

Randomized Controlled Trial of Spaced Education for Pediatric Residency Education

ERIN F. MATHES, MD
 ILONA J. FRIEDEN, MD
 CHRISTINE S. CHO, MD, MPH, MED
 CHRISTY KIM BOSCARDIN, PHD

Abstract

Background Spaced education (SE) has shown promise as an instructional tool that uses repeated exposure to the same questions, but information on its utility in graduate medical education is limited, particularly in assessing knowledge gain with outcome measures that are different from repeat exposure to the intervention questions.

Objective We examined whether SE is an effective instructional tool for pediatrics residents learning dermatology using an outcome measure that included both unique and isomorphic questions.

Methods We randomized 81 pediatrics residents into 2 groups. Group A completed an SE course on atopic dermatitis and warts and molluscum. Group B

completed an SE course on acne and melanocytic nevi. Each course consisted of 24 validated SE items (question, answer, and explanation) delivered 2 at a time in 2 e-mails per week. Both groups completed a pretest and posttest on all 4 topics. Each group served as the comparison for the other group.

Results Fifty residents (60%) completed the study. The course did not have a statistically significant effect on the posttest scores for either group. Overall, test scores were low. Eighty-eight percent of residents indicated that they would like to participate in future SE courses.

Conclusions Using primarily novel posttest questions, this study did not demonstrate the significant knowledge gains that other investigators have found with SE.

Introduction

As work hour limits change and time for formal education in residency decreases, primary care residencies struggle to fit essential subspecialty instruction into limited didactic time. Despite the high frequency of skin-related complaints in pediatrics, dermatology remains a significant knowledge gap for pediatricians.¹⁻¹¹ Online learning tools can provide educational opportunities unrestricted by time or location.

Online spaced education (SE) has been shown to be an effective and acceptable method for addressing knowledge

and practice gaps in medical education.¹²⁻¹⁷ The SE approach is based on the spacing effect—the principle that exposure to concepts distributed and repeated over time improves learning.¹⁸ Online SE with questions uses both the spacing effect and the testing effect (the principle that testing improves knowledge retention while evaluating a learner's performance). In urology studies, Kerfoot et al¹⁶ demonstrated knowledge gains that persist for up to 2 years.

The method's promising results as an instruction tool have led to the increasing use of SE in medical education. Kerfoot and colleagues measured the efficacy of SE with improved scores on repeated administration of the same questions.^{13,14,16,17} We sought to assess the generalizability of SE across different domains by creating an innovative SE program with clinical images of 4 common skin conditions. Our hypothesis was that SE would be an acceptable learning method and would significantly improve pediatrics residents' knowledge of pediatric dermatology when knowledge gain and application were assessed with an outcome measure different from the intervention (unique and isomorphic questions).

Methods

Design

Participating residents were randomized into 2 groups (groups A and B) and exposed to 2 different courses. Group

All authors are at the University of California, San Francisco. **Erin F. Mathes, MD**, is Assistant Professor of Dermatology and Pediatrics; **Ilona J. Frieden, MD**, is Professor of Dermatology and Pediatrics; **Christine S. Cho, MD, MPH, Med**, is Assistant Professor of Clinical Pediatrics and Clinical Emergency Medicine; and **Christy Kim Boscardin, PhD**, is Assistant Professor, Department of Medicine, Office of Medical Education.

Funding: This work was funded in part by a joint fellowship award to Dr Mathes from the Society for Pediatric Dermatology and the Dermatology Foundation.

Conflict of interest: The authors declare they have no competing interests.

The authors would like to thank B. Price Kerfoot for his guidance in the planning stages of the study and spaceded.com for providing a secure site for the study.

Corresponding author: Erin F. Mathes, MD, 1701 Divisadero Street, Box 0316, San Francisco, CA 94143-0316, 415.353.7880, mathese@derm.ucsf.edu

Received February 18, 2013; revisions received September 18, 2013, and November 12, 2013; accepted December 2, 2013.

DOI: <http://dx.doi.org/10.4300/JGME-D-13-00056.1>

A took an SE course on atopic dermatitis (AD) and warts and molluscum (WM). Group B took an SE course on acne and melanocytic nevi (MN). All residents took the same pretest (to ensure adequate randomization) and posttest. This design allowed us to compare group A's scores on AD/WM posttest items with group B's scores on the same items (comparison group) and to compare group B's scores on acne/MN items with group A's scores on the same items.

Sample

Of the possible 158 pediatrics residents from 2 residency programs (University of California, San Francisco [UCSF], $n = 36$, 44%; and Children's Hospital and Research Center Oakland [CHRCO], $n = 45$, 56%), 81 volunteered to participate. Our a priori power calculation indicated that a sample of 84 residents would achieve a power of 0.9 to detect a moderate standardized effect size of 0.5 (Cohen's d statistic) with a 2-sided 0.05 significance level. The assumption of a moderate effect size was based on previously reported moderate to large effect sizes in SE studies.^{13,17}

The Institutional Review Board of CHRCO and the Committee on Human Research of UCSF approved this study. Residents implied consent by taking the pretest.

