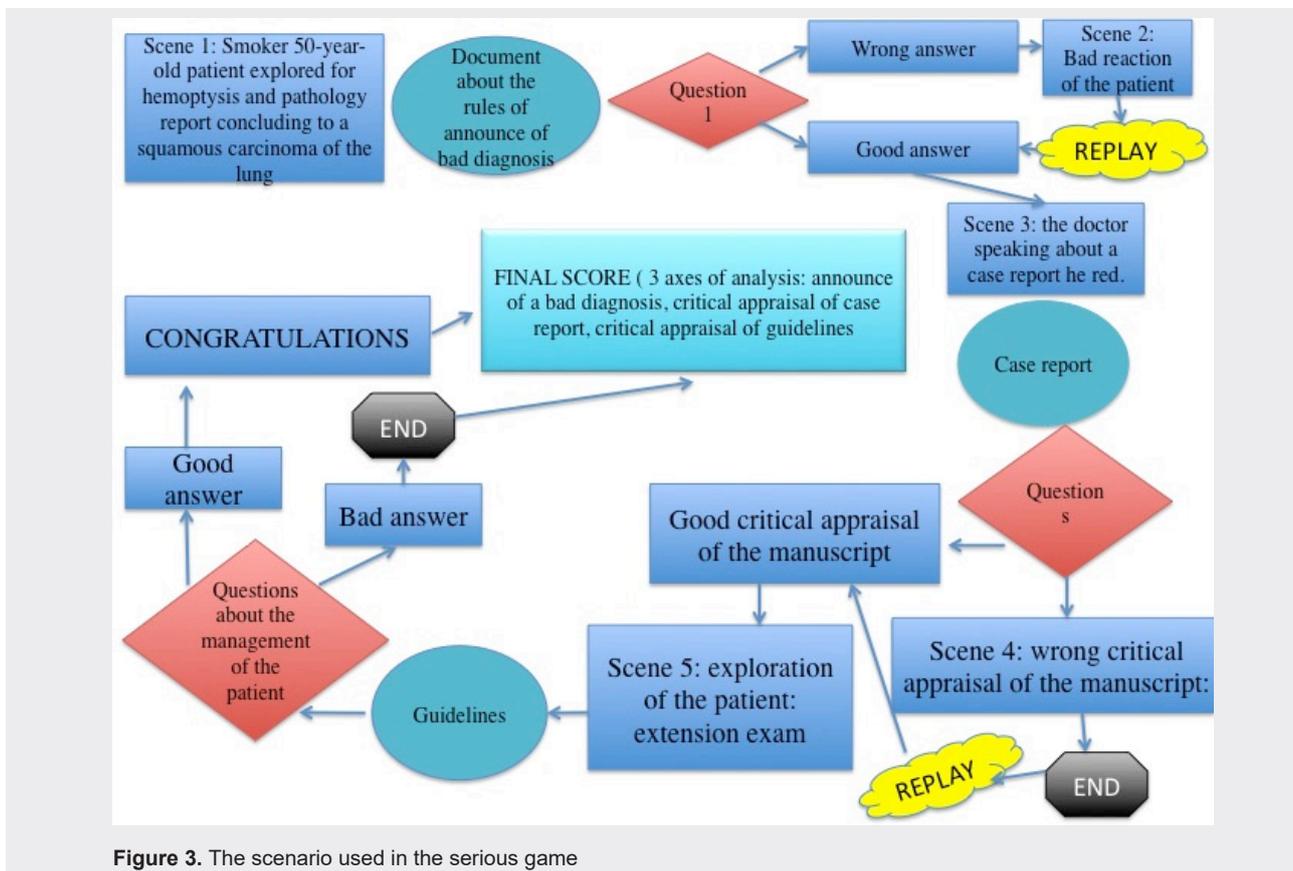


Figure 3. The scenario used in the serious game

Final test: The final test was inspired from the Fresno test (14). 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 weren't 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.

Judgment criteria: The principal judgment criteria consisted of the students' scores test. The secondary judgment criteria consisted of the students' satisfaction

Statistics: Statistical tests used to compare the scores between both groups consisted of non parametric tests with alpha levels set at 0.05. The authors used SPSS software (version 20.0).

Ethics: 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 informed consent from each respondent.

Clinical trial registration: This clinical trial was registered under the clinicalTrial.gov ID: NCT04076163

RESULTS

Forty-eight students were enrolled in the study (24 in the intervention group and 24 in the control group). The mean age of students was 21 years. Twelve men and 36 women agreed to participate in the study. Table 1 illustrates the characteristics of the students in both groups.

Table 1. Demographic characteristics of the students included in the study.

All students: Mean age: 21 years [average 19-22]. Sex ratio M/F: 12/36	
Control group	Intervention group
Mean age: 21 years [average 20-21]	Mean age: 21 years [average 19-22]
Sex ratio M/F: 7/17	Sex ratio: 6/18

The mean score of the students in the control group was 14/20 versus 13.5/20 in the intervention group. There was no significant difference between males and females. There was no significant difference between both groups ($p > 0.05$). The main results are represented in table 2.

