Effect of a Cognitive Aid on Adherence to Perioperative Assessment and Management Guidelines for the Cardiac Evaluation of Noncardiac Surgical Patients


ABSTRACT

Background: The 2007 American College of Cardiologists/American Heart Association Guidelines on Perioperative Cardiac Evaluation and Care for Noncardiac Surgery is the standard for perioperative cardiac evaluation. Recent work has shown that residents and anesthesiologists do not apply these guidelines when tested. This research hypothesized that a decision support tool would improve adherence to this consensus guideline.

Methods: Anesthesiology residents at four training programs participated in an unblinded, prospective, randomized, cross-over trial in which they completed two tests covering clinical scenarios. One quiz was completed from memory and one with the aid of an electronic decision support tool. Performance was evaluated by overall score (% correct), number of incorrect answers with possibly increased cost or risk of care, and the amount of time required to complete the quizzes both with and without the cognitive aid. The primary outcome was the proportion of correct responses attributable to the use of the decision support tool.

Results: All anesthesiology residents at four institutions were recruited and 111 residents participated. Use of the decision support tool resulted in a 25% improvement in adherence to guidelines compared with memory alone ($P < 0.0001$), and participants made 77% fewer incorrect responses that would have resulted in increased costs. Use of the tool was associated with a 3.4-min increase in time to complete the test ($P < 0.001$).

Conclusions: Use of an electronic decision support tool significantly improved adherence to the guidelines as compared with memory alone. The decision support tool also prevented inappropriate management steps possibly associated with increased healthcare costs. (Anesthesiology 2014; 120:1339-53)
Decision Support for Perioperative Assessment

Four anesthesiology training programs participated in this prospective randomized trial with cross-over design which was completed between February and May 2013. After investigational review board approval was granted at all institutions (Charleston, South Carolina; Nashville, Tennessee; Lexington, Kentucky; Chapel Hill, North Carolina), all residents (total 211) were recruited via electronic communication, with 111 (52.6%) self-selecting to consent to participate (fig. 1). As no baseline performance measure was known for the study population, an a priori power analysis was not performed. The recruited sample size \( n = 111 \) provided 80% power to detect very subtle increases (approximately 0.3 SD units) in the primary outcome, the proportion of correct responses attributable to the use of the cognitive aid, assuming two-sided hypothesis testing and an \( \alpha \) level of 0.05. Secondary outcomes, described in the last paragraph of the Materials and Methods section, included the performance of test-takers related to the types of errors and efficiency with which they completed the tests. Multi-site compliance and logistics were regulated in the Department of Anesthesia and Perioperative Medicine at the Medical University of South Carolina (MUSC). Randomization was done using computer-generated number generation and was assigned by MUSC as described below. The programs that participated in the study were MUSC (Charleston, South Carolina), Vanderbilt University Medical Center (VUMC, Nashville, Tennessee), University of North Carolina (Chapel Hill, North Carolina), and University of Kentucky (Lexington, Kentucky).

Demographic information of participants including year in training, sex, and race was collected. Participants were then randomly assigned to groups A to D by a statistician with no personal experience with any of the participants. Group assignment A to D determined the order of quiz- zes and the availability of the DST for each quiz (fig. 1). Differences between group demographics were assessed via chi-square tests.

The DST containing the AHA/ACC evaluation and management guidelines was presented in electronic form on either an Apple iPad or iPhone (Apple, Inc., Cupertino, CA). A subset of the authors (M.D.M. and W.R.H.) designed the logic of the DST based directly on the AHA/ACC guidelines. An iOS programmer created the executable application. Various screenshots of the DST can be seen in appendix 1, which demonstrate how a user would navigate the ACC/AHA guidelines using the DST. Participants had no previous exposure or formal training to the DST before this study. The DST was distributed electronically to participants on the day of testing and the software expired immediately after use, so that participants tested on subsequent days would not have confounding access to the DST before their testing session.

A single, optional, online didactic lecture provided a short introduction to the ACC/AHA perioperative cardiovascular guidelines and an introduction to the DST. The same online training was used at all study sites. This didactic
was intentionally brief (22 min), as its goal was orientation to the use of the DST rather than extensive education and review of the ACC/AHA guidelines. The goal of keeping this brief was to test the true capacity of residents to apply guidelines based on their existing functional knowledge of these published guidelines, rather than testing them immediately after an extensive review, as was used in two previous studies.\(^4\,5\) Orientation was performed a minimum of 2 weeks before implementation of the cognitive aid at all sites to prevent score inflation associated with very recent exposure. Testing was performed with local participation in a standardized (classroom) environment, which included supervision from trained study personnel who administered the test and gave access to the DST according to group assignment. All participants were excused from active clinical duties and asked to silence all pagers except hospital-wide mayday pagers when appropriate. Testing of participants was accomplished within a 10-day period at all sites.

To build on the recent works by Vigoda et al.,\(^4\,5\) MCQ stems were created with a correct answer representing each evaluation path and terminal end point of the ACC/AHA perioperative cardiac evaluation algorithm. Although there is no defined guideline for the design and assessment of MCQs, our process was similar to that previously described.\(^4\,5\) The clinical stems and corresponding MCQ answer options were assessed for construct and content validity by evaluating whether they actually represented an adequate test of knowledge of the ACC/AHA guidelines under consideration, both in the specifics of each question being asked and in the scope of the questions as a whole concerning the information contained in the guidelines. This was done by having a cohort of anesthesiology faculty from three institutions who are actively involved in preoperative assessment take the quizzes and assess the manner in which the content of the quiz reflected the content of the guidelines and adequately tested the scope of knowledge represented in the guidelines. This procedure followed a modified Delphi technique method, and such the quizzes went through several iterative changes that incorporated faculty input until no further improvements to the questions were suggested. As the grading of the quiz was objective with correct answers agreed upon by the group of faculty that evaluated the quiz, there was no reliability assessment of grading. There was also no reliability measure of participant performance as each quiz was only taken once and under different testing conditions (i.e., memory alone vs. cognitive aid). The stems and MCQs can be reviewed in appendix 2.

