Using the Timmer Scale to Standardize Pediatric Dentistry Residents’ Scientific Appraisal Skills

Samah Omar, DDS, MSD*, Janet Bauer, DDS, MSEd, MSPH, MBA, Afsaneh Matin, DDS, Amanjyot Bians, MAS, Jung-Wei Chen, DDS, MS, PhD, Monica Ocampo

*Corresponding Author: somar@llu.edu

Abstract

Introduction: There is no consensus on the best methodology to apply evidence-based practice principles to develop a systematic approach to improve critical appraisal or research design evaluation skills in advanced education journal clubs. Methods: We implemented a tool-based approach for our pediatric dentistry residents’ journal club centered on the use of a study quality assessment tool, the Timmer scale. The tool consisted of 19 standard questions that evaluated the research methodology, data collection, statistical analysis, and reporting of the findings of each article. Learners first underwent a 4-hour training session on study quality assessment. They were then assigned to read articles from monthly issues of core journals and appraised the quality of each article using the Timmer scale and submitted their scores in advance of the group session. Then, during a 1-hour journal club, the group came to a consensus on the Timmer scale score, and group and individual scores were compared to the course director’s scores as prompts for feedback and further discussion. Results: Over 3 years, 24 pediatric dentistry residents participated in the course. A noticeable improvement in the pediatric dentistry residents’ performance was noticed, with the discrepancy between their scores and group scores improving over time. Discussion: Using a quality assessment tool in journal clubs appeared to improve the residents’ ability to critically assess articles in a systematic way. Additionally, the tool was useful for assessing residents’ performance over time.

Keywords
Quality Assessment, Journal Club, Evidence-Based Practice, Pediatric Dentistry

Educational Objectives

By the end of this activity, participants will be able to:

1. Develop a systematic approach for critically appraising articles based on simple research design principles.
2. Evaluate the literature using a quality assessment tool (Timmer scale).
3. Improve critical thinking skills and article appraisal abilities.
4. Develop evidence-based practice principles in a journal club setup.

Introduction

Journal club seminars are typically a forum intended to review, discuss, and evaluate the professional literature from a specific discipline. Historically, pedagogical studies looking at the effectiveness of the journal club approach in health care education have mainly focused on the advantages, periodicity, and article selection.1,9 Numerous models have been published on how to establish and conduct journal clubs.1,3,5,8,11 However, none of these presented a specific validated tool or applied evidence-based practice (EBP) to develop a systematic approach to critical appraisal or research design evaluation. More recent studies of journal club pedagogy focused on the assessment of research articles using evidence-based methodologies.11,12 Ahmadi et al. performed a systematic review and found that while journal clubs seem to be the preferred way of teaching critical appraisal skills and evaluating the literature, there is a lack of consistency in terms of the appropriate format of the journal clubs and their effectiveness in determining the quality of individual articles.10 To the best of our knowledge, this is the first presentation of a systematic approach to teaching journal club using a validated tool.

The Timmer scale13 is an assessment tool developed to evaluate article quality. It has shown evidence of validity for assessing...
the quality of abstracts submitted to professional research conferences. The tool consists of 19 standard questions that evaluate the research methodology, data collection, statistical analysis, and reporting of the findings. Because it uses a scale rather than a checklist and articulates the domains in which a research article can be systematically analyzed, we decided to use it with pediatric dentistry residents to practice the skill of literature appraisal.

Since the 1970s, multiple assessment instruments have been introduced to provide methodology for critical appraisal in systematic reviews of the professional literature, including: Consolidated Standards of Reporting Trials (CONSORT), Timmer, Jadad, Assessment of Multiple Systematic Reviews (AMSTAR), Revised Assessment of Multiple Systematic Reviews (R-AMSTAR), Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), and Grading of Recommendations Assessment, Development and Evaluation (GRADE). While most of these tools have multiple forms for different study designs like Oxford and the Journal of the American Medical Association, Timmer, which has been validated and determined to be reliable for assessing research design methodology, has one form with standardized questions that critique the design, statistical methodology, and reporting of data. The specific study design is factored in using a specific formula that takes into account the randomization. It also has a numerical scale that allowed ranking of articles for quality. As a result, it is easier to train and calibrate residents to use the Timmer scale.

