The American Journal of Surgery (2009) 197, 89–95

The American Journal of Surgery®

Association for Surgical Education

Online spaced education to teach urology to medical students: a multi-institutional randomized trial

B. Price Kerfoot, M.D., Ed.M.[a,b,*], Erica Brotschi, M.D.[a,c]

[a]*Veterans Affairs Boston Healthcare System, Boston, MA, USA;* [b]*Harvard Medical School, Boston, MA, USA;* [c]*Boston University School of Medicine, Boston, MA, USA*

KEYWORDS:
Medical students;
Urology;
Undergraduate medical education;
Educational technology

Abstract

BACKGROUND: We investigated whether online spaced education could prospectively improve students' acquisition and retention of knowledge.

METHODS: One hundred fifteen third-year medical students at 2 schools were randomized to receive weekly/biweekly spaced education e-mails on 2 of 4 urology topics: prostate cancer (PC) and prostate-specific antigen (PSA) screening, or benign prostatic hyperplasia (BPH) and erectile dysfunction (ED). E-mails began in month 1 of their third year. During their 3-month surgery clerkships, students completed a 28-item validated pre-test on all 4 topics, 8 web-based teaching cases, and a 28-item post-test. This test was administered again a mean of 280 days later to assess long-term retention.

RESULTS: Under an intention-to-treat analysis, students who received the spaced education e-mails demonstrated significant, topic-specific increases in pre-test scores ($P < .001$ and $P = .03$ for PC/PSA and BPH/ED, respectively). Spaced education improved long-term retention of PC/PSA ($P = .04$) but not of BPH/ED ($P = .60$).

CONCLUSIONS: Spaced education delivered prospectively can generate significant, topic-specific learning.

Long-term retention of knowledge is a critical precondition for that knowledge to result in substantive improvements in clinicians' behavior. Unfortunately, knowledge learned by trainees is often quickly forgotten.[1] Ebbinghaus in the late 19th century demonstrated that forgetting is a natural psychological phenomenon that should be expected.[2,3] In a recent randomized trial of web-based teaching in urology to medical students, we documented a substantial decay in urologic knowledge 5 months after the completion of the online program.[4] In fact, if the forgetting curve (plot of memory retention over time) from this trial is extrapolated to 12 months, it appears that little-to-no urologic knowledge is retained. Such forgetting raises the important question as to whether the educational process itself might be tailored to improve students' retention of the curricular material.

"Spaced education" refers to online educational programs that are structured to take advantage of the pedagogical benefits of the "spacing effect."[5,6] The spacing effect is the psychological principle that educational encounters that

This study was supported by the Pellegrino Foundations, the Research Career Development Award Program of the Veterans Affairs Health Services Research & Development Service, and the American Urological Association Foundation (Linthicum, MD), and Astellas Pharma US, Inc.

The views expressed in this article are those of the authors and do not necessarily reflect the position and policy of the United States Federal Government or the Department of Veterans Affairs. No official endorsement should be inferred.

* Corresponding author. Tel.: +1-774-286-9230; fax: +1-857-364-6561.

E-mail address: price.kerfoot@gmail.com

Manuscript received June 13, 2007; revised manuscript October 16, 2007

0002-9610/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.amjsurg.2007.10.026

are spaced and repeated over time (spaced distribution) result in more efficient learning and improved learning retention, compared to massed distribution of the educational encounters (bolus education).[5–7] In the 2004–5 academic year, we conducted a randomized trial which demonstrated that weekly e-mails containing clinical scenarios and questions sent to students upon the completion of their surgery clerkship could significantly improve students' long-term retention of urologic knowledge.[8] The question remained, though, whether educational material presented prospectively using the spaced education methodology could generate significant, topic-specific learning in urology, not just improve its retention. Herein, we report the results of a multi-institutional randomized trial that investigated this question. In addition, we investigated whether spaced education could be utilized prospectively (1) to prime students to achieve greater urologic learning during their surgery clerkships, and (2) to improve students' abilities to self-assess their own levels of urologic knowledge.