Table 2. The different final test scores of critical appraisal attributed to the students

	Mean scores	Mean scores (Men)	Mean scores (Women)
Intervention group	15.02	14.5	15.5
Control group	14.75	14.25	15.25
P value	>0.05	>0.05	>0.05

Concerning the self-accomplishment feeling and the motivation, the students in the intervention group were more motivated and more likely to reproduce the experience with peers. Besides, they expressed their feeling of self-accomplishment concerning the critical appraisal practice's skill. Table 3 illustrates the

different results.

Table 3. The satisfaction questionnaire illustrating the different mean scores attributed by the students.

	Intervention group	Control group	p values
Means +/- SD			
I enjoyed the critical appraisal game learning	20.03+/-1.09	17.5+/-1.34	p=0.03
I improved my skills in critical appraisal of medical literature	45.13+/-0.97	20+/-2.23	p=0.00
I was motivated	45.24+/-0.5	23+/-1.23	p=0.025
I feel more comfortable with health problem-solving	20.4+/-1.2	14+/-0.97	p=0.01
I would like to share this learning experience with other learners	46.09+/-1.14	20+/-0.87	p=0.025

DISCUSSION

This study revealed no significant difference between the final scores of the intervention group (n=24) and the control group (n=24) using serious games when dealing with critical appraisal learning in comparison with traditional face-to-face methods. In a systematic review about the quality of evidence games, Gorbanev, et al. reported the lack of evidence concerning the efficiency of serious games (13). Other studies reported superiority of serious games in comparison with active methods. In another study, the authors compared clinical reasoning scores between serious games-based learning and problem-based learning. Students using serious games achieved better results than the others (13). This study's results put emphasis on the similarity of the pedagogical basis beyond both methods. In fact, serious games incorporate pedagogical principles consisting in motivation, interactivity, repetition and elements of scoring (14). Some authors have reported pedagogical advantages of serious games concerning the learning and evaluation of empathy (15, 16, 17). Olivier, et al. reported positive effects of a serious game on empathy and prejudice of psychology students towards persons with disabilities (19). Few studies in the literature have assessed the efficiency of serious games in comparison to traditional methods. This may be explained by the lack of use of serious games in medical education. In fact, the majority of the serious games described in the literature were resource-intensive and developed by professionals with rare free software or platforms (19–23). Most tutors have the tendency to avoid digital-based learning methods because they seem uncomfortable with these technologies. On the other hand, the students are more keen on them. There are many requirements

necessary to develop a serious game including expertise in medicine, education and technology development (24, 25, 26). All these skills may be difficult to achieve for tutors even if there are multiple softwares' packages that are available and that are easy to use and sufficient to achieve educational purposes. In a systematic review purposing to assess the efficiency of serious games, Wang, et al reported the heterogeneity of the studies reported in the literature with huge game types and different approaches (27). Even if serious games seemed not to increase the students' skills about critical appraisal, they seemed very motivating in this study. In fact, the students of the intervention group expressed a high level of motivation and mentioned that this kind of technology increased the relation learner-tutor. Serious games using avatars increases the intrinsic motivation of the students including their sense of autonomy and achievement (14). Busari, et al. reported in an inquiry-based study of serious games that the students' motivation toward serious games depended on their familiarity with games and the game's realism (28). In a literature review, McLeod, et al. concluded that serious games influence the extrinsic motivation much more than internal motivation which is mandatory in order to achieve long learning achievements (29).

The major limits of this study consist of the absence of calculation of the trainees needed to be included. This was due to the fact that the students are allocated by the Faculty to the different Departments. Besides, the low number of the students included didn't allow us to be conclusive and put emphasis on the necessity of taking our results with caution.

CONCLUSION

This study results highlight the absence of significant difference between using face-to-face learning methods or serious game-based learning when dealing with critical appraisal practice. On the other hand, the serious games used enhanced the students' motivation and self-accomplishment feeling. This may encourage using serious games in association with face-to-face learning in order to maintain the students' motivation.

Acknowledgements : We do thank the students who performed their training in the Department of Pathology of Abderrahman Mami Hospital.

