**Title Page**

**Title of Project:** Spaced Education to Optimize Prostate Cancer Screening

**Principal Investigator and Team Members:**
B. Price Kerfoot, MD, EdM – Principal Investigator
Paul R. Conlin, MD – Primary Mentor

**Organization:** Harvard University (Harvard Medical School [HMS])

**Inclusive Dates of Project:** 9/30/2007 to 9/29/2011

**Federal Project Officers.**
(1) Grants Management Specialist: Al Deal
Email: adeal@ahrq.gov Phone: (301) 427-1806 Fax: (301) 427-1462
(2) Program Official: Kay Anderson
Email: kay.anderson@ahrq.hhs.gov Phone: (301) 427-1555 Fax: (301) 427-1562

**Grant Award Number:** 1K08HS017147-01

## (1) Structured Abstract

**Purpose:** To investigate whether an email-based intervention termed 'Spaced Education' (SE) could reduce clinicians' inappropriate screening for prostate cancer.

**Scope:** Prostate cancer screening with prostate-specific antigen (PSA) is frequently performed counter to clinical practice guidelines.

**Methods:** Ninety-five primary care clinicians in eight Veterans Affairs medical centers enrolled. Participants were randomized into two cohorts: SE clinicians received four isomorphic cycles of nine emails over 36 weeks (0-2 emails per week), while control clinicians received no intervention. Each email presented a clinical scenario and asked whether it was appropriate to obtain a PSA test. Participants received immediate feedback after submitting their answers. Patients with PSA testing for nonscreening reasons were excluded using a validated protocol. Inappropriate testing was defined as use of PSA for prostate cancer screening in patients >76 or <40 years old. Logistic regression with adjustment for patient clustering by clinician was performed.

**Results:**
During the 36-week intervention period (weeks 1-36), clinicians receiving SE emails ordered significantly fewer inappropriate PSA screening tests than did control clinicians (10.5% vs 14.2%; p=0.041). Over the 72-week period following the intervention (weeks 37-108), SE clinicians continued to order fewer inappropriate tests compared with controls (7.8% vs 13.1%; p=0.011), representing a 40% relative reduction in inappropriate screening. The screening differences between cohorts did not erode significantly over this follow-up period (p=0.63 for cohort-time interaction). SE durably improves the prostate-cancer screening behaviors of clinicians and represents a promising new methodology to improve patient care across healthcare systems.

**Key Words:**
Educational Technology; Clinical Practice Guideline; Prostate Cancer; Prostate-Specific Antigen

**(2) Purpose (Objectives of Study).**

Specific Aim 1: To develop a validated test instrument to reliably measure providers' knowledge of evidence-based PSA screening guidelines.

Specific Aim 2: To develop a validated spaced education intervention to instruct primary care providers on how to utilize PSA screening in accordance with evidence-based CPGs.

Specific Aim 3: To conduct a multi-institutional, randomized controlled trial to evaluate the impact of this spaced education intervention on providers' knowledge of PSA screening guidelines
(secondary outcome measure).

Specific Aim 4: In this same randomized trial, to evaluate the impact of the spaced education intervention on PSA utilization patterns by primary care providers (primary outcome measure).

**(3) Scope (Background, Context, Incidence, Prevalence).**

Much of the value of education depends on the life span of acquired knowledge. Even so, memory research has contributed little to educational practice over the past century, largely due to its experimental focus on short-term retention of memory content. Though a few researchers have focused on long-term retention of knowledge, the insights derived through their work have not been effectively incorporated into educational practice.

We have developed a novel method of online education (termed 'spaced education') that improves long-term retention of learning by harnessing two core memory research findings: the spacing effect and the testing effect. The 'spacing effect' refers to the finding that information that is presented and repeated over spaced intervals of time is learned and retained more efficiently than information presented at a single time point. The spacing effect appears to have a distinct neurophysiological basis. 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. The 'testing effect' refers to the research finding that the process of testing does not merely assess the knowledge levels of individuals. Rather, it alters the learning process itself so that new knowledge is retained more effectively.

In randomized controlled trials, we have shown that the spaced education methodology improves knowledge acquisition and boosts learning retention. We currently deliver spaced education via periodic emails that contain case scenarios and multiple-choice questions. Upon submitting answers to each question online, learners receive immediate feedback and educational material. The questions (or topics) are then repeated over spaced intervals of time to harness the pedagogical benefits of the spacing effect.