Intervention

The structure of our SE courses is detailed in TABLE 1. Each SE course consisted of 24 SE items (12 for each of 2 topics, AD/WM or acne/MN).

Content Validity Analysis for Selection of Test Items

All test items addressed objectives modified from the American Board of Pediatrics Content Outline¹⁹ and the Academic Pediatric Association Educational Guidelines for Pediatric Residency²⁰ for acne, WM, MN, and AD. Fourteen academic pediatric dermatologists with an average of 11.5 years in practice (range, 1–30 years) from 8

What was known

Spaced education (SE), on the principle that exposure to content distributed and repeated over time improves learning, has not undergone significant testing in residents.

What is new

Research using posttest assessment that differed from the original questions in a population of pediatrics residents learning dermatology content did not reproduce knowledge gains found in prior studies of SE.

Limitations

The small sample resulted in insufficient power to detect small difference; the low pretest internal consistency, potentially due to guessing by participants.

Bottom line

Although well accepted by learners, SE was not found to be effective in this study population.

institutions in the United States reviewed and rated 32 objectives on a 4-point scale of importance for pediatrics residency curriculum. We used the content validity index and asymmetric confidence interval to quantify the relevance of the proposed content areas.

An item bank with 100 test items was developed based on the validated objectives. Items were reviewed by 10 pediatric dermatologists and pediatricians and modified. Nineteen dermatology residents and recently graduated pediatricians pilot tested the items to determine item difficulty and discrimination. Items that were too easy (100% correct) or with low discriminatory power (0.00) were omitted or modified for the final assessment tools (2 questions omitted, 8 modified). The Cronbach coefficient α was 0.73 based on pilot data.

Instrument

Resident knowledge was measured by the pretest and posttest. The pretest and posttest consisted of 28 primarily

TABLE 1 STRUCTURE OF SPACED EDUCATION (SE) COURSES

SE items (24 per course)	1. Question with self-assessment	Photograph and clinical scenario Multiple-choice question
	2. Brief Didactic	Correct answer Other learners' answers Explanation
	3. Practice Effect	Repeated exposure to question
Course parameters	SE item delivery	2 questions per e-mail 2 e-mails per week
	Answer incorrect	Resent every 8 days until correct
	Answer correct	Resent in 21 days
	Answered correctly twice	Question retired

TABLE 2 PRETEST AND POSTTEST SCORE DIFFERENCES BETWEEN GROUP A AND GROUP B

Scores	Group A (AD and WM), n = 27	Group B (Acne and MN), n = 23	P Value
AD and WM item pretest scores	52% (SD = 3%)	59% (SD = 3%)	.06
AD and WM item posttest scores	46% (SD = 5%)	52% (SD = 4%) comparison group	.38
Acne and MN item pretest scores	49% (SD = 2%)	55% (SD = 3%)	.19
Acne and MN item posttest scores	50% (SD = 6%) comparison group	65% (SD = 5%)	.10

Abbreviations: AD, atopic dermatitis; WM, warts and molluscum; MN, melanocytic nevi.

unique knowledge test items with clinical photographs targeting all 4 topics in the intervention (7 items per category). The pretest included questions on demographics, dermatology experience, and knowledge self-assessment. The posttest included questions on acceptability of the SE program and time to complete each SE item. The Cronbach coefficient α was 0.46 for the pretest and 0.91 for the posttest.

Procedure

All pediatrics residents at UCSF and CHRCO were invited to take the pretest. Stratified random sampling using postgraduate year (PGY) level and a random number generator randomized residents into groups A and B; 27 residents were in group A and 23 in group B. One of the authors (E.F.M.) was the course administrator and was not blinded to resident assignments. The courses were administered via spaceded.com from February 2010 to July 2010 with a choice of standard website or mobile access. We recorded how many items were retired as a measure of implementation. At the end of the study period, regardless of whether they had completed all of the items in the SE course, residents were given a posttest.

Analysis

First, analysis of variance was used to compare differences in the group outcomes. Then, analyses of covariance were performed to examine the effect of the 2 interventions (group A versus group B) on posttest scores. Pretest scores and PGY were included as covariates in the model.

Results

Baseline Resident Characteristics

Eighty-one residents were randomized. Pretest and posttest data were available for only 50 residents (62% retention rate). Groups A and B had similar demographics and pretest scores (TABLE 2).