The response options were also categorized into four groups: correct, incorrect and associated with increased risk to the patient, incorrect and associated with increased cost to the patient or medical system, and incorrect and associated with both increased risk and cost. The categorization of incorrect responses was created by two authors (W.R.H. and M.D.M.) and subsequently validated by the remaining authors. Categorization was based on the intervention each answer option would require (i.e., obtaining an unnecessary electrocardiogram posed no risk to a patient needing nonemergent surgery but would increase cost and therefore would be categorized as incorrect and associated with increased cost to the patient or medical system). Participants were instructed to select the single best answer for each question from memory or as directed by the DST depending on the testing condition (with or without DST). The answers given to each item were recorded in a database along with the duration required to complete each nine-question quiz.

**Statistical Analyses**

Generalized linear mixed models (GLMMs) were used to evaluate the data because participants took two separate quizzes under two testing environments (with and without DST).\(^5\) The GLMMs were used to determine the independent effect of the use of the DST on overall scores, on individual quiz item scores, and on the duration of time spent taking the quizzes. These models adjusted for the potential influence of the specific quiz being administered (quiz 1 vs. 2), the order in which the quiz was taken (first vs. second), the presence or absence of the DST, and the study site (MUSC, University of North Carolina, University of Kentucky, or VUMC). By including these covariates, the GLMMs provided the opportunity for us to conduct F tests on each of these, to determine whether any of them had any independent association with the participants’ scores and quiz durations, and they provided, via the intraclass correlation coefficient, a means of estimating the reliability between subjects’ scores on the two quizzes. Interaction terms between the use of the DST and quiz order were assessed and included in the final models when they were moderately significant (\(P < 0.10\)) because of the potential for the DST to be differentially beneficial when taking the first versus second quiz (i.e., to account for a possible learning effect from use of the DST first). Overall quiz scores, durations, and number of incorrect responses that would result in increased healthcare costs were treated as continuous variables with a linear link function in their GLMM, whereas individual quiz item scores were treated as binary variables with logit link functions in their GLMM. Random participant effects were used in the GLMMs to account for correlation between scores within the same participants. For continuous measures, adjusted mean scores (i.e., scores that adjusted for order, site, and quiz effects) were estimated along with their 95% CIs using the GLMMs. For binary responses, the use of the DST was quantified using adjusted probabilities of correct responses, along with odds ratios and their 95% CIs. For each quiz, we also investigated internal consistency using Cronbach’s alpha. Analyses were conducted using SAS v9.3 (SAS Institute Inc., Cary, NC) Proc MIXED, Proc GLIMMIX, and the Proc CORR procedures, and \(P\) values less than 0.05 were considered statistically significant.
Results
A total of 111 residents (of 211 recruited) participated among the four sites (MUSC 37, University of Kentucky 33, University of North Carolina 25, and VUMC 16). One participant at MUSC consented to participate and finished 5 of 18 questions before being paged for an emergent intubation; his results were removed from the dataset. All datasets were complete with no ambiguity in answer selection or other potential source of error apparent to the authors. Of the 111 residents participating, 36 (32%) were postgraduate year (PGY) 4 or higher, 29 (26%) were PGY-3, 30 (27%) were PGY-2, and 16 (14%) were PGY-1 or interns (table 1 for demographic breakdown by group). Based on sex, race (Caucasian vs. non-Caucasian), or PGY, there were no significant differences between the four groups. Cronbach’s alpha coefficients were calculated to be 0.66 for quiz 1 and 0.65 for quiz 2, indicating a moderate level of internal consistency for each.

Table 2 summarizes the findings of the GLMM for participants’ overall scores. The intraclass correlation coefficient was estimated to be 0.26, indicating that subjects’ scores on quiz 1 tended to be moderately correlated with their scores on quiz 2. The scores on the quizzes (quiz 1 and 2) were comparable (P > 0.05), validating that there was not a significant difference in difficulty of the quiz. However, there was a clear order effect, meaning that subjects tended to score better on the second quiz when compared with the first, regardless of whether it was quiz 1 or 2. The results indicated that the DST had a statistically significant impact on both the first (P < 0.0001) and second (P < 0.0001) quiz scores, and the DST benefit was significantly more pronounced (P < 0.01) when used during the first quiz as compared with being used during the second quiz, as seen in table 2. On average, the DST resulted in participants scoring 2.9 points higher when the DST was used during their first quiz and 1.6 points higher when used during their second quiz, which represents an absolute improvement in performance of 32 and 18%, respectively, and a relative improvement of 60.4 and 23.9%, respectively. The DST by quiz interaction was not significant, providing some evidence of the DST’s generalizability. There was no significant (P = 0.98) difference in scores across the four sites (table 2). PGY also had no significant effect on performance, except that subjects with more years of training were significantly less likely (P = 0.04) to have an incorrect response that would increase risk to the patient. The odds of an incorrect response that increased risk to the patient declined by 31% with each additional year of training (odds ratio, 0.69; 95% CI, 0.48 to 0.99).