Methods

Study participants

In July 2005, all 330 third-year students at Harvard Medical School (HMS) and Boston University School of Medicine (BUSM) in the 2005–6 academic year (176 and 154 students, respectively) were invited via e-mail to participate in the study. Participation was voluntary. There were no exclusion criteria. Students were given a $20 bookstore gift certificate upon completion of the final (delayed) test. Faculty members at both medical schools were blinded to student participation and to which cohort the students were randomized. Institutional review board approval was obtained for this protocol.

Standard online education program in urology

The spaced education intervention described here was delivered as an adjunct to a core online education program in urology that all third-year students at each medical school were required to complete during their 3-month surgery clerkships. This standard online education program was conducted over the duration of their 1- to 2-week surgical-subspecialty elective and consisted of the following: (1) a validated 28-item multiple-choice question pre-test focusing on the core topics of prostate cancer (PC), screening with prostate-specific antigen (PSA), benign prostatic hyperplasia (BPH), and erectile dysfunction (ED); (2) 8 validated web-based teaching cases on these 4 topics; and (3) the same 28 multiple-choice questions as a post-test. The validity and efficacy of these web-based teaching program have previously been established.[4] We used a previously validated 28-item multiple choice test on the 4 core urology topics that has demonstrated a Cronbach alpha reliability of .76.[9] To optimize the comparability of the test results over time, the identical test was used through-out the study. On these tests, students were asked to report their utilization of the educational materials. The test was administered online via the HMS MyCourses web-based course management system. Students' participation in the program and scores on these tests were reported to the clerkship directors.

Development of the spaced education items

The content of the spaced education items and test questions was based on a validated curriculum in clinical urology and specifically focused on the 4 core urology topics outlined above.[10] Each spaced education item contained an evaluative component (a multiple choice question) and an educational component (the correct answer, a take-home message, and explanations of the correct and incorrect answers). An example of a spaced education item is provided in the Appendix. Ten spaced education items were selected for each topic area from a set of previously validated items with demonstrated educational efficacy.[8]

Study design and organization

This multi-institutional randomized controlled trial was conducted from July 2005 to September 2006. Students were stratified by medical school and surgery clerkship date and then were block randomized to 1 of 2 cohorts: cohort A students received spaced education e-mails on the topics of PC/PSA, and cohort B students received spaced education e-mails on the topics of BPH/ED.

The spaced education e-mails were sent to students starting in July 2005. To take advantage of the educational merits of the spacing effect, the educational material was distributed in 3 cycles or repetitions (Figure 1). The temporal relation of the cycles (including their overlap) was established to create expanding time intervals between presentations of a given clinical scenario. This structure has been shown to promote greater retention of learning.[11,12] For example, the first clinical scenario in cycle 1 was presented in week 1, presented again in week 6 (as a 5-week

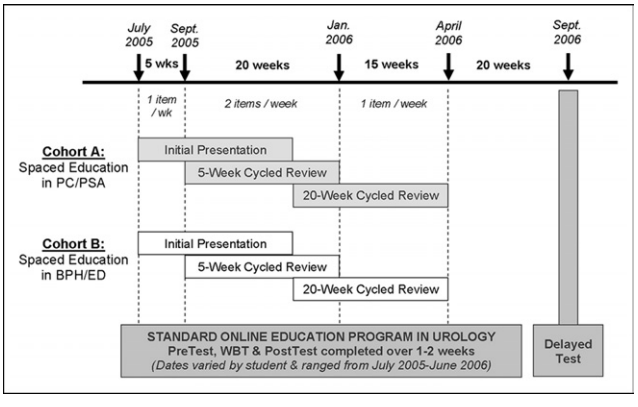

Figure 1 Structure of the randomized controlled trial.

cycled review), and presented for a final time in week 21 (as a 20-week cycled review). It has been our experience from other trials that the repetition is not considered burdensome, but rather is seen by the participants as a means to test and reinforce their prior learning. The entire spaced education program ran 40 weeks: students were sent 1 e-mail per week during weeks 1–5 and 26–40 and 2 e-mails per week during weeks 6–25.