Efficacy and Acceptability of SE Modules

No difference was found between posttest scores on AD/WM items for group A compared with its comparison group B. For the questions on acne/MN, group B's score was higher than that of the comparison group A (TABLE 2). Although this difference was not statistically significant, based on the calculations of the effect size (Cohen's d) using the estimated marginal means, the difference was in the moderate effect size range ($d = 0.45$).

After controlling for PGY and pretest scores, the differences between groups A and B were not statistically significant (TABLE 3). Furthermore, PGY was positively associated with performance on AD/WM items but not on acne/MN items. The implementation variable (number of items retired) was not significantly associated with performance for AD/WM or acne/MN items. The average number of items retired was 16 of 24 for group A and 17 of 24 for group B.

When asked if they would like to participate in future programs using SE, 88% (44 of 50) of the residents who completed the posttest answered yes. Approximately three-quarters of the residents completed the questions in < 5 minutes a day.

Discussion

The promising results for SE have led to its increasing use in medical education, but to date there has been relatively little application in graduate medical education. This is the first study to expand research on SE into dermatology education. This study also assessed knowledge gain and application with an outcome measure that was different from intervention questions. Neither online module had a statistically significant effect on posttest scores.

Eighty-eight percent of residents indicated that they would like to participate in future SE courses. This high level of acceptability has been found in other SE studies¹²⁻¹⁷ but may reflect selection bias, as residents who complete

TABLE 3 EFFECT OF THE SPACED EDUCATION MODULES ON AD & WM AND ACNE & MN KNOWLEDGE, CONTROLLING FOR POSTGRADUATE YEAR AND PRETEST SCORE^a

	AD and WM (14 Items)		Acne and MN (14 Items)	
	Beta Coefficients (95% CI)	P Value	Beta Coefficients (95% CI)	P Value 95% CI
Treatment versus comparison	-0.83 (-2.77, 1.11)	.40	2.25 (-0.10, 4.61)	.06
Postgraduate year	1.19 (0.05, 2.33)	.04	0.98 (-0.48, 2.43)	.19
Pretest score	0.30 (-0.22, 0.81)	.26	-0.02 (-0.59, 0.55)	.95

Abbreviations: AD, atopic dermatitis; WM, warts and molluscum; MN, melanocytic nevi; CI, confidence interval.

^aN = 50.

the studies may be more motivated by SE. Nonetheless, SE addresses components of adult learning theory, such as self-directed learning and integration with the demands of busy life.²¹⁻²³

Our findings differed from those of prior studies showing that SE leads to significant gains in short-term learning and long-term retention in medical education. In designing their studies, Kerfoot et al used repeat exposures to the same questions to demonstrate learning.^{13,14,16,17} Most of the questions on our pretest and posttest were unique, and there were few identical and isomorphic anchor questions. Our participants' knowledge retention was not significant. One explanation may be that previous findings do not reflect knowledge acquisition but memorization and/or recall of question answers. We were also using the commercially available spaced.com delivery system without monetary incentives, so our enrollment and retention were lower than in previous studies, although likely more realistic and generalizable.

Another important finding in our study was that residents from both residency programs had low test scores. Given low baseline knowledge, it is possible that our SE courses did not offer substantial enough didactics to teach residents difficult and nuanced concepts to which they had not been exposed or that contradicted their prior learning.

Our study has several limitations. We had a relatively small number of subjects and a relatively high attrition rate, resulting in insufficient power to detect smaller differences between groups. Our test questions were validated and piloted, but not on more junior pediatrics residents, which could partly explain the low test scores. In addition, minor differences in the content covered by the pretests, intervention tests, and posttests may have contributed to the differences in performance for the 2 courses. Although the pretest reliability was lower than the posttest reliability, we hypothesize that guessing in the pretest contributed to

lower internal consistency. The pretest was used to confirm that the groups had similar baseline knowledge and was not used to demonstrate knowledge improvement within a group. The pilot test reliability estimates of 0.73 gave us some assurance of reliability of the test items before the study. What is not known is the effect of the roughly two-thirds of items that were retired because of successful completion in the course, how these relate to items on the final test, and whether they were comparable for the 2 courses. Differences in the numbers of retired items may partially explain some of the surprising results. Future research might also assess whether the timing of when items are encountered by learners within the SE course (and the time to testing) affects retention.

Conclusion

Our study did not demonstrate the impressive knowledge gains that other investigators have found with SE. Although its flexibility and acceptability make SE a potentially attractive option for teaching in residency, this study indicates that SE cannot replace didactics or clinical rotations. Further study in other content areas using novel questions to measure performance is needed to determine the effectiveness and best use of SE in residency education.