Table 3 summarizes the findings of the GLMM assessing the impact of the factors associated with the duration of time it took participants to take their quizzes. Durations were comparable on the two quizzes, and participants’ second quizzes took less time, on average, than their first quizzes. Use of the cognitive aid did increase the quiz duration on both their first (P < 0.0001) and second (P < 0.0001) quizzes, with the added duration being moderately longer on the first quiz compared with their second (4.4 vs. 2.4 added minutes, P = 0.06). There were differences in the quiz durations by site, with participants taking longer time at University of Kentucky; UNC = University of North Carolina; VUMC = Vanderbilt University Medical Center.

Table 4 summarizes the findings of the GLMMs assessing the impact of the cognitive aid on each of the individual quiz items. The cognitive aid was associated with a statistically significant improvement of correct responses for scenario question numbers 1, 3, 4, 7, 8, and 9, even after adjusting for quiz type, quiz order, and study site.

Use of the DST also had statistically significant impact on the number of incorrect quiz responses that were considered to increase healthcare costs, but not on incorrect responses that

<table>
<thead>
<tr>
<th>Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>31</td>
<td>26</td>
<td>26</td>
<td>28</td>
<td>111</td>
</tr>
<tr>
<td>Male</td>
<td>22</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>78 (70%)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>27</td>
<td>21</td>
<td>22</td>
<td>25</td>
<td>95 (86%)</td>
</tr>
<tr>
<td>PGY-1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>16 (14%)</td>
</tr>
<tr>
<td>PGY-2</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>30 (27%)</td>
</tr>
<tr>
<td>PGY-3</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>29 (26%)</td>
</tr>
<tr>
<td>PGY-4</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>36 (32%)</td>
</tr>
</tbody>
</table>

Demographics: the number of participants in groups A–D is shown (N). Further detail includes the number of male, Caucasian, and PGY 1–4 participants. N = number; PGY = postgraduate year.

Table 2. Results of the Generalized Linear Mixed Model Analyses Assessing the Impact of Cognitive Aid on Participants’ Overall Scores

<table>
<thead>
<tr>
<th>Effect</th>
<th>Adjusted Mean Score (Number Correct)</th>
<th>95% CI</th>
<th>P Value‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiz 1</td>
<td>6.9</td>
<td>6.7–7.2</td>
<td>0.74</td>
</tr>
<tr>
<td>Quiz 2</td>
<td>6.9</td>
<td>6.6–7.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>6.9</td>
<td>6.6–7.3</td>
<td>0.98</td>
</tr>
<tr>
<td>UNC</td>
<td>6.9</td>
<td>6.5–7.3</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MUSC</td>
<td>6.8</td>
<td>6.5–7.2</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>VUMC</td>
<td>6.9</td>
<td>6.4–7.4</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>First quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without cognitive aid</td>
<td>4.8</td>
<td>4.5–5.2</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>With cognitive aid</td>
<td>7.7</td>
<td>7.4–8.1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Impact of cognitive aid*</td>
<td>2.9†</td>
<td>2.4–3.4</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Second quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without cognitive aid</td>
<td>6.7</td>
<td>6.4–7.1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>With cognitive aid</td>
<td>8.3</td>
<td>8.0–8.7</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Impact of cognitive aid*</td>
<td>1.6‡</td>
<td>1.1–2.1</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* Impact of cognitive aid is defined as the difference between the scores with and without the cognitive aid. † P < 0.01 comparing impact of cognitive aid on first quiz compared with impact of cognitive aid on second quiz. ‡ Obtained via F tests based on results from the generalized linear mixed models.

MUSC = Medical University of South Carolina; UK = University of Kentucky; UNC = University of North Carolina; VUMC = Vanderbilt University Medical Center.

Decision Support for Perioperative Assessment

Anesthesiology 2014; 120:1339-53

Hand et al.
### Table 3. Results of the Analyses Assessing the Impact of Cognitive Aid on Participants' Quiz Durations

<table>
<thead>
<tr>
<th>Effect</th>
<th>Adjusted Quiz Duration (in Minutes)</th>
<th>95% CI</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiz 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With cognitive aid</td>
<td>9.8</td>
<td>9.1–10.4</td>
<td>0.30</td>
</tr>
<tr>
<td>Without cognitive aid</td>
<td>10.2</td>
<td>9.5–10.8</td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>10.9</td>
<td>10.0–11.8</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>UNC</td>
<td>9.0</td>
<td>8.0–10.1</td>
<td></td>
</tr>
<tr>
<td>MUSC</td>
<td>11.0</td>
<td>10.1–11.9</td>
<td></td>
</tr>
<tr>
<td>VUMC</td>
<td>9.0</td>
<td>7.6–10.3</td>
<td></td>
</tr>
<tr>
<td>First quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without cognitive aid</td>
<td>8.5</td>
<td>7.7–9.4</td>
<td></td>
</tr>
<tr>
<td>With cognitive aid</td>
<td>12.9</td>
<td>12.0–13.9</td>
<td></td>
</tr>
<tr>
<td>Impact of cognitive aid*</td>
<td>4.4</td>
<td>3.1–5.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Second quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without cognitive aid</td>
<td>8.0</td>
<td>7.1–8.9</td>
<td></td>
</tr>
<tr>
<td>With cognitive aid</td>
<td>10.4</td>
<td>9.6–11.3</td>
<td></td>
</tr>
<tr>
<td>Impact of cognitive aid*</td>
<td>2.4</td>
<td>1.2–3.7</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* Impact of cognitive aid is defined as the difference between the scores with and without the cognitive aid. † Obtained via F tests based on results from the generalized linear mixed models.