Students were sent the spaced education e-mails up to and during their surgery clerkship (durations ranged from 0–40 weeks, depending on when a student's surgery clerkship was scheduled). When a student completed the core urology education program, he or she would no longer be sent the spaced education e-mails. The study was structured in this manner to increase the average duration between the termination of the spaced education and the completion of the delayed test in order to better assess the long-term retention of the material. In September 2006, the 28-item delayed test on all four core urology topics was administered online to the students. To determine students' abilities to self-assess their own urologic knowledge, students were asked the following question on the delayed test: "Please give a percentile estimate of your performance on the PC/PSA questions (or BPH/ED questions) relative to your classmates."

Outcomes and measures

The primary outcome measure was the acquisition of learning over time from the spaced education items as measured by scores on the pre-test. Secondary outcome measures included (1) the retention of learning from the spaced education items as measured by scores on the delayed test; (2) the priming of learning from the core urology education program by the spaced education items; and (3) the ability of students to accurately self-assess their performance on the delayed test.

Statistical analyses

The data from all students who completed the pre-test were included in an intention-to treat analysis. Pre-test data were carried forward, if needed, to impute any missing post-test and/or delayed test data; this conservative approach fixed all interval gains in knowledge at 0 for those students who did not complete the post-test- or delayed-test. Topic-specific test scores were normalized to a percentage scale, with the minimum score of 0% and a maximum of 100%. Two-tailed Student t tests were used to test the statistical significance of changes in knowledge. Dose-response analyses were performed with Pearson correlation. Given the nonparametric nature of the utilization characteristics, statistical comparisons were performed with the Mann-Whitney U test. Intervention effect sizes for learning were measured by means of Cohen's d, which was calculated by dividing mean scores or score increases by pooled standard deviations.[13] Cohen's d expresses the difference between the means in terms of standard deviation units, with .2 generally considered as a small effect, .5 as a moderate effect, and .8 as a large effect.[14] The accuracy of students' self-assessment of their performance on delayed test was analyzed by t test comparison of the absolute values of the difference between each student's estimated and actual percentile ranking relative to their peers. Statistical calculations were performed with SPSS for Windows 13.0 (SPSS, Inc, Chicago, IL).

Results

One hundred fifteen students enrolled in the trial. The baseline characteristics of the randomized students were similar between cohorts (Table 1), although a small but statistically significant gender difference was present ($P = .03$). The pre-test, post-test, and delayed test were completed by 89% (102/115), 86% (99/115), and 69% (79/115) of participants, respectively (Figure 2). Attrition was similar between cohorts. Students in each cohort reported reading a similar percentage of the spaced education e-mails ($P = .22$). Cohort cross-over reported by students was not significant (data not shown). Students in cohorts A and B reported completing 94% and 97% of the WBT teaching modules, respectively ($P = .40$).

Under an intention-to-treat analysis, students who received the spaced education e-mails demonstrated a significant and topic-specific increase in their pre-test scores compared to controls ($P < .001$ and $P = .03$ for PC/PSA and BPH/ED, respectively, Figure 3), corresponding to Cohen effect sizes of .83 and .44. A significant, topic-specific dose-response was present. In cohort A, the duration of the PC/PSA spaced education e-mails correlated significantly with PC/PSA pre-test scores (Pearson $r = .33$, $P = .02$) but did not correlate with BPH/ED pre-test scores ($r = .05$, $P = .71$). The reverse was true in cohort B: the duration of the BPH/ED spaced education e-mails correlated significantly with BPH/ED pre-test scores ($r = .29$, $P = .04$) but did not correlate with PC/PSA pre-test scores ($r = .10$, $P = .50$).