Methods

In 2014 we implemented a new systematic approach to evaluating and critiquing the articles discussed in our weekly journal club course using the Timmer scale. Journal club attendance was mandatory for all residents enrolled in the program and optional for dental students, staff, and the other program faculty. The discussion was usually facilitated by the course director with input from other attendees. The course director had prior training in principles of EBP and article review methodologies. A 4-hour training session was provided to all residents prior to the beginning of the journal club course (Appendices A and B). This training session introduced the residents to the objectives of EBP, basic research design principles, and trained the participants to use the quality assessment tool (Timmer scale). During this session, the residents practiced scoring a random article using the Timmer scale, with the course director’s guidance. The residents were also provided with a list of scientific articles (Appendix G) related to implementing EBP in dentistry to review before the training session.

Three or four different articles were read and assessed for every weekly session. Selected articles were acquired from the most current Pediatric Dentistry journal volume. The assigned article schedule was posted every 3 months or as a new journal issue was published. A minimum of 10-14 days was given to the residents to review and score the articles. Only clinical trials and systematic reviews were scored; in-vitro studies, literature reviews, and case reports were not scored. Unscored articles were included in the general discussion and were evaluated for the clinical and scientific relevance.

After reading an article, residents calculated a summary score (SS) using the Quality Assessment Score (QAS) Sheet (Appendix C) and the Study Design and Total Possible Points Form (Appendix D) which were simplified color-coded forms adapted from the original 19-question Timmer tool by the course director with permission from the creators. Residents first completed the QAS sheet. Each question was scored as completely applies (2 points), partially applies (1 point), or does not apply (0 points). The only exception was question 7, which was scored as 2 if the study was randomized and 0 if it was not. After each question was scored, the scores were summed for the QAS, which ranged from 0 to 38. After the QAS was determined, a study design score (SDS) and total possible points (TPP), assigned to the specific study design, were identified from the reported research methodology (Appendix D). All scores were then compiled to calculate an SS using the following formulas: For studies without randomized sample subjects, $SS = (QAS + SDS)/TPP$. For studies with randomized sample subjects, $SS = (QAS + SDS + 1)/TPP$. Calculated SS was a fraction number between 0.09 (if QAS = 0) and 1 (if QAS = 38); the closer the number to 1, the higher the quality of the study. In addition to the scoring, the residents were asked to describe the strengths and weaknesses of each article in a separate form (Appendix E). All article scores were sent to the course director 2 days prior to the journal club meeting for review. These scores were used for evaluating each resident’s performance.

The course director independently scored the assigned articles prior to the seminar and summarized the research design, QAS, SDS, TPP, and SS for each article using a summary sheet. The course director compiled the scores from all residents, faculty, and the group score, and presented them to the class in a graph format (Figures 1-4) weekly. Each resident
was assigned a random number at the beginning of the year so they could track their individual performance when the graph was shared in class. Immediate feedback was given anonymously to the whole group during the session and each article was discussed and critiqued by all participants, facilitated by the course director. After these discussions, the group determined a consensus group score for SS and QAS for each article.
At the end of every quarter, all participants filled out a course evaluation survey (Appendix F). When major misinterpretations of the critical appraisal process were identified, relevant reading assignments were used in 20-30 minutes remedial topic-focused training sessions (Appendix G). Information extracted from the survey along with the session’s feedback were used to modify and improve the precourse training session. Based on this information, we assigned readings to be reviewed before the training session (Appendix G).
The performance of residents was evaluated over a 3-year period to determine the impact of implementing the Timmer scale\textsuperscript{13} on our pediatric dentistry residents' ability to assess articles in a journal club environment. This was done based on the residents' performance, by comparing their scores to course director and group consensus scores over time. Secondly, the authors evaluated the applicability of the Timmer instrument as a teaching tool to assess teaching outcomes and improve evidence-based instructional material. We evaluated the process by comparing the residents' scores over the 3-year period.

**Results**

Eight pediatric dentistry residents (four first-year and four second-year) participated in journal club course every year over the 3 year study period (24 participants in total). The participants’ QAS scores for each year of implementation are presented in Figures 1-3. A noticeable improvement in the residents' critical appraisal skills was evidenced by the decrease in the discrepancy between their scores and group consensus scores over the course of each year of implementation. When looking at the residents' scores in the fall quarter, greater variation in their scores was noticed. Residents' performance improved over time as the scores became closer in value to each other and to the group consensus scores as they reached the end of the spring quarter (Figures 1-3). Discrepancies between the QAS graph (Figure 3) and the SS graph (Figure 4) indicated mistakes in study design identification or inaccurate SS formula calculation. When comparing residents' performance in different years, we found that there was lower discrepancy among the scores in the third year of implementation (Figure 3) compared to the first year of implementation (Figure 1).