MUSC = Medical University of South Carolina; UK = University of Kentucky; UNC = University of North Carolina; VUMC = Vanderbilt University Medical Center.

### Table 4. Results of GLMMs Assessing the Impact of the Cognitive Aid on Individual Questions

<table>
<thead>
<tr>
<th>Quiz Item Number</th>
<th>Adjusted* Probability of Responding Correctly, %</th>
<th>Odds Ratio and 95% CI</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Without cognitive aid</td>
<td>55.3</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>1: With cognitive aid</td>
<td>94.6</td>
<td>14.1 (6.6–30.2)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>2: Without cognitive aid</td>
<td>83.3</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>2: With cognitive aid</td>
<td>89.9</td>
<td>1.8 (0.9–3.6)</td>
<td>0.11</td>
</tr>
<tr>
<td>3: Without cognitive aid</td>
<td>64.0</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>3: With cognitive aid</td>
<td>93.3</td>
<td>7.8 (3.4–18.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>4: Without cognitive aid</td>
<td>66.4</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>4: With cognitive aid</td>
<td>92.2</td>
<td>6.0 (2.5–14.0)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>5: Without cognitive aid</td>
<td>89.3</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>5: With cognitive aid</td>
<td>93.8</td>
<td>1.8 (0.7–4.7)</td>
<td>0.20</td>
</tr>
<tr>
<td>6: Without cognitive aid</td>
<td>73.1†</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>6: With cognitive aid</td>
<td>99.1†</td>
<td>38.8 (0.0–61.436)</td>
<td>0.33</td>
</tr>
<tr>
<td>7: Without cognitive aid</td>
<td>38.4</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>7: With cognitive aid</td>
<td>79.6</td>
<td>6.3 (3.8–10.3)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>8: Without cognitive aid</td>
<td>66.5</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>8: With cognitive aid</td>
<td>93.9</td>
<td>7.7 (2.7–22.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>9: Without cognitive aid</td>
<td>48.0</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>9: With cognitive aid</td>
<td>88.4</td>
<td>8.3 (4.0–17.1)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* Probability of correct responses was adjusted for quiz type (1 vs. 2), order (first vs. second), study site (University of Kentucky, University of North Carolina, Medical University of South Carolina, or Vanderbilt), and participant effects using GLMMs. † Obtained via F tests based on results from the GLMMs. On quiz item number 6 in quiz 2, all participants responded correctly (with and without the cognitive aid), and this resulted in non estimable parameters within the GLMM; thus, the results only reflect the impact of the cognitive aid within quiz 1.

GLMM = generalized linear mixed model.

### Table 5. Results of GLMMs Assessing the Impact of the Cognitive Aid on Incorrect Responses That Would Result in Increased Cost

<table>
<thead>
<tr>
<th>Incorrect responses that increase costs</th>
<th>Adjusted* Mean (Number Incorrect)</th>
<th>95% CI</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>First quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without cognitive aid</td>
<td>4.0</td>
<td>3.7–4.4</td>
<td></td>
</tr>
<tr>
<td>With cognitive aid</td>
<td>1.0</td>
<td>0.6–1.3</td>
<td></td>
</tr>
<tr>
<td>Impact of cognitive aid‡</td>
<td>−3.1</td>
<td>−3.6 to −2.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Second quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without cognitive aid</td>
<td>2.1</td>
<td>1.8–2.5</td>
<td></td>
</tr>
<tr>
<td>With cognitive aid</td>
<td>0.5</td>
<td>0.2–0.9</td>
<td></td>
</tr>
<tr>
<td>Impact of cognitive aid‡</td>
<td>−1.6</td>
<td>−2.1 to −1.1</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* Mean number of incorrect responses were adjusted for quiz type (1 vs. 2), order (first vs. second), study site (University of Kentucky, University of North Carolina, Medical University of South Carolina, or Vanderbilt), and participant effects using GLMMs. † Obtained via F tests based on results from the GLMMs. ‡ Impact of cognitive aid is defined as the difference between the scores with and without the cognitive aid. GLMM = generalized linear mixed model.

Discussion

Our results demonstrate three important findings. First, the use of an electronic DST increased adherence to the 2007 ACC/AHA guidelines. Second, use of the electronic DST reduced management errors associated with increased costs, such as ordering unnecessary tests and consults. Third, use of the DST resulted in a longer time to complete the test.

The electronic, smart-phone–based DST improved absolute performance, as measured by adherence to the ACC/AHA preoperative assessment algorithm, by an average of 25% as compared with performance from memory alone. This result was consistent across residents at all years of training at four different institutions. In specific, the DST provided a statistically significant benefit for six of the nine MCQ scenarios (table 4). The recent works by Vigoda et al.4,5 highlight the inability of residents and practicing anesthesiologists to apply the ACC/AHA guidelines on perioperative cardiovascular evaluation and care for noncardiac surgery. Our results among residents taking the first test without the assistance of DST correctly answered an average of 4.8/9 (53%) questions. Similar to Vigoda’s studies, our rate of error beckons for a mechanism of improvement—and the improvement is seen with our simple DST. Under the same would increase risk to patients. Overall, there were 468 incorrect selections. Of the incorrect selections, 387 (83% of incorrect) were assigned to the “increased cost” category, 38 (8%) to the “increased risk” category, and 43 (9%) to the “increased cost and risk” category. Table 5 highlights the fact that the impact of the DST on quiz items healthcare costs was statistically significant (P < 0.0001) for both quizzes. Interestingly, the impact was greater if the DST was used during the participants’ first quiz when compared with their second (P < 0.001).
first-quiz conditions, the DST group correctly answered 7.7/9 (86%) questions. The efficacy of the DST is consistent with other treatment algorithms that have been assessed with DSTs with excellent uptake and success.19–24,36–38 Perhaps, the reason DSTs were adopted for Advanced Cardiac Life Support is related to the high-stakes and high-stress clinical environments in which code scenarios usually occur, but this research advocates the use of a DST in the calm and routine preoperative care setting. The authors’ propose that better adherence to published guidelines may, in fact, reduce the number of high-stakes events being managed with a DST due to higher quality care.