After completion of the WBT modules, there were no significant differences in topic-specific post-test scores between cohorts ($P = .40$ and $P = .06$ for PC/PSA and BPH/ED, respectively, Figure 3).

On average, the delayed test was completed by students 40 weeks (280 days; range 77–434) after submission of the post-test, with no significant differences in duration between cohorts. Even though PC/PSA knowledge was similar between cohorts at the time of the post-test, students who received spaced education in PC/PSA preferentially retained this topic-specific knowledge, compared to students who did not receive the spaced education on this topic ($P = .04$; effect size .41, Figure 3). The spaced education program on BPH/ED did not generate an improvement in retention ($P = .60$).

On their delayed test, students were asked to rate their interest in the urology topics on a 5-point Likert-type scale

Table 1 Characteristics of the 115 randomized students

	Cohort A	Cohort B
	Spaced education on PC/PSA	Spaced education on BPH/ED
Participants included in randomization	58	57
Medical school		
Boston University School of Medicine	25 (43%)	25 (44%)
Harvard Medical School	33 (57%)	32 (56%)
Clerkship dates		
July–September 2005	20 (34%)	21 (37%)
October–December 2005	15 (26%)	14 (25%)
January–March 2006	13 (22%)	12 (21%)
April–June 2006	10 (17%)	10 (18%)
Site of clerkship		
Beth Israel Deaconess Med.	12 (21%)	12 (21%)
Boston Med. Center–East Newton	15 (26%)	11 (19%)
Boston Med. Center–Menino	3 (5%)	5 (9%)
Brigham & Women's Hospital	7 (12%)	10 (18%)
Cape Cod Hospital	2 (3%)	2 (4%)
Mass. General Hospital	14 (24%)	10 (18%)
Quincy Hospital	0 (0%)	5 (9%)
VA Boston Healthcare System	5 (9%)	2 (4%)
Gender		
Female	22 (38%)	34 (60%)
Male	36 (62%)	23 (40%)
Degree program		
M.D.	43 (74%)	44 (77%)
M.D./Ph.D.	8 (14%)	7 (12%)
No response	7 (12%)	6 (11%)
Completed urology rotation	22 (38%)	17 (30%)
Age (y)	25.8 (SD 2.5)	25.4 (SD 2.2)
Self-assessment of prior "knowledge base in urology"*	2.1 (SD .8)	2.0 (SD .8)

Percentages may not total 100% due to rounding.
*5-point scale: 1 = poor, 5 = excellent.

(1 = not at all interested, 5 = very interested). Overall, students rated the topics of PC/PSA as more interesting than BPH/ED ($P < .001$), but there were no topic-specific differences in interest between cohorts. Students also reported that the maximum total number of spaced education e-mails that they would want to receive each week would be a mean 5.1 (median 4.0) if multiple specialties besides urology started conduct spaced education e-mail programs.

When asked to assess their own performance on the PC/PSA questions on the delayed test relative to their peers, students who received the spaced education in this topic area were significantly more accurate in their self-assessment ($P = .02$, Figure 4). Similarly, students who received the spaced education on BPH/ED also could more accurately assess their performance on these items relative to their peers, but this difference did not reach statistical significance ($P = .06$).

Comments

The results of this randomized trial demonstrate that spaced education delivered prospectively can generate significant topic-specific learning in urology. Spaced education also improved the long-term retention of PC/PSA knowledge, but it did not improve retention of BPH/ED knowledge. These results likely underestimate the full educational value of the spaced education program since more than 50% of participating students received less than 2 full cycles of the educational material. Our findings confirm the results from a parallel study involving 537 urology residents in the United States and Canada.[15] In this randomized trial, those residents who received spaced education demonstrated significantly greater knowledge acquisition and retention than those in the bolus cohort. The pedagogical merits of the spacing effect extend well beyond online education or the learning of clinical knowledge. For example, a recent randomized trial compared massed (1 day) versus distributed (weekly) training of surgical residents in microvascular anastomosis skills. Those residents in the weekly training sessions had significantly greater skill retention and were better able to transfer these skills to a live, anesthetized rat model.[16] A distinct neurophysiological basis for the spacing effect has been identified. A recent study demonstrated that spaced learning by rats improves neuronal longevity in the hippocampus and that the strength of the rats' memories correlates with the number of new cells in this region of their brains.[17]

