Scientific Evidence Underlying the ACC/AHA Clinical Practice Guidelines

Pierluigi Tricoci, MD, MHS, PhD
Joseph M. Allen, MA
Judith M. Kramer, MD, MS
Robert M. Califf, MD
Sidney C. Smith Jr, MD

Clinical Practice Guidelines are systematically developed statements to assist practitioners with decisions about appropriate health care for specific patients’ circumstances. Guidelines are often assumed to be the epitome of evidence-based medicine. Yet, guideline recommendations imply not only an evaluation of the evidence but also a value judgment based on personal or organizational preferences regarding the various risks and benefits of a medical intervention for a population.

For more than 20 years, the American College of Cardiology (ACC) and the American Heart Association (AHA) have released clinical practice guidelines to provide recommendations on care of patients with cardiovascular disease. The ACC/AHA guidelines currently use a grading schema based on level of evidence and class of recommendation (available at http://www.acc.org and http://www.aha.org). The level of evidence classification combines an objective description of the existence and the types of studies supporting the recommendation and expert consensus, according to 1 of the following 3 categories:

- Level of evidence A: recommendation based on evidence from multiple randomized trials or meta-analyses
- Level of evidence B: recommendation based on evidence from a single randomized trial or nonrandomized studies
- Level of evidence C: recommendation based on expert opinion, case studies, or standards of care

The class of recommendation designation indicates the strength of a recommendation and requires guideline writers not only to make a judgment

See also p 870 and Patient Page.

Context The joint cardiovascular practice guidelines of the American College of Cardiology (ACC) and the American Heart Association (AHA) have become important documents for guiding cardiology practice and establishing benchmarks for quality of care.

Objective To describe the evolution of recommendations in ACC/AHA cardiovascular guidelines and the distribution of recommendations across classes of recommendations and levels of evidence.

Data Sources and Study Selection Data from all ACC/AHA practice guidelines issued from 1984 to September 2008 were abstracted by personnel in the ACC Science and Quality Division. Fifty-three guidelines on 22 topics, including a total of 7196 recommendations, were abstracted.

Data Extraction The number of recommendations and the distribution of classes of recommendation (I, II, and III) and levels of evidence (A, B, and C) were determined. The subset of guidelines that were current as of September 2008 was evaluated to describe changes in recommendations between the first and current versions as well as patterns in levels of evidence used in the current versions.

Results Among guidelines with at least 1 revision or update by September 2008, the number of recommendations increased from 1330 to 1973 (+48%) from the first to the current version, with the largest increase observed in use of class II recommendations. Considering the 16 current guidelines reporting levels of evidence, only 314 recommendations of 2711 total are classified as level of evidence A (median, 11%), whereas 1246 (median, 48%) are level of evidence C. Level of evidence significantly varies across categories of guidelines (disease, intervention, or diagnostic) and across individual guidelines. Recommendations with level of evidence A are mostly concentrated in class I, but only 245 of 1305 class I recommendations have level of evidence A (median, 19%).

Conclusions Recommendations issued in current ACC/AHA clinical practice guidelines are largely developed from lower levels of evidence or expert opinion. The proportion of recommendations for which there is no conclusive evidence is also growing. These findings highlight the need to improve the process of writing guidelines and to expand the evidence base from which clinical practice guidelines are derived.

JAMA. 2009;301(8):831-841

Author Affiliations: Division of Cardiology and Duke Clinical Research Institute (Dr Tricoci), Division of General Internal Medicine and Duke Center for Education and Research on Therapeutics (Dr Kramer), and Division of Cardiology and Duke Translational Medicine Institute (Dr Califf), Duke University, Durham, North Carolina; American College of Cardiology Science and Quality Division, Washington, DC (Mr Allen); and Center for Cardiovascular Science and Medicine, University of North Carolina, Chapel Hill (Dr Smith).

Corresponding Author: Pierluigi Tricoci, MD, MHS, PhD, Duke Clinical Research Institute, 2400 Pratt St, Room 0311, Terrace Level, Durham, NC 27705 (trico001@dcri.duke.edu).
about the relative strengths and weaknesses of the study data but also to make a value judgment about the relative importance of the risks and benefits identified by the evidence and to synthesize conflicting findings among multiple studies. Definitions of the classes of recommendation are as follows:

- **Class I**: conditions for which there is evidence and/or general agreement that a given procedure or treatment is useful and effective
- **Class II**: conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment
- **Class IIa**: weight of evidence/opinion is in favor of usefulness/efficacy
- **Class IIb**: usefulness/efficacy is less well established by evidence/opinion
- **Class III**: conditions for which there is evidence and/or general agreement that the procedure/treatment is not useful/effective and in some cases may be harmful.