The DST provided statistically significant improvement for six of the nine possible scenarios when evaluating performance on both quiz 1 and 2. Scenario 6 did not reach statistical significance because all respondents correctly answered this question on quiz 2, creating nonestimable parameter for the GLMM; we remain encouraged by the absolute improvement despite the mathematical nuance created by the lack of an estimable error rate. That said the composite percentages of 73 versus 99% for testing without and with DST, respectively, certainly seem to show benefits consistent with the other scenarios showing statistically significant improvement. For clinical scenarios number 2 and 5, statistical significance was not reached despite the fact that the scores with DST were among the highest for any scenario. We are encouraged that the absolute percentages were higher with the DST and believe with a large sample size we would have had the power to demonstrate a statistically significant improvement. The high scores for these scenarios in the memory-alone group were similar to these in the DST group.

The second finding of interest is that 83% of the incorrect options selected by participants would increase the cost of care either to the patient directly or the medical system providing care to the patient (e.g., requesting an unwarranted electrocardiogram or choosing to “delay until blood pressure controlled”). During the design of this study, the authors associated each possible incorrect MCQ option with increasing the cost, risk, or both cost and risk to the patient. Participants testing with or without the DST did well avoiding selections that increased risk to the patient; however, 83% of incorrect selections included decisions to overtreat or delay procedures at an increased cost to the patient and healthcare system. When using the DST, participants made 77% fewer incorrect responses that would have resulted in increased costs compared with answering from memory alone (3.1/4 for quiz 1 and 1.6/2.1 for quiz 2).

The third finding of interest is that participants required an average of 3.4 min (204 s) longer to complete the quizzes when using the DST. This equates to a mere 23 additional seconds per question (204 s + 9 questions) required to improve adherence to an important guideline in perioperative medicine. Hospitals face the demand to deliver higher quality care with greater efficiency than ever before. However, we posit that the investment of fractionally more time using a DST would be more than offset by the correct application of best-practice guideline. We also expect that efficiency using this DST would improve with familiarity—as previously described, the residents had almost no previous exposure to the DST software to prevent any excessive learning effect that would have confounded the “from memory” test scores. These 23 s may be seen as an investment in quality as they portended improved adherence to the guideline in question. In comparison, “the average surgical ‘time out’ at VUMC consumes 57 s (average of last 172,000 electronically observed time outs), and the ‘time out’ has been credited with reducing surgical site and procedure errors.” (Jesse Ehrenfeld, MD, Associate Professor, Department of Anesthesiology, Vanderbilt University, Nashville, Tennessee) Several studies have shown that deviation from published guidelines increases morbidity and overall hospital cost. As any DST is considered for adoption, one must recognize that implementation requires planning and education as there is often hidden work associated with new technologies.39–41 With adequate education, orientation to a DST, and rapid practitioner adoption, the authors believe the benefits shown in this study could be magnified.

A limitation is that we evaluated performance of anesthesiologists-in-training, not practicing physicians. Improved performance by nonresidents cannot be inferred from this research, despite similarly poor performance by practicing physicians demonstrated by Vigoda et al.4,5 In addition, although the results of this study seem generalizable among several institutions, the generalizability of the use of an electronic DST beyond the scope of the ACC/AHA guidelines cannot be made at this time. Future research needs to address whether similar tools will be of benefit for other perioperative management guidelines.42–44 It would be intuitive to assume that any such DST could improve adherence to guidelines, but we and others have shown that simply having guidelines in front of the clinician does not guarantee adherence.4,45,46 Specific work is needed to investigate the human factors in each application. In addition, the quizzes used in this study were not created using rigorous psychometric methods; given their moderate degree of consistency (Cronbach’s alpha coefficients), it is possible that the DST might perform differently (better or worse) if different assessment tools were used.

We found that the use of the DST was associated with more work per simulated patient encounter. Future research in this domain of implementation science needs to address whether this differential amount of time exists in simulated encounters with human standardized patients and in clinical settings. Our design could not determine the effect that order had beyond a simple description of percent correct. For example, we were unable to identify which type of errors might be prevented by a test taker who did or did not have previous exposure to the DST. We, therefore, are unable to make specific statements about prospective benefits of DST use.

Our small sample size prevented us from making statements about each question stem independently and resulted in the conclusion that the DST improved performance on only six of nine MCQ stems. Sample size was estimated based
on overall percent correct and an improvement with the DST. We did not perform a reliability assessment (e.g., test–retest) to test whether the grade achieved precisely represents a stable description of participant knowledge. Finally, we did not achieve perfect adherence to and application of the AHA/ACC guidelines under consideration, even though the DST contained all of the proper logic for such application. This failure is likely multifactorial: participant error, lack of attention, lack of knowledge (e.g., what constitutes an emergent case), or lack of familiarity with the DST itself. The design of the study prevented substantial practice navigating the DST to maintain a valid baseline for scores without DST assistance.