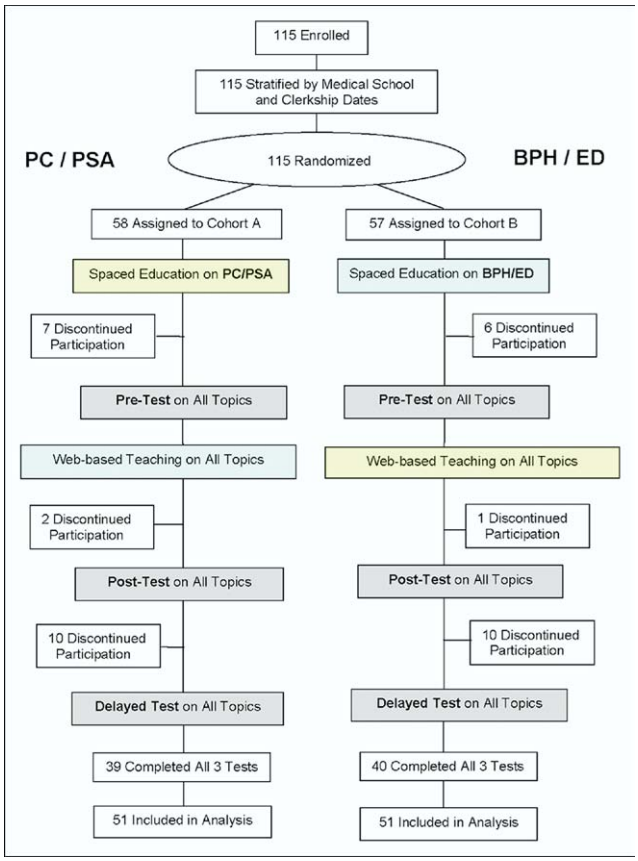

Figure 2 Consolidated Standards of Reporting Trials (CONSORT) flow chart of randomized controlled trial.

It is not clear why the spaced education program in BPH/ED to urology students did not generate improvements in retention, as did the PC/PSA program. This may be due to students' relatively limited interest in BPH/ED. The core concepts and management principles of BPH and ED that we teach to the students are less well-defined than those of PC and PSA. As such, these topics may be more difficult to teach using the spaced education question–answer format. The spaced education also did not prime the students who received it to learn more from their core urology education program of web-based teaching modules. This finding may be an artifact due to a ceiling effect from our test instrument, but more likely, the educational impact of the bolus web-based teaching modules overwhelmed the learning gains from spaced education in the short-term.

In spite of the prospective spaced education program, students in both cohorts demonstrated a substantial decline in their urology knowledge in between the post-test and delayed test in both topic areas. Thus, while prospective spaced education can improve learning and retention, it does not appear to be enough in-of-itself to shift this urology learning into long-term memory. This finding argues that structured reinforcement of the important curricular content is required throughout the medical school curriculum (and beyond). Our prior work has demonstrated that such a maintenance program of spaced education is acceptable and effective.[8]

Of note, spaced education did improve students' long-term abilities to accurately self-assess their own knowledge proficiency relative to their peers, although statistical significance was achieved for this outcome measure in only 1 of the 2 cohorts. This is an important finding since physician education after the end of formal training is largely based on physicians' self-directed educational activities. It is thus critical that physicians be able to accurately self-assess their true knowledge deficiencies and thus their true learning needs. Unfortunately, there is a great deal of psychological evidence that humans (which presumably includes physicians) are quite limited in their abilities to accurately self-assess their performance.[18–20] Spaced education holds promise as a method to improve physicians' and physician trainees' skills of self-assessment. The key may be the immediate and direct feedback provided by the spaced education e-mails. While the cycled reviews of the material reinforce the educational material, they may also be reinforcing the learners' more accurate perceptions of their performance. One potential limitation of our analysis is that