REFERENCES

- Ahmadi N, Dubois L, McKenzie M, Brown CJ, MacLean AR, McLeod RS. Role of Evidence-Based Reviews in Surgery in teaching critical appraisal skills and in journal clubs. *Can J Surg.* 2013;56(4):98–102.
- Horsley T, Hyde C, Santesso N, Parkes J, Milne R, Stewart R. Teaching critical appraisal skills in healthcare settings (Review) Summary of findings for the main comparison. *Public Health.* 2011;(11):1-10
- Marus A. Teaching Students How to Read and Write Science . *Acad Med.* 2003;78(12):1235–9.
- Erazo-Palacios J, Gonzalez CR, Garcia-Mendoza B. Towards a Web Generator of Programming Games for Primary School Children. *Inter J Eng Pedag.* 2022;12:98-114
- Kataoka Y, Maeno T, Inaba T, Ninn S, Suzuki M, Maeno T. A qualitative study of factors promoting EBM learning among medical students in Japan. *Int J Med Educ.* 2022;13:215–20.
- Bahtiar AR, Segara AJT. Design of Smart Gamification In village Tourism: An Indonesian Case Study. *iJEP.* 2020;10:82-93.
- Mlika M, Ben Hassine L, Braham E, Mezni F. L'e-portfolio : un outil d'apprentissage dans un laboratoire de pathologie ?. L' innovation pédagogique pour améliorer l' apprentissage. 2015;2:20–30.
- Mlika M, Hassine L Ben, Braham E, Mrabet A, Mezni F. About the association of different methods of learning in the training of medical students in a Pathology Lab. *SJAMS.* 2015;3:1149–53.
- Mlika M, Khelil M Ben, Amri C, Mrabet A, Mezni F El. Using freemind software to teach pathology to medical students. *Int J Pedagog Learn.* 2017;12(1):17–24.
- A controlled randomized trial about using a checklist-based critical appraisal.
- Mlika M, Dziri C, Ben Khelil M, Zakhama L, Mezni F. *Tunis Med.* 2019;97(6):779-785.
- Ooi KH, Cheo T, Shook G, Soon T, Leong CN. Spontaneous regression of locally advanced nonsmall cell lung cancer. *Medicine.* 2018;97:2–5.
- Guidelines CP. Early and locally advanced non-small-cell lung cancer (NSCLC) : ESMO Clinical Practice Guidelines for diagnosis , treatment and follow-up †. *Ann Oncol.* 2017;28(Supplement 4):1–21.
- Gorbanev I, Agudelo-Iondoño S, González RA, Cortes A. A systematic review of serious games in medical education : quality of evidence and pedagogical strategy. *Med Educ Online.* 2018;23(1).
- Halm MA. Evaluating the impact of EBP education: Development of a modified Fresno Test for Acute care nursing. *Worldviews Evid Based Nurs.* 2018;15:272-280
- Chon S, Timmermann F, Dratsch T, Schuelper N, Berth F, Datta RR, et al. Serious Games in Surgical Medical Education : A Virtual Emergency Department as a Tool for Teaching Clinical Reasoning to Medical Students *JMIR serious games.* 2019;7:1–11.
- Jia E, Chee M, Prabhakaran L, Neo LP, Carpio GAC, Jia A, et al. Play and Learn with Patients — Designing and Evaluating Techniques : A Randomized Controlled Trial. *Games Health J.* 2019;8(3):1–8.
- Kafai YB, Burke Q, Kafai FYB, Kafai YB. Constructionist Gaming : Understanding the Benefits of Making Games for Learning Constructionist Gaming : Understanding the Benefits of Making Games for Learning. *Educ Psychologist.* 2016;1520.
- Anisimova T, Ganeeva A, Sharafeeva L. Development of digital skills and engineering thinking in students as part of the digital summer project. *iJEP.* 2021;11:69-81.
- Olivier L. The effect o f a serious game on empathy and prejudice of psychology students towards pe rsons with disabilities. *AJOD.* 2013;1–10.
- Chan W, Qin J, Chui Y, Heng P, Member S. A Serious Game for Learning Ultrasound-Guided Needle Placement Skills. *IEEE.* 2012;16(6):1032–42.
- Hannig A, Kuth N, Özman M, Jonas S, Spreckelsen C.

- eMedOffice : A web-based collaborative serious game for teaching optimal design of a medical practice. *BMC Med Educ.* 2012;104
23. Lagro J, Pol MHJ Van De, Laan A, Huijbregts-verheyden FJ, Fluit LCR, Rikkert MGMO. A Randomized Controlled Trial on Teaching Geriatric Medical Decision Making and Cost Consciousness With the Serious Game GeriatriX. *JAMDA.* 2014;4–9.
 24. Trial RC, Haubruck P, Nickel F, Ober J, Walker T, Bergdolt C. Evaluation of App-Based Serious Gaming as a Training Method in Teaching Chest Tube Insertion to Medical Students. *J Med Internet Res.* 2018;20:1-5.
 25. Kaczmarczyk J, Davidson R, Bryden D, Haselden S. Learning decision making through serious games. *Clin Teacher.* 2015;1–6.
 26. Greig ME, Christopher M. A serious game skills competition increases voluntary usage and proficiency of a virtual reality laparoscopic simulator during first-year surgical residents ' simulation curriculum. *Surg Ebdosc.* 2016;0–7.
 27. Olszewski AE, Wolbrink TA. Serious Gaming in Medical Education. *Technical Reports.* 2016;0(0):1–14.
 28. Wang R, Demaria S, Goldberg A, Katz D. A Systematic Review of Serious Games in Training Health Care Professionals. *Serious Games in Health Care Training.* 2016;11(1):41–51.
 29. Busari JO, Yaldiz H. Serious games as an educational strategy for management and leadership development in postgraduate medical education – an exploratory inquiry. *Adv Med Educ Practice.* 2018;571–9.
 30. Mcleod A, Hewitt B, Gibbs D, Kristof C. Evaluating Motivation for the Use of an Electronic Health Record Simulation Game. *Perspectives Health Info Management.* 2017;1:1-21.