In summary, we demonstrated that a simple DST improved resident adherence to the ACC/AHA guidelines on perioperative cardiovascular evaluation and care for noncardiac surgery evaluated via MCQs with clinical stems. There was a significant improvement in overall performance of residents when using the DST. Residents selected fewer incorrect options in six of nine scenarios and primarily reduced incorrect selections that would increase the cost of medical care. As anesthesiologists strive to practice evidence-based medicine, these results indicate that it may be time to embrace decision support technology during the preoperative assessment and planning phase of patient care. Future research needs to address the implementation of such tools in the clinical setting.

Acknowledgments

This study was supported by the Foundation for Anesthesia Education and Research (Rochester, Minnesota); Research in Education Grant (Principal Investigator: Dr. McEvoy). This project was also supported by the South Carolina Clinical and Translational Research Institute, Medical University of South Carolina’s Clinical and Translational Institute (Charleston, South Carolina); National Institutes of Health/National Center for Advancing Translational Sciences (Bethesda, Maryland) grant no. UL1TR000062; and National Center for Research Resources award no. UL1RR029882 (Charleston, South Carolina).

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Hand: Department of Anesthesia and Perioperative Medicine, Medical University of South Carolina, 167 Ashley Avenue, Suite 301, Charleston, South Carolina 29425. handw@musc.edu. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

References


22. Vause J: Adherence to evidence-based guidelines is the key to improved health outcomes for general practice patients: YES. J Prim Health Care 2012; 4:156–8


47. Practice alert for the perioperative management of patients with coronary artery stents: A report by the American Society of Anesthesiologists Committee on Standards and Practice Parameters. ANESTHESIOLOGY 2009; 110:22–3


Appendix 1. Various Screenshots from the Decision Support Tool as Seen by a User. The Screenshots Demonstrate Both User Interface (Questions to Answer) and Specific End Points of the Preoperative Testing Algorithm. The Actual Decision Support Tool Uses a Touchscreen Interface and Only One Screen Is Visible at Any Time.

(Continued)
Appendix 1. (Continued)

Decision Support for Perioperative Assessment

Appendix 1. (Continued)
Appendix 1. (Continued)

Appendix 2. Multiple Choice Questions. Quizzes A and B. The Scenario Descriptions and Multiple Choice Questions Were Presented with and without the Decision Support Tool Available and in Assigned Order to Each Participant

<table>
<thead>
<tr>
<th>P #</th>
<th>Cognitive Aid</th>
<th>YES</th>
<th>NO</th>
<th>Time Start:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiz A</td>
<td>Order:</td>
<td>1st</td>
<td>2nd</td>
<td>Time Finish:</td>
</tr>
</tbody>
</table>

1. 65-yr-old male with 40 pack-year tobacco history, obesity, and anxiety disorder was recently diagnosed with laryngeal cancer presenting for radical neck dissection. He does no physical activity due to social anxiety. His father died of heart disease at 45 yr. The patient’s preoperative vital signs are T: 36.5, HR: 82, BP: 155/72, RR: 16.

Multiple choice:
A. Proceed with planned surgery.
B. Refer for treadmill cardiac stress test.
C. Need ECG before I can make a recommendation.
D. Delay surgery until BP is <140/70.
E. Refer for cardiac consultation.

2. 61-yr-old female with controlled hypertension, insulin-dependent diabetes (Hbg A1c 9.9%), and gastro-esophageal reflux disease is scheduled for metatarsal tendon release for bilateral “hammer toes.” Glucose ranges from 110 to 230 mg/dl. She works as a waitress and walks to and from work, worsening her foot pain. The patient’s preoperative vital signs are T: 36.5, HR: 78, BP: 124/82, RR: 16.

Multiple choice:
A. Proceed to surgery.
B. Order a stress test (dobutamine echo, exercise, or nuclear imaging).
C. Start β-blocker therapy and delay surgery until heart rate adequately controlled.
D. Delay surgery until Hgb A1c <7%.
E. Patient needs ECG before I can make a recommendation.
3. 72-yr-old male with a history of poorly controlled diabetes on insulin therapy, CHF, TIA with residual weakness on his left side, hypertension, and a 50 pack-year history of smoking presents for elective repair of 6.2-cm abdominal aortic aneurysm. He is able to walk only one block at a time because of left sided weakness. The patient’s preoperative vital signs are T: 36.5, HR: 98, BP: 165/85, RR: 18.

Multiple choice:
A. Proceed to planned surgery.
B. Patient needs to have PFTs before proceeding.
C. Consider testing if it will change management.
D. Delay surgery overnight to collect records from outside providers.
E. Obtain immediate cardiology consultation.

4. 4.71-yr-old male with poorly controlled insulin-dependent diabetes and hypertension which led to renal dysfunction and peripheral neuropathy describes his activity as being limited to dressing, brushing his teeth, and clicking the television remote control. His preoperative creatinine is 2.3 mg/dl. Colonoscopy revealed colon cancer and you are evaluating him for a colectomy in 5 days. The patient’s preoperative vital signs are T: 36.5, HR: 82, BP: 144/78, RR: 14.