**Discussion**

The utilization of an established quality assessment tool and structured worksheets improved residents' critical appraisal skills in a journal club setting. At the beginning of the fall quarter, residents were able to score the articles fairly, but did not show full understanding of the concepts demonstrated by the variability in the residents' scores. At the end of the year, resident scores showed improvement by scoring near the group consensus score. Before the implementation of this tool-based approach, the residents were using their own subjective judgments that were based on arbitrary evaluations influenced by their diverse backgrounds and experiences rather than a systematic and scientific approach. The structured Timmer scale\textsuperscript{13} helped residents focus on specific elements of the research methodology and data interpretation, all of which improved their critical appraisal skills.

In recent years, many advanced educational and training programs have adopted different methods and modules to integrate critical appraisal skills in their curriculum using EBP methodology.\textsuperscript{1,3,5,8,11} However, no consensus on a standard model has emerged. Although a different approach, our findings were consistent with Moher,\textsuperscript{4} who found that an evidence-based medicine curriculum improved residents' performance after attending at least six journal club sessions. Our results also add to the current body of literature on evaluating assessment instruments for implementation into journal club settings.\textsuperscript{7,12,26-31} Atzema reports that using a validated and well-structured assessment instrument may enhance the residents' assessment skills of research design methodology and reporting of findings.\textsuperscript{27} It may also give residents confidence to read research articles and extract clinically relevant quality evidence that will help them improve clinical care.

When comparing overall performance over several years, we noticed less variation among residents in the third year (Figure 3) compared to the first year (Figure 1). This demonstrated improvement in the instructional course over the years. This also indicated that critical appraisal and quality assessment in research design and methodology is a process that is acquired over a period of time and highlighted the importance of continuity and consistency of training. This supported the findings of earlier research in which authors questioned the retention of EBP knowledge and skills after graduation from training programs and suggested that skills and knowledge might be lost if not updated by continuous training or engagement in local journal clubs.\textsuperscript{8,32-34} Plotting all data in a graph prior to class helped the course director identify resident performance outliers or discrepancies between the residents' QAS or SS and the group scores. Discrepancies indicated errors in identifying study design, randomization, or SS calculation errors. Sharing the progress report in the format of a color-coded graph every session had a positive influence on participant's motivation and allowed for quick visualization of individual performance over time and identification of errors and mistakes. Discrepancies between QAS and SS graphs indicated errors in study design identification or inaccurate SS formula calculation. For example, after examining the graphs in Figures 3 and 4, it was readily apparent that two residents (R5 and R7) had discrepancies in QAS and SS of articles S3 and S4, indicating an error that needed
evaluation and discussion. The same was true for R7 relative to article W3. The resident made a mistake in identifying the research design, affecting the SS calculation. This error in the research design was discussed with the group to clarify any confusion. Based on residents’ feedback from course evaluations (Appendix F) and identification of common challenges in critical appraisal, modifications were implemented in the following year’s introductory course as described in the Methods.

Limitations of this educational approach included using a convenience sample of articles for the assessment. Thus, at each time period, all assessed articles differed in content, rendering it difficult to standardize. Our advanced training program is a small 2-year program, resulting in a small participant size of eight residents at all times. The heterogeneity of the participants, along with the small sample size, made it hard to validate our observations with statistical analysis. Longer-term studies with larger sample sizes would be beneficial to further support our findings.

We concluded that incorporating a validated assessment tool like the Timmer scale into journal club enhanced residents’ QAS and improved their approach to evaluating published research. The Timmer assessment tool13 was useful to monitor and evaluate residents’ performance in assessing quality of articles when compared to group consensus in a journal club setting. It is also a valuable teaching tool to improve critical appraisal and evidence-based teaching in residency programs.

Appendices

A. Introductory Course Material (EBP).pptx
B. Journal Club Course Introduction.pptx
C. Quality Assessment Score Sheet.docx
D. Study Design and Total Possible Points Form.docx
E. Articles Evaluation Form.docx
F. Course Evaluation Form.docx
G. Preclass and Remediation Reading Assignments.docx

All appendices are peer reviewed as integral parts of the Original Publication.