In medical education, the most important outcome measure is not the generation of new knowledge but rather the generation of new knowledge that is effectively translated into improved practice patterns and patient outcomes. To assess whether spaced education could durably improve clinical behavior, we selected screening for prostate cancer with prostate-specific antigen (PSA) as our experimental system. PSA is serine-protease released by prostatic epithelium for which serum levels correlate with prostate cancer risk. There is great controversy about if and when PSA screening should be performed, because there are limited data showing that PSA screening reduces overall mortality. On the other hand, there is little to no controversy as to when PSA should not be used for screening. None of the major clinical practice guidelines (CPGs) recommend that PSA screening be routinely performed in asymptomatic men younger than 40 years of age, greater than 75 years of age, or with less than a 10-year life expectancy. Even so, PSA screening is frequently performed counter to CPGs. A nationwide study of Medicare and Veterans Affairs (VA) records found that over 50% of men aged 75+ underwent PSA screening in 2003. In addition, a provider-level analysis demonstrated that 19% of all PSA screening tests ordered by clinicians from 1997-2004 were performed counter to guidelines.

**(4) Methods (Study Design, Data Sources/Collection, Interventions, Measures, Limitations).**

In our randomized trial, we investigated whether spaced education could be utilized to durably reduce inappropriate PSA screening by clinicians. Primary care providers (physicians, nurse practitioners, and physician assistants) from eight Veterans Affairs (VA) hospitals in the New England region were recruited via email to participate. The spaced education intervention consisted of 36 items, each of which consisted of an evaluative component (a multiple-choice question based on a clinical scenario) and an educational component (the answer and explanation). The content of the items was developed from published CPGs. Nine spaced education items were constructed to address whether PSA screening should or should not be performed in variety of clinical contexts. Three additional sets of nine items (for cycles 2-4) were constructed to be identical in focus to the original set, except for minor alterations to the clinical contexts to reduce the sense of repetition for study participants. We developed and pilot-tested multiple-choice test items on this same content. Fourteen questions were selected for inclusion in the test based on item difficulty, point-biserial correlation, and Kuder-Richardson 20 score. The 36 spaced education items (four cycles of nine items) and 14 test questions were independently content validated by three physicians (n=2 internists and n=1 urologist).

The spaced education items were delivered to PCPs at designated time intervals via an automated email delivery system. The email presented the clinical scenario and question (evaluative component). Upon clicking a hyperlink in the email, a web page opened that allowed the PCP to submit an answer to the question. The answer was downloaded to a central server, and PCPs were immediately presented with a web page displaying the correct answer to the question and an explanation of the curricular learning point (the educational component).

This multi-institutional, randomized controlled trial was conducted from January 2007 through February 2009. PCPs were stratified by hospital and block randomized into two cohorts: the spaced education cohort received four cycles of nine emails over a 36-week period (0-2 emails per week), while those in the control cohort received no intervention (and represented the standard of education in the VA system). The timing of four spaced education cycles was established to take advantage of the educational benefits of the spacing effect. All the material presented in cycle 1 was presented again (with minor variations in clinical context) as 3-week, 6-week, and 18-24 week cycled reviews. For example, isomorphic variations of the spaced education item presented in week 1 (in cycle 1) were re-sent to PCPs in week 4 (in cycle 2), week 10 (in cycle 3), and week 19 (in cycle 4). This expanding pattern of spacing intervals has been demonstrated to improve retention of learning compared with fixed intervals. The time intervals between spaced education cycles were established based on psychology research findings to optimize long-term retention of learning. The 14-item test was administered to participants at enrollment (test 1) and at weeks 18 (test 2) and 36 (test 3). At week 36, participants also completed a short survey that asked the amount of time required to answer each spaced education item.