Multiple choice:
A. Stress test is needed before proceeding to OR.
B. Start β-blockers and delay surgery 1 month.
C. Refer for 2-D echocardiogram.
D. Need ECG before I can make recommendation.
E. Proceed to surgery with heart rate control or consider noninvasive testing if it will change management.

5. 68-yr-old female with poorly controlled diabetes and hypertension presents after an automobile accident with an open femur fracture and cold right foot. The patient is stable and answers your questions. She is unable to do more than her daily chores without becoming short of breath and had a stroke 2 yr ago. The patient’s preoperative vital signs are T: 36.5, HR: 98, BP: 156/55, RR: 14.

Multiple choice:
A. Obtain immediate cardiology consultation.
B. Patient needs to have 2-D echocardiogram before I make recommendation.
C. Need ECG before I make recommendation.
D. Consider noninvasive testing if it would change management.
E. Proceed to planned surgery.

6. 62-yr-old male with 40 pack-year smoking history and ischemic stroke 2 yr ago presents for an elective femoral-popliteal bypass due to claudication. He can walk several blocks but is ultimately limited by leg pain. The patient’s preoperative vital signs are T: 36.5, HR: 78, BP: 178/82, RR: 12.

Multiple choice:
A. Refer for cardiology consultation.
B. Need ECG before I can make recommendation.
C. Delay surgery until BP <140/70.
D. Proceed to surgery with heart rate control or consider noninvasive testing if it will change management.
E. Stress test is required before going to OR.

7. 66-yr-old male active smoker (1ppd × 20 yr) with severe osteoarthritis is scheduled for total knee arthroplasty. He suffered an MI 10 weeks ago but did not require percutaneous intervention and is completing rigorous cardiac rehabilitation despite knee pain. He was discharged with β-blocker, statin, and ACE inhibitor. He is compliant with medications and endorses only knee pain as limitation to exertion. The patient’s preoperative vital signs are T: 36.5, HR: 62, BP: 118/59, RR: 14.

Multiple choice:
A. Surgery should be delayed at least 3 months after a myocardial infarction.
B. Patient should have a stress test (dobutamine echo, exercise, or nuclear image) before surgery.
C. Proceed with planned surgery.
D. Surgery should be delayed at least 6 months after a myocardial infarction.
E. Need ECG and pulmonary function tests before I can make a recommendation.
8. 72-yr-old male with history of calf claudication who has type 2 diabetes requiring insulin therapy, hypertension and chronic renal insufficiency (Cr 2.2 mg/dl) is scheduled for femoral-popliteal bypass. He endorses worsening chest pain and shortness of breath while doing household chores. The patient’s preoperative vital signs are T: 36.5, HR: 55, BP: 145/85, RR: 18.

Multiple choice:
A. Patient needs a cardiac catheterization immediately.
B. Patient should have a stress test, β-blockade, and cardiology consult before surgery.
C. Proceed with surgery if today’s ECG is unchanged from ECG performed 5 months ago.
D. Proceed with planned surgery with HR control.
E. Proceed with planned surgery.

9. 68-yr-old male presents with poorly controlled diabetes on insulin therapy, hypertension, chronic kidney dysfunction with creatinine of 2.1, congestive heart failure, and 40 pack-year history of smoking. He is unable to walk more than three blocks secondary to fatigue and some knee pain. He is presenting for total knee replacement. The patient’s preoperative vital signs are T: 36.5, HR: 98, BP: 165/85, RR: 18.

Multiple choice:
A. Proceed to planned surgery.
B. Patient needs to have PFTs before I make recommendation.
C. Need ECG before I make recommendation.
D. Proceed to surgery with heart rate control or consider noninvasive testing if it will change management.
E. Obtain immediate cardiology consultation.

P # ________ Cognitive Aid YES NO Time Start:__________

Quiz B Order: 1st 2nd Time Finish:_________

1. 66-yr-old female with uncontrolled diabetes with diabetic nephropathy (creatinine of 2.1), hypertension, TIA 2 yr ago with no residual symptoms, and a symptomatic pituitary adenoma presents for elective transphenoidal resection of pituitary tumor. She manages her ADLs but does not leave her residence due to fatigue. The patient’s preoperative vital signs are T: 36.5, HR: 98, BP: 165/85, RR: 18.

Multiple choice:
A. Proceed to planned surgery.
B. Patient needs to have 2-D echocardiogram before I make recommendation.
C. Need ECG before I make recommendation.
D. Proceed to surgery with heart rate control or consider noninvasive testing if it will change management.
E. Obtain immediate cardiology consultation.

2. 67-yr-old female with uncontrolled diabetes, hypertension, and dialysis-dependent renal disease presents for same-day evaluation from dialysis clinic. She has developed a clot and infection in her A-V fistula, which has since developed subcutaneous crepitus. Patient is posted for exploration and revision. She has poor functional capacity and her last stress test (last week) was consistent with inducible ischemia. Patient has scheduled an appointment with her cardiologist in 1 week to discuss intervention. The patient’s preoperative vital signs are T: 36.5, HR: 128, BP: 156/55, RR: 14.

Multiple choice:
A. Obtain immediate cardiology consultation.
B. Patient needs to have 2-D echocardiogram before I make recommendation.
C. Need ECG before I make recommendation.
D. Consider noninvasive testing if it would change management.
E. Proceed to surgery.

3. 70-yr-old female who had a myocardial infarction 6 yr ago presents for preoperative evaluation for nephrectomy to treat renal cell carcinoma. She has chronic renal insufficiency (Cr 2.5 mg/dl), hypertension, and indicates she has to rest after walking two blocks on flat ground but symptoms resolve quickly with rest. The patient’s preoperative vital signs are T: 36.5, HR: 82, BP: 144/78, RR: 14.