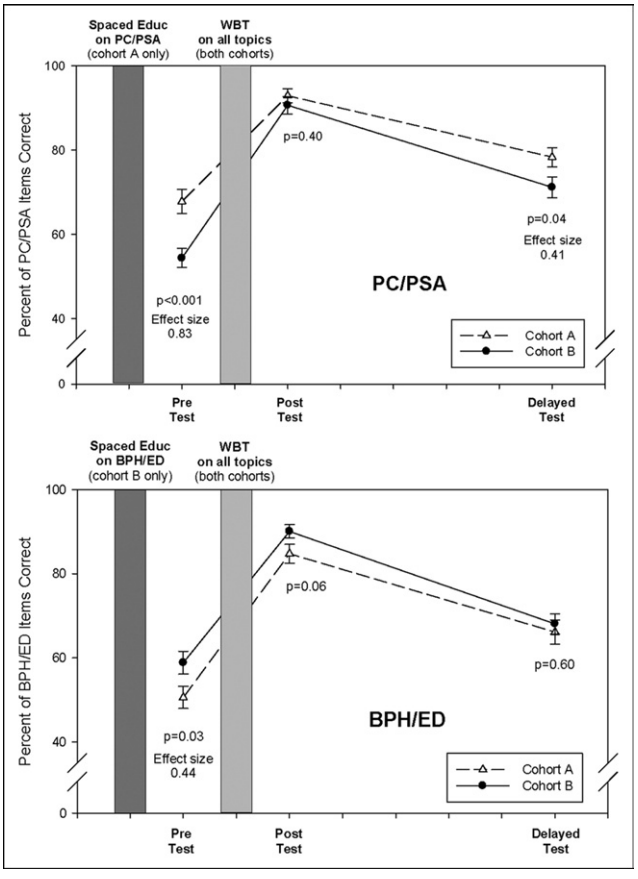

Figure 3 Topic-specific test scores by cohort. The identical 28-item test covering all 4 core urology topics was administered to students 3 times over the duration of the study. Students in cohort A (white triangles) received the spaced education e-mails on PC/PSA and students in cohort B (black circles) received the spaced education e-mails on BPH/ED. The pre-test and post-test were completed 1–2 weeks apart; the delayed test was completed a mean of 280 days (40 weeks) after submission of the post-test.