Samah Omar, DDS, MSD: Associate Professor, Pediatric Dentistry Department, School of Dentistry, Loma Linda University; ORCID: https://orcid.org/0000-0001-6110-0839

Janet Bauer, DDS, MSED, MSPH, MBA: Professor Emerita, UCLA School of Dentistry and School of Dentistry, Loma Linda University.

Afsaneh Matin, DDS: Assistant Professor, Pediatric Dentistry Department, School of Dentistry, Loma Linda University.

Amanjot Bians, MAS: Former Volunteer Research Assistant, Center of Dental Research, School of Dentistry, Loma Linda University.

Jung-Wei Chen, DDS, MS, PhD: Professor and Program Director, Advanced Specialty Education Program of Pediatric Dentistry, School of Dentistry, Loma Linda University.

Monica Ocampo: Former Volunteer Research Assistant, Center of Dental Research, School of Dentistry, Loma Linda University.

Disclosures
None to report.

Funding/Support
None to report.

Ethical Approval
Reported as not applicable.

References

1. Tam KW, Tsai LW, Wu CC, Wei PL, Wei CF, Chen SC. Using vote cards to encourage active participation and to improve critical appraisal skills in evidence-based medicine journal clubs. J Eval Clin Pract. 2011;17(4):827-831. https://doi.org/10.1111/j.1365-2753.2011.01711.x

2. McLeod RS, MacRae HM, McKenzie ME, Victor JC, Brasel KJ, Evidence Based Reviews in Surgery Steering Committee. A moderated journal club is more effective than an internet journal club in teaching critical appraisal skills: results of a multicenter randomized controlled trial. J Am Coll Surg. 2010;211(6):769-776. https://doi.org/10.1016/j.jamcollsurg.2010.08.016

3. Carpenter CR, Kane BG, Carter M, Lucas R, Wilbur LG, Graffeo CS. Incorporating evidence-based medicine into resident education: a CORD survey of faculty and resident expectations. Acad Emerg Med. 2010;17(Suppl 2):S54-S61. https://doi.org/10.1111/j.1553-2712.2010.00889.x

4. Mohr NM, Stoltze AJ, Harland KK, Van Heukelum JN, Hogrefe CP, Ahmed A. An evidence-based medicine curriculum implemented in journal club improves resident performance on the Fresno test. J Emerg Med. 2015;48(2):222-229. https://doi.org/10.1016/j.jemermed.2014.09.011

5. Bednarczyk J, Pauls M, Fridfinnsson J, Weldon E. Characteristics of evidence-based medicine training in Royal College of Physicians and Surgeons of Canada emergency medicine residencies—a national survey of program directors. BMC Med Educ. 2014;14:57. https://doi.org/10.1186/1472-6920-14-57

6. Pato MT, Cobb RT, Lusskin SL, Schardt C. Journal club for faculty or residents: a model for lifelong learning and maintenance of certification. Int Rev Psychiatry. 2013;25(3):276-283. https://doi.org/10.3109/09540261.2013.793172
7. Jude S, Antaki F. Approach to presenting a clinical journal club. *Gastroenterology*. 2014;146(7):1591-1593. https://doi.org/10.1053/j.gastro.2014.04.024

8. Serghi A, Goebert DA, Andrade NN, Hishinuma ES, Lunsford RM, Matsuda NM. One model of residency journal clubs with multifaceted support. *Teach Learn Med.* 2015;27(3):329-340. https://doi.org/10.1080/10401334.2015.1044658

9. Bowen AC, Conell TC, Bryant PA. Evaluating a web-based paediatric infectious diseases journal club: more than just critical appraisal? *BMC Med Educ.* 2014;14:242. https://doi.org/10.1186/s12909-014-0242-9

10. Ahmadi N, McKenzie ME, MacLean A, et al.; Evidence-Based Reviews in Surgery Steering Group. Teaching evidence based medicine to surgery residents-is journal club the best format? A systematic review of the literature. *J Surg Educ.* 2012;69(1):91-100. https://doi.org/10.1016/j.jsurg.2011.07.004

11. 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):E98-E102. https://doi.org/10.1503/cjs.007512