References

- 1 Tunnessen WW Jr. A survey of skin disorders seen in pediatric general and dermatology clinics. *Pediatr Dermatol*. 1984;1(3):219-222.
- 2 Hayden GF. Skin diseases encountered in a pediatric clinic. A one-year prospective study. *Am J Dis Child*. 1985;139(1):36-38.
- 3 Dolan OM, Bingham EA, Glasgow JF, Burrows D, Corbett JR. An audit of dermatology in a paediatric accident and emergency department. *J Accid Emerg Med*. 1994;11(3):158-161.
- 4 Auvin S, Imiela A, Catteau B, Hue V, Martinot A. Paediatric skin disorders encountered in an emergency hospital facility: a prospective study. *Acta Derm Venereol*. 2004;84(6):451-454.
- 5 Hansra NK, O'Sullivan P, Chen CL, Berger TG. Medical school dermatology curriculum: are we adequately preparing primary care physicians? *J Am Acad Dermatol*. 2009;61(1):23.e1-29.e1. Epub 2009 May 5.
- 6 Prose N. Dermatology training during the pediatric residency. *Clin Pediatr (Phila)*. 1988;27(2):100-103.

- 7 Macnab A, Martin J, Duffy D, Murray G. Measurement of how well a paediatric training programme prepares graduates for their chosen career paths. *Med Educ.* 1998;32(4):362–366.
- 8 Hester EJ, McNealy KM, Kelloff JN, Diaz PH, Weston WL, Morelli JG, et al. Demand outstrips supply of US pediatric dermatologists: results from a national survey. *J Am Acad Dermatol.* 2004;50(3):431–434.
- 9 Pletcher BA, Rimsza ME, Cull WL, Shipman SA, Shugerman RP, O'Connor KG. Primary care pediatricians' satisfaction with subspecialty care, perceived supply and barriers to care. *J Pediatr.* 2010;156(6):1011–1015.
- 10 Enk CD, Gilead L, Smolovich I, Cohen R. Diagnostic performance and retention of acquired skills after dermatology elective. *Int J Dermatol.* 2003;42(10):812–815.
- 11 Sherertz EF. Learning dermatology on a dermatology elective. *Int J Dermatol.* 1990;29(5):345–348.
- 12 Kerfoot BP, Brotschi E. Online SE to teach urology to medical students: a multi-institutional randomized trial. *Am J Surg.* 2009;197(1):89–95. Epub 2008 July 9.
- 13 Kerfoot BP, Kearney MC, Connelly D, Ritchey ML. Interactive spaced education to assess and improve knowledge of clinical practice guidelines: a randomized controlled trial. *Ann Surg.* 2009;249(5):744–749.
- 14 Kerfoot BP, Fu Y, Baker H, Connelly D, Ritchey ML, Genega EM. Online spaced education generates transfer and improves long-term retention of diagnostic skills: a randomized controlled trial. *J Am Coll Surg.* 2010;211(3):331.e1–337.e1. Epub 2010 July 13.
- 15 Shaw T, Long A, Chopra S, Kerfoot BP. Impact on clinical behavior of face-to-face continuing medical education blended with online spaced education: a randomized controlled trial. *J Contin Educ Health Prof.* 2011;31(2):103–108.
- 16 Kerfoot BP. Learning benefits of on-line spaced education persist for 2 years. *J Urol.* 2009;181(6):2671–2673. Epub 2009 April 16.
- 17 Kerfoot BP, Baker H, Pangaro L, Agarwal K, Taffet G, Mechaber AJ, et al. An online spaced-education game to teach and assess medical students: a multi-institutional prospective trial. *Acad Med.* 2012;87(10):1443–1449.
- 18 Pashler H, Rohrer D, Cepeda NJ, Carpenter SK. Enhancing learning and retarding forgetting: choices and consequences. *Psychon Bull Rev.* 2007;14(2):187–193.
- 19 American Board of Pediatrics. Content Outline: General Pediatrics. <https://www.abp.org/abpwebsite/moc/cognitiveexpertisecureexam/contentandpreparation/gpmoc.pdf>. Accessed May 16, 2009.
- 20 Kittredge D, Baldwin CD, Bar-on ME, Beach PS, Trimm RF, eds. APA Educational Guidelines for Pediatric Residency. 2004. www.ambpeds.org/egweb. Accessed May 17, 2009.
- 21 Kopp SL, Smith HM. Developing effective web-based regional anesthesia education: a randomized study evaluating case-based versus non-case-based module design. *Reg Anesth Pain Med.* 2011;36(4):336–342.
- 22 Kosower E, Berman N. Comparison of pediatric resident and faculty learning styles: implications for medical education. *Am J Med Sci.* 1996;312(5):214–218.
- 23 Jack MC, Kenkare SB, Saville BR, Beidler SK, Saba SC, West AN, et al. Improving education under work-hour restrictions: comparing learning and teaching preferences of faculty, residents, and students. *J Surg Educ.* 2010;67(5):290–296.