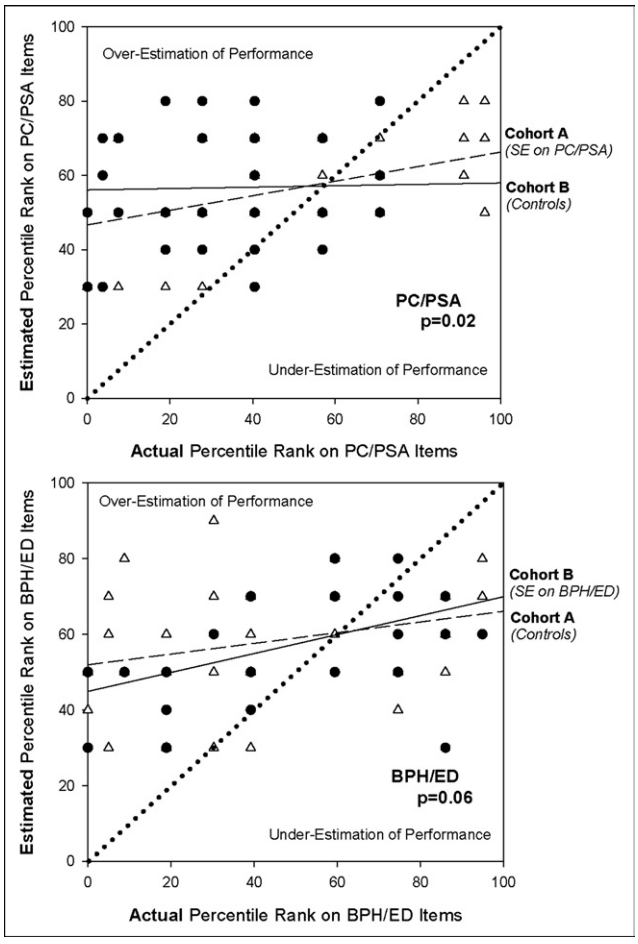

Figure 4 Accuracy of student self-assessment of performance on delayed test. In each plot, the dotted diagonal line from the origin represents accurate self-assessment of performance, the points in the upper left represent students who have overestimated their performance relative to their peers, and the points in the lower right represent students who have underestimated their performance relative to their peers.

students were asked to assess their percentile performance relative to their peers, rather than against an absolute standard of competency. There were 2 reasons that we elected for this approach. First, validated standards of competency are extremely difficult to establish and currently do not exist for medical student education in urology. Second, percentile-based comparisons of physician performance are increasingly more common in the United States as efforts are made to improve physician performance overall.[21] For example, in an attempt to lower death rates from coronary artery bypass grafting, health officials in Massachusetts now post online the surgeon-specific 30-day mortality rates for this procedure, as well as the names of the surgeons whose adjusted death rates are significantly higher or lower than their peers.[22]

This study has several limitations, including the moderate response rate of students on the delayed test and the focused nature of the educational content. In addition, the structure of our study does not allow us to determine the extent to which the improvements in learning and retention were due to the spaced education methodology itself or due to the additional exposure to the educational content provided by the spaced education program. Strengths of the study include its randomized controlled design, the inclusion of students from more than 1 medical school, and its focus on long-term retention of learning by the students.

In summary, spaced education delivered prospectively can generate significant improvements in topic-specific knowledge and improve students' abilities to self-assess their own levels of urologic knowledge. Further research is needed to determine how best to solidify this learning into longer-term memory.

Acknowledgment

The authors thank Dr. Daniel D. Federman for his support of the project and his critical revision of the manuscript.

References

1. McKenna SP, Glendon AI. Occupational first aid training: decay in cardiopulmonary resuscitation (CPR) skills. J Occup Psychol 1985;58: 109–17.
2. Baddeley A. *Your Memory: A User's Guide*. Buffalo, NY: Firefly Books; 2004.
3. Ebbinghaus H. Memory: *A Contribution to Experimental Psychology*. New York, NY: Dover Publications, Inc., 1964.
4. Kerfoot BP, Baker H, Jackson TL, et al. A multi-institutional randomized controlled trial of adjuvant web-based teaching to medical students. Acad Med 2006;81:224–30.
5. Glenberg AM, Lehmann TS. Spacing repetitions over 1 week. Memory Cognition 1980;8:528–38.
6. Toppino TC, Kasserman JE, Mracek WA. The effect of spacing repetitions on the recognition memory of young children and adults. J Exp Child Psychol 1991;51:123–38.
7. Landauer TK, Bjork RA. Optimum rehearsal patterns and name learning. In: Gruneberg MM, Morris PE, Sykes RN, editors. *Practical Aspects of Memory*. New York, NY: Academic Press; 1978:625–32.
8. Kerfoot BP, DeWolf WC, Masser BA, et al. Spaced education improves the retention of clinical knowledge by medical students: a randomised controlled trial. Med Educ 2007;41:23–31.
9. Kerfoot BP, Baker H, Volkan K, et al. Development of validated instrument to measure medical student learning in clinical urology: a step toward evidence based education. J Urol 2004;172:282–5.
10. Kerfoot BP, Baker H, Volkan K, et al. Development and initial evaluation of a novel urology curriculum for medical students. J Urol 2004;172:278–81.
11. Morris PE, Fritz CO. The improved name game: better use of expanding retrieval practice. Memory 2002;10:259–66.
12. Bjork RA. Retrieval practice and the maintenance of knowledge. In: Gruneberg MM, Morris PE, Sykes RN, editors. *Practical Aspects of Memory: Current Research and Issues*. New York, NY: Wiley; 1988: 396–401.
13. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. Ed. 2. Hillsdale, NJ: Erlbaum; 1988.
14. Maxwell SE, Delaney HD. *Designing Experiments and Analyzing Data: A Model Comparison Approach*. Belmont, CA: Wadsworth, 1990.