Thus, level of evidence C and class II indicate, respectively, recommendations lacking supporting evidence and those subject to uncertainties about the appropriate medical decision.

The significant increase in the quantity of scientific literature concerning cardiovascular disease published in recent years (along with the number of technical and medical advances)—if aimed to address unresolved issues confronting guideline writers—should have resulted in guideline recommendations with more certainty and supporting evidence. However, whether guidelines have truly evolved in this direction has not been systematically investigated.

Thus, we performed a systematic review of the ACC/AHA clinical practice guidelines issued from 1984 to September 2008, with intent of examining all guidelines published since 1984 for changes associated with the use of class of recommendation grading schema both for individual guidelines and categories of guidelines and evaluating the adequacy of evidence behind current guideline recommendations. Our ultimate goal was to elucidate possible gaps that may limit the evidence-based foundations of ACC/AHA guidelines and to highlight potential opportunities for improvement.

**METHODS**

All ACC/AHA practice guidelines issued from 1984 to September 2008 were abstracted by personnel in the ACC Science and Quality Division to obtain the number of recommendations within each class of recommendation and the distribution of level of evidence designations across all guidelines. The recommendations are clearly displayed statements highlighted in each guideline document and separated from the remainder of the text. Each recommendation contains a specific designation reflecting the class of recommendation and the level of evidence. Therefore, the abstraction of the data performed for this analysis only reflected the content of the documents and was not subject to any judgment by the abstractors.

Current guidelines were defined as those posted on the ACC Web site on September 30, 2008 (http://www.acc.org/qualityandscience/clinical/topic/topic.htm#guidelines). The review included only comprehensive guideline documents; focused updates were not included because these represent an update on a limited number of recommendations and are not reflective of an entire topic. Guidelines were classified into the following categories: (1) disease-based guidelines; (2) interventional procedure–based guidelines; and (3) diagnostic procedure–based guidelines.

The aim of the analysis was to report the distribution of recommendations across classes of recommendation and levels of evidence. For current guidelines for which at least 1 previous version was available, changes in the use of classes of recommendation were evaluated by comparing the first version with the current version. Because levels of evidence were introduced only in 1998 and not consistently adopted after they were introduced, only 6 of 17 current guidelines have a previous version reporting level of evidence and were suitable to assess changes (atrial fibrillation, heart failure, stable angina, unstable angina, pacemaker, and percutaneous coronary intervention). We reported the use of class of recommendations and level of evidence in current guidelines, defined as above.

Because individual guidelines may vary widely in the numbers of recommendations, in order to weigh equally each guideline subject within each category (ie, disease, interventional, or diagnostic), the summary of the distribution of guideline recommendations within a category across the grading schemes is shown as the median of the percentage reported for each guideline subject in question. Median values are also reported to summarize the changes in each of the categories.

**RESULTS**

**Historical Summary**

From 1984 to September 2008, the ACC/AHA Joint Task Force issued 53 guidelines on 22 topics, including a total of 7196 recommendations. Among the 53 guidelines, 24 were disease-based, 15 were interventional procedure–based, and 14 were diagnostic procedure–based. In 1990, the class II recommendation was expanded to include classes IIa and IIb. As of September 2008, 17 of the 53 guidelines were listed as the current guidelines on the ACC Web site.

**Observed Changes Over Time**

The ACC/AHA guidelines are periodically updated. Among the current guidelines, 12 are revisions of previously issued documents. The mean time elapsing from the publication of a version to the update was 4.6 years (SD, 1.8 years) for disease-based, 5.4 years (SD, 2.1 years) for interventional procedure–based, and 8.2 years (SD, 2.4 years) for diagnostic procedure–based guidelines.

Considering only the current guidelines with at least 1 revision, the total number of recommendations in

---

©2009 American Medical Association. All rights reserved.
creased from 1330 to 1973 (48% increase). The raw increase in number of recommendations was higher for diagnostic procedure–based guidelines (242 additional recommendations) than for interventional procedure–based (130 additional recommendations) or disease-based (101 additional recommendations) guidelines (Table 1).

Overall, the guidelines shifted to more class II recommendations and fewer class III recommendations, while the use of class I recommendations remained fairly constant over time (Table 1). Among disease and interventional guidelines, there was a definite trend toward more class II recommendations, while the proportion of class I recommendations decreased. In diagnostic guidelines, there was a greater increase in class I recommendations and a decrease in class II recommendations. In addition, the proportion of class III recommendations decreased among all guidelines, but especially for interventional guidelines.