The primary outcome measure was the difference in the percentage of inappropriate PSA screening performed by PCPs in the spaced education and control cohorts. Based on published clinical guidelines and reports, inappropriate PSA utilization was defined as the use of PSA for prostate cancer screening in patients older than age 76, younger than age 40, or with an estimated life expectancy of fewer than 10 years. For the average man in the United States, an estimated life expectancy of 10 years is reached at age 76. No adjustments were made to account for patient comorbidities that might alter life expectancy. We dichotomized appropriateness of screening based on the age of the patient at the time of the screening test. Data on PSA testing and patient characteristics were extracted from VA electronic databases for the male patients whose PSA levels were tested by participating clinicians during the trial period. Patient age for determining appropriateness of PSA screening was defined at the time

of each individual PSA test. Patients with PSA testing for reasons other than screening were excluded from the database, following a previously validated protocol. An intention-to-treat analysis of PSA screening outcomes was performed. The percentages of inappropriate PSA screening tests were calculated by dividing the number of inappropriate PSA screening tests ordered by a cohort by the total number of PSA screening tests ordered by that cohort. Logistic regression with adjustment for clustering by provider was performed using a generalized estimating equation approach. Secondary outcome measures included (1) the change in test scores between cohorts measured over time and (2) the change in performance on the spaced education items measured over time.

**(5) Results (Principal Findings, Outcomes, Discussion, Conclusions, Significance, Implications).** Ninety-five of 260 PCPs enrolled in the trial. Participants' baseline demographic characteristics and attrition over the course of the trial were similar between randomized cohorts. In the spaced education cohort, the four cycles of nine emails were completed by 84%, 88%, 94% and 92% of clinicians, respectively. Spaced education clinicians reported spending a median 2.5 minutes (IQR 1.5-5.0) to complete each item. The estimated duration to complete the entire program of 36 emails was a median 90 minutes (estimated IQR 54-180). The percentage of spaced education items answered correctly rose from 72.0% (SD 20.9) in cycle 1 to 85.3% (SD 16.6) in cycle 2, to 88.3% (17.2) in cycle 3, and to 90.1% (SD 15.5) in cycle 4. All three multiple-choice tests were completed by 92% and 85% of providers in the spaced education and control cohorts, respectively (p=0.23). Average Cronbach alpha reliability (internal consistency) of the test instrument was 0.71. Test 1 scores (week 0) were similar between cohorts (mean 72%, p=0.89). Mean test scores in the spaced education cohort were 96% (SD 8) and 95% (SD 8) for tests 2 (week 18) and 3 (week 36), respectively. Corresponding test scores in the control cohort were 73% (SD 20) and 76% (SD 17) ($p<0.001$ for both cross-cohort comparisons). The cross-cohort differences represent Cohen effect sizes of 1.2 and 1.1 for test 2 and 3, respectively.

Over the 36-week intervention period, clinicians in the spaced education cohort ordered fewer inappropriate PSA screening tests compared with control clinicians (762 vs 1145, respectively). Spaced education clinicians also ordered fewer PSA screening tests overall compared with clinicians in the control cohort (7244 vs 8173, respectively). In logistic regression models, the percentage of inappropriate PSA screening was significantly reduced among spaced education clinicians compared with controls (10.5% vs 14.2%, respectively; p=0.041). This between-cohort difference in the percentage of inappropriate PSA screening increased significantly over intervention period (p=0.019 for interaction between cohort and time). The impact of the intervention was unaffected by clinicians' age, gender, or provider type.

In the 72-week follow-up period, clinicians in the spaced education cohort continued to order fewer inappropriate PSA screening tests (1028 vs 1906 by control clinicians) and fewer PSA screening tests overall (13,089 vs 14,488 by controls). Over this time period, the percentage of inappropriate PSA screening continued to be significantly lower among spaced education clinicians compared with controls (7.8% vs 13.1%, respectively; p=0.011), representing a 40% relative reduction in inappropriate screening. The screening differences between cohorts did not erode significantly over the 72-week follow-up period (p=0.63 for interaction between cohort and time).

Our results show that online spaced education can significantly reduce clinicians' inappropriate screening for prostate cancer and align their clinical practice patterns more closely with CPG standards In addition, our study is the first to show that online education can produce demonstrable improvements in clinical practice that persist for more than 1 year after the intervention. The fact that this modest intervention (36 interactive emails over a 36-week period) generated such substantial and durable results suggests that spaced education is

a potent methodology for continuing education of health professionals. With content tailored to meet specific needs, online spaced education is the type of intervention that can readily be deployed across healthcare systems to improve the quality of patient care. Because the spaced education methodology is content neutral, it also has the potential to improve long-term learning across a broad range of nonmedical topics and across all levels of education (from primary school to university). Additional research is needed to establish spaced education's efficacy in these domains.