15. Kerfoot BP, Baker HE, Koch MO, et al. Randomized, controlled trial of spaced education to urology residents in the United States and Canada. J Urol 2007;177:1481–7.
16. Moulton CA, Dubrowski A, Macrae H, et al. Teaching surgical skills: what kind of practice makes perfect? A randomized, controlled trial. Ann Surg 2006;244:400–9.
17. Sisti HM, Glass AL, Shors TJ. Neurogenesis and the spacing effect: learning over time enhances memory and the survival of new neurons. Learn Mem 2007;14:368–75.
18. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA 2006;296:1094–102.
19. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol 1999;77:1121–34.
20. Dunning D. *Self-Insight: Roadblocks and Detours on the Path to Knowing Thyself*. New York, NY: Psychology Press; 2005.
21. Gwande A. *Better: A Surgeon's Notes on Performance*. New York, NY: Metropolitan Books; 2007.
22. Adult coronary artery bypass grafting surgery in adults in Massachusetts: January 1, 2002 - December 21, 2004; Surgeon-specific standardized mortality incidence rates. http://www.massdac.org/reports/SurgeonSpecificRates2002to2004.pdf; accessed May 15, 2007.

Appendix: Example of spaced education e-mail

Spaced education in urology: benign prostatic hyperplasia

Your healthy 63-year old patient, Mr. Gore, completes an International Prostate Symptom Score (IPSS) questionnaire which documents his moderate urinary symptoms (weak urinary stream, urgency and nocturia × 4). His urinalysis is normal, his post-void residual is low (15 cc), and his prostate is mildly enlarged on examination. What is the next step in the management of urinary symptoms?

(A) referral to a urologist for uroflowmetry
(B) referral to a urologist for urodynamics
(C) referral to a urologist for a TURP (transurethral resection of the prostate)
(D) start empiric therapy with an alpha-blocker
(E) start empiric therapy with a 5-alpha reductase inhibitor

Scroll down for the answer . . . .

Correct answer: D
Start empiric therapy with an alpha-blocker

Take-Home Message:

Primary care physicians should feel free to start an empiric trial of alpha-blocker therapy (hytrin, flomax, etc.) for their patients with lower urinary tract symptoms. Referral to a urologist is indicated if the alpha-blocker is not effective, if the symptoms worsen, or if the patient cannot tolerate the medication (usually due to dizziness).

Explanation of Incorrect Answers:

(A) referral to a urologist for uroflowmetry: while this test will document the patient's urinary flow rate (<10cc/sec suggestive of obstruction), an alpha-blocker can be started in the absence of this information.
(B) referral to a urologist for urodynamics: while this test will document the patient's bladder mechanics, it is usually best to give a trial of an alpha-blocker prior to this procedure.
(C) referral to a urologist for a TURP (transurethral resection of the prostate): A TURP can be very effective to relieve obstruction caused by the prostate, but the procedure may not be necessary if the patient responds well to an alpha-blocker.
(E) start empiric therapy with a 5-alpha reductase inhibitor: recent data suggest that an alpha blocker plus a 5-alpha reductase inhibitor may be more effective than the alpha blocker alone. Due to the increased side effect profile of this combination therapy, most doctors start with an empiric trial of the alpha-blocker alone.

Reference: McConnell, et al. N Engl J Med 349:2387–98, 2003.