Multiple choice:
A. PFTs needed before proceeding to OR.
B. Start β-blockers and delay surgery 4 weeks.
C. Refer for 2-D echocardiogram.
D. Need ECG before I can make recommendation.
E. Proceed to surgery with heart rate control or consider noninvasive testing if it will change management.
4. 66-yr-old female with hypertension and hypothyroidism presents for evaluation before her scheduled lumbar laminectomy and fusion. Her leg pain prevents her from standing and she is not sure what exertion she is capable of. Patient takes a statin for hypercholesterolemia and thyroid hormone replacement. The patient's preoperative vital signs are T: 36.5, HR: 82, BP: 125/72, RR: 16.

Multiple choice:
A. Proceed with planned surgery.
B. Refer for dobutamine stress test due to inactivity.
C. Need ECG before I can make a recommendation.
D. Delay surgery until BP is <140/70.
E. Refer for cardiac consultation.

5. 66-yr-old female with hypertension, chronic renal insufficiency (Cr 2.2 mg/dl), and type 2 diabetes endorses fatigue while folding laundry and new chest pressure walking to mailbox on flat land. The patient presents for elective knee arthroplasty for degenerative joint disease. The patient's preoperative vital signs are T: 36.5, HR: 55, BP: 145/85, RR: 18.

Multiple choice:
A. Patient needs a cardiac catheterization immediately.
B. Patient should have a stress test, β-blockade, and cardiology consult before surgery.
C. Proceed with surgery if today's ECG is unchanged from ECG performed 5 months ago.
D. Proceed with planned surgery with HR control.
E. Proceed with planned surgery.

6. 64-yr-old female with poorly controlled type 2 diabetes (Hbg A1c 10.1%), controlled hypertension, and gout presents for laparoscopic cholecystectomy. Blood glucose ranges from 110 to 230 mg/dl and she says she is able to keep up with her peers in a twice-per-week aerobics class. The patient's preoperative vital signs are T: 36.5, HR: 78, BP: 124/82, RR: 14.

Multiple choice:
A. Proceed to surgery.
B. Order a stress test (dobutamine echo, exercise, or nuclear imaging).
C. Start β-blocker therapy and delay surgery until heart rate adequately controlled.
D. Delay surgery until Hgb A1c <7%.
E. Patient needs ECG before I can make a recommendation.

7. 70-yr-old female with long smoking history (1 ppd) presents for lumbar laminectomy due to radicular pain unsuccessfully treated by steroid injections. After her myocardial infarction two and a half months ago, she was medically treated with β-blocker, ACE inhibitor, and statin. She swims 20 min without stopping every day. The patient's preoperative vital signs are T: 36.5, HR: 62, BP: 118/59, RR: 14.

Multiple choice:
A. Surgery should be delayed at least 3 months after a myocardial infarction.
B. Patient should have a stress test (dobutamine echo, exercise, or nuclear image) before surgery.
C. Proceed with planned surgery.
D. Surgery should be delayed at least 6 months after a myocardial infarction.
E. Need ECG and pulmonary function tests before I can make a recommendation.

8. 74-yr-old female with controlled hypertension, uncontrolled type 2 diabetes, and peripheral vascular disease presents for evaluation before having a newly diagnosed 5.6-cm abdominal aortic aneurysm repair. Her exertional capacity is limited to light housework before fatiguing. The patient's preoperative vital signs are T: 36.5, HR: 78, BP: 178/82, RR: 12.

Multiple choice:
A. Refer for cardiology consultation.
B. Need ECG before I can make recommendation.
C. Delay surgery until BP <140/70.
D. Proceed to surgery with heart rate control or consider noninvasive testing if it will change management.
E. Stress test is required before going to OR.
9. 74-yr-old male who suffered an ischemic stroke 2 yr ago presents with stable calf claudication after two blocks that limits exertion, uncontrolled hypertension, and uncontrolled insulin-dependent diabetes is scheduled for elective ileo-femoral bypass. Patient states he does not think he has any heart issues and has not visited a cardiologist since his stroke. The patient's preoperative vital signs are T: 36.5, HR: 98, BP: 165/85, RR: 18.

Multiple choice:
A. Proceed to planned surgery.
B. Patient needs to have PFTs before I make recommendation.
C. Proceed to surgery with heart rate control or consider noninvasive testing if it will change management.
D. Delay surgery overnight to collect records from outside providers.
E. Obtain immediate cardiology consultation.

ANESTHESIOLOGY REFLECTIONS FROM THE WOOD LIBRARY-MUSEUM

Advertising Maltine with Coca Wine

The roots of the Maltine tree (left) were advertised as “concentrated extract of malted wheat, oats, and barley.” A major branch of the Maltine tree was the tonic “Maltine with Coca Wine,” each ounce of which contains “thirty grains of assayed Huanaco Coca leaves…” According to an 1894 issue of the National Medical Review, the “Coca boosts the patient and the maltine furnishes the peg that prevents him from slipping back.” Because Maltine had been widely distributed to the public as a stimulant in beverages and in foods (such as Maltine with Coca Wine), cocaine was rapidly accepted by laymen when used as a local anesthetic by dentists and physicians. (Copyright © the American Society of Anesthesiologists, Inc.)

George S. Bause, M.D., M.P.H., Honorary Curator, ASA’s Wood Library-Museum of Anesthesiology, Park Ridge, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYC@aol.com.