Overall, among current guidelines, there were 1124 class II recommendations of 3044 total recommendations, with a median of 41% (interquartile range [IQR], 29%-51%) of recommendations in class II across the guidelines (Table 2).

**Level of Evidence**

From the introduction of levels of evidence in 1998 through September 2008, 33 guidelines have been released, of which 27 adopted level of evidence classification and 6 did not. Among current guidelines, only echocardiography guidelines do not report level of evidence. The 16 current guidelines reporting levels of evidence, comprising a total of 2711 recommendations, classify 314 recommendations as level of evidence A (median, 11% [IQR, 6%-16%]), whereas 1246 have level of evidence C (median, 48% [IQR, 26%-57%]) (Table 2).

Among disease-based guidelines, which generally have a greater proportion of level of evidence A, there is great variability regarding the use of levels of evidence. Unstable angina/non–ST-segment elevation myocardial infarction, heart failure, and secondary prevention guidelines have more than 20% recommendations with level of evidence A, whereas valvular heart disease guidelines have only 1 recommendation (320 total; 0.3%) with level of evidence A (Table 2). Individually, most of the current guidelines include more than 50% level of evidence C recommendations, with valvular heart disease guidelines having the highest percentage at 71% (226/320).

Level of evidence A recommendations are mostly concentrated in class I (Table 1). Nonetheless, among all 1305 class I recommendations of guidelines reporting level of evidence, only 245 have level of evidence A (median, 19% [IQR, 11%-30%]); whereas 481 (median, 36% [IQR, 20%-50%]) have level of evidence C. Only 6 of 17 current guidelines have a previous version reporting level of evidence. Such guidelines were those updated more frequently. In this small subset, compared with the first versions reporting levels of evidence, there was a median of 6 additional level A recommendations (IQR, 5-11), 13 level B recommendations (IQR, 3-29), and 24 level C recommendations (IQR, 14-25).

**COMMENT**

The ACC/AHA guidelines—as an established guidance for management of cardiovascular disease—have progressively increased the number of recommendations, but these recommendations largely reflect a lower certainty of evidence. Furthermore, in current guidelines, level of evidence C—indicating recommendations based solely on expert opinion, case studies, or “standard of care”—is the most frequent designation. These findings point to consistent gaps in evidence about medical practices and the need to generate the research required to close gaps in knowledge.

**Guidelines’ Role and Evolution of ACC/AHA Practice Guidelines**

There is a broad consensus that medical practice should be based on evidence about outcomes of therapies and interventions and in agreement with the values and preferences of the patient. This construct of evidence-based medicine is predicated on the existence of a body of evidence of benefits and risks that can be distilled into value judgments about reasonable actions that are expressed in clinical practice guidelines.

During the last decade, the need for development of guidelines has increased because of advances in development of drugs and devices resulting in greater complexity for the diagnosis and treatment of cardiovascular diseases. Potential increases in health costs and risks due to marketing-driven, uncontrollable use of novel clinical options also make guidelines increasingly important. Furthermore, the most solid evidence in guidelines is now used to develop performance measures, which are used, in turn, to judge the quality of practice, often in the context of differential payment or public reporting.

In this setting, the ACC/AHA guidelines have assumed a critical role in the establishment of standards of cardiac care and in providing benchmarks to define quality of care. As such, it is important to recognize current limitations of the ACC/AHA guidelines to identify potential areas for improvement.

**How Solid Is the Base of Evidence in Current ACC/AHA Practice Guidelines?**

The considerable increase in the number of guideline recommendations across all guidelines through the current versions has not been uniformly supported by an increased volume of definitive evidence. In fact, while the overall proportion of recommendations labeled as class I has remained relatively constant, the greatest increase in guidelines recommendations has been among those subject to uncertainties, namely class II. Across guidelines, the median of recommendations in class II is currently 41%.