12. Dirschl DR, Tornetta P III, Bhandari M. Designing, conducting, and evaluating journal clubs in orthopaedic surgery. *Clin Orthop Relat Res.* 2003;413:146-157. https://doi.org/10.1097/01.blo.0000081203.51121.25

13. Timmer A, Sutherland LR, Hilsden RJ. Development and evaluation of a quality score for abstracts. *BMC Med Res Methodol.* 2003;3:2. https://doi.org/10.1186/1471-2288-3-2

14. Altman DG, Schulz KF, Moher D, et al.; CONSORT GROUP. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134(8):663-694. https://doi.org/10.7326/0003-4819-134-8-200104170-00012

15. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17(1):1-12. https://doi.org/10.1016/0197-2456(95)00134-4

16. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol.* 2007;7(10). https://doi.org/10.1186/1471-2288-7-10

17. Kung J, Chiappelli F, Cajulis OS, et al. From systematic reviews to clinical recommendations for evidence-based health care: validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *Open Dent J.* 2010;4:84-91. https://doi.org/10.2174/1874210601004020084

18. Welcome to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) website! PRISMA. Accessed 2018. http://prisma-statement.org

19. Atkins D, Best D, Briss PA, et al.; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ.* 2004;328(7454):1490. https://doi.org/10.1136/bmj.328.7454.1490

20. Phi L, Ajaj RA, Ramchandani MH, et al. Expanding the grading of recommendations assessment, development, and evaluation (Ex-GRADE) for evidence-based clinical recommendations: validation study. *Open Dent J.* 2012;6:31-40. https://doi.org/10.2174/1874210601206010031

21. Teich ST, Heima M, Lang L. Dental students’ use of AMSTAR to critically appraise systematic reviews. *J Dent Educ.* 2015;79(9):1031-1039.

22. Guyatt G, Drummond Rennie, Meade MO, Cook DJ. *User’s Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice.* 3rd Edition. American Medical Association. 2015. Accessed December 27, 2017. https://jamaevidence.com/resource/520

23. *Critical Appraisal Tools.* Centre for Evidence-Based Medicine. Accessed 2018. https://www.cebm.ox.ac.uk/resources/ebm-tools/critical-appraisal-tools

24. Agahi RH, Navabi N, Shahrvan A, Ghassemi A. Critical appraisal of reporting randomized clinical trials published in Iranian dental journals during 2003-2010. *J Dent (Tehran).* 2014;11(3):310-318.

25. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol.* 2009;62(10):1013-1020. https://doi.org/10.1016/j.jclinepi.2008.10.009

26. Grant WD. An evidence-based journal club for dental residents in a GPR program. *J Dent Educ.* 2005;69(6):681-686.

27. Atzema C. Presenting at journal club: a guide. *Ann Emerg Med.* 2004;44(2):169-174. https://doi.org/10.1016/j.annemergmed.2004.03.035

28. Lao WS, Puligandla P, Baird R. A pilot investigation of a pediatric surgery journal club. *J Pediatr Surg.* 2014;49(5):811-814. https://doi.org/10.1016/j.jpedsurg.2014.02.068

29. Merjohr GK. The evidence-based study club: a practical means to facilitate EBD awareness and implementation. *J Evid Base Dent Pract.* 2008;8(3):152-154. https://doi.org/10.1016/j.jebdp.2008.05.017

30. Moharani RS, Rahimi E, Najafi A, Khashayar P, Khajavi MR, Meysamie AP. Teaching critical appraisal and statistics in anesthesia journal club. *QJM.* 2009;102(2):139-141. https://doi.org/10.1093/qjmed/hcn131

31. Deenadayalan Y, Gimmer-Somers K, Prior M, Kumar S. How to run an effective journal club: a systematic review. *J Eval Clin Pract.* 2008;14(5):898-911. https://doi.org/10.1111/j.1365-2753.2008.01050.x

32. Harris J, Kearley K, Heneghan C, et al. Are journal clubs effective in supporting evidence-based decision making? A systematic...
33. 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

34. Matthews DC. Journal clubs most effective if tailored to learner needs. *Evid Based Dent*. 2011;12(3):92-93. https://doi.org/10.1038/sj.ebd.6400818

Received: April 20, 2020
Accepted: November 14, 2020
Published: February 12, 2021