By facilitating the interactive answering of questions and allowing for the automated delivery and spacing of educational items, the online environment appears to be particularly well suited for taking advantage of the spacing and testing effects. We now utilize an adaptive algorithm that improves learning efficiency by customizing the spacing and content of the educational items for each learner. Even so, the benefits of the spacing and testing effects are not limited to online education. Given their robust impact on long-term retention of learning, it is unclear why they have not been systematically incorporated into classroom learning. One exception has been the construction of 'spiral curricula,' in which content material is iteratively revisited over a longitudinal education program. Another has been the use of an 'exam a day' in college courses to improve retention of learning. Though some suggest that the spacing and testing effects emphasize fact-based learning, our results demonstrate that an educational methodology based on these effects can durably impact much higher levels of learning.

Our study raises some important but unanswered questions. First, it is not clear how spaced education produced substantial improvements in PSA screening behaviors despite clinician's high baseline knowledge of CPGs (mean baseline test score 72%). The fact that spaced education was effective suggests that it does more than just improve knowledge. We hypothesize that spaced education may also improve the translation of knowledge into clinical practice and/or reduce providers' clinical inertia in adhering to screening CPGs. Second, the educational benefit of cycles 3 and 4 is not clear, given that clinicians' spaced education performance appeared to plateau after cycle 2. However, over-learning (the repeated presentation of content after mastery has been achieved) can significantly improve long-term retention of learning. It is also possible that, once saturation of learning is reached, the spaced education intervention becomes more of a reminder system than an educational tool. Third, it is not clear how long the improvements in practice patterns will persist after this 72-week follow-up period. Still untested is whether a maintenance program of spaced education with substantially increased spacing intervals may prevent regression back to baseline practices. Finally, the optimal frequency, spacing intervals, content length, and duration for spaced education programs have yet to be established 'Email fatigue' is a potential concern, but spaced education programs that deliver more content more frequently (e.g., daily emails containing two questions) are well accepted by students and physicians.

Limitations of our study include the narrow focus of the spaced education intervention on inappropriate PSA screening and our binary outcomes measure (screened inappropriately or not). There are also multiple strengths to the study, including the novelty of the online educational intervention, the randomized controlled design, the inclusion of multiple provider types, and the study's focus on long-term educational and behavioral outcomes.

**All specific aims were successfully achieved:**

► Specific Aim 1: A set of 20 provisional test items based on PSA screening CPGs was developed, with emphasis on evaluating when PSA screening should not be performed. Each test item was content validated by a panel of three physicians, revised to improve clarity and clinical validity, and then pilot tested by 30 physicians to determine the test items' psychometric properties. Based on the pilot test results, a 14-item test was constructed, with a Cronbach alpha reliability of 0.71.

► Specific Aim 2: Thirty-six clinical scenarios were developed based on the PSA screening CPGs to exemplify the circumstances under which PSA screening should not be performed. These scenarios were content validated by a panel of three physicians, and revisions were made to improve clarity and accuracy.

► Specific Aims 3 & 4: These results are outlined in detail above. In summary, clinicians receiving interactive spaced education (SE) emails over 36 weeks ordered significantly fewer inappropriate PSA screening tests than did control clinicians (762 vs 1146, respectively; $p<0.001$). The percentage of inappropriate PSA screening was significantly reduced among SE clinicians compared with controls (10.5% vs 14.2%; $p=0.041$). Differences in inappropriate PSA screening between cohorts increased significantly over the duration of the trial ($p=0.018$ for time-cohort interaction).

**Significance & Implications:**

Inappropriate screening for prostate cancer among elderly men leads to diagnostic and therapeutic procedures of questionable benefit and incurs substantial psychological and financial costs. If the spaced education program is successful, the improved prostate cancer screening patterns would reduce patient anxiety, reduce patients' loss of work, reduce the number of inappropriate referrals to the urology service, reduce the rate of unnecessary prostate biopsies, and result in a substantial cost saving for the health system as a whole. In addition, the spaced education methodology could serve as a model through which a range of clinical practice guidelines could be implemented more effectively.

**(6) List of Publications:**

Kerfoot BP, Lawler EV, Sokolovskaya G, Gagnon D, Conlin PR. Durable improvements in prostate cancer screening from online spaced education: a randomized controlled trial. Am J Prev Med 2010;39:472-8.