Level of evidence provides the link between recommendations and evidence base. Although there is significant variation among individual guidelines in the use of levels of evidence, the most solid evidence in guidelines is now used to develop performance measures, which are used, in turn, to judge the quality of practice, often in the context of differential payment or public reporting.
### Table 1. Change in the Number of Recommendations and Distribution Across Classes of Recommendation Between First Guideline Version and Current Version

<table>
<thead>
<tr>
<th>Guidelines by Year of Publication</th>
<th>Disease guidelines</th>
<th>Perioperative evaluation</th>
<th>Interventional guidelines</th>
<th>Diagnostic guidelines</th>
<th>Summary of all guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class I</td>
<td>Class II</td>
<td>Class III</td>
<td>No./Total (%)</td>
<td>Class I</td>
</tr>
<tr>
<td></td>
<td>No./Total (%)</td>
<td>Change, %</td>
<td>No./Total (%)</td>
<td>Change, %</td>
<td>No./Total (%)</td>
</tr>
<tr>
<td>Disease guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td></td>
<td></td>
<td>46/95 (48.4)</td>
<td>−23.7</td>
<td>38/95 (40)</td>
</tr>
<tr>
<td>2001</td>
<td>46/95 (48.4)</td>
<td>−23.7</td>
<td>38/95 (40)</td>
<td>−23.7</td>
<td>11/95 (11.6)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>73/127 (57.5)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>1995</td>
<td>73/127 (57.5)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>2005</td>
<td>66/129 (51.2)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>73/127 (57.5)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>1995</td>
<td>73/127 (57.5)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>2005</td>
<td>66/129 (51.2)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>73/127 (57.5)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>1995</td>
<td>73/127 (57.5)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>2005</td>
<td>66/129 (51.2)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>73/127 (57.5)</td>
<td>−11.0</td>
<td>44/129 (34.1)</td>
<td>−10.2</td>
<td>19/129 (14.7)</td>
</tr>
</tbody>
</table>

| Change in No. of recommendations   | +315               | +195                    | +63                      |
| Change in distribution across classes, median (IQR), % | 0.2 (−9.7 to 5.7) | 15.3 (6.9 to 25.4) | −16.4 (−28.4 to −2.2) |

**Abbreviations:** CABG, coronary artery bypass graft; IQR, interquartile range; PCI, percutaneous coronary intervention.

**a** Class I: conditions for which there is evidence and/or general agreement that a given procedure or treatment is useful and effective. Class II: conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment. Class III: conditions for which there is evidence and/or general agreement that the procedure or treatment is not useful/protective and in some cases may be harmful.
available evidence supporting recommendations, the median of level of evidence A recommendations is only 11% across current guidelines, whereas the most common grade assigned is level of evidence C, indicating little to no objective empirical evidence for the recommended action. The continued paucity of adequate evidence from randomized clinical trials is made most obvious by individual guidelines such as valvular heart disease, which has only 1 level of evidence A recommendation and yet has 71% level C recommendations. Thus, expert opinion remains a dominant driver of clinical practice, particularly in certain topic areas, highlighting the need for clinical research in these fields. Interestingly, our findings are reflective of a specialty—cardiology—that has a large pool of research to draw on for its care recommendations. Guidelines in other medical areas in which large clinical trials are performed less frequently may have an even weaker evidence-based foundation.

**Implications for Guideline Writing**

The current format of the ACC/AHA practice guidelines aims to provide recommendations to a broad set of possible decision points for each disease or condition. But whether guidelines should result from knowledge only and should not contain recommendations such as those in class II or level of evidence C is a matter of debate. Another organization, the US Preventive Services Task Force, has a different policy in guideline writing that avoids issuing recommendations that are not supported by evidence.

The main argument in favor of comprehensive documents is that patient care needs to be delivered and decisions made even in situations that have not been the subject of large randomized clinical trials. Physicians may need more guidance particularly in making decisions when extensive evidence is lacking. Alternatively, one might ar-

---

**Table 2. Distribution of Class of Recommendation and Level of Evidence in Current Guidelines**

<table>
<thead>
<tr>
<th>Guidelines</th>
<th>Year</th>
<th>Class of Recommendations</th>
<th>Level of Evidence</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>Disease guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>2006</td>
<td>41/111 (36.9)</td>
<td>55/111 (49.5)</td>
<td>15/111 (13.5)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>2005</td>
<td>66/112 (51.2)</td>
<td>44/112 (4.1)</td>
<td>19/129 (14.7)</td>
</tr>
<tr>
<td>Peripheral artery disease</td>
<td>2005</td>
<td>147/237 (62.6)</td>
<td>68/237 (28.1)</td>
<td>22/237 (9.4)</td>
</tr>
<tr>
<td>STEMI</td>
<td>2004</td>
<td>248/422 (58.8)</td>
<td>123/422 (29.1)</td>
<td>51/422 (12.1)</td>
</tr>
<tr>
<td>Perioperative evaluation</td>
<td>2007</td>
<td>13/50 (26.0)</td>
<td>27/50 (54.0)</td>
<td>10/50 (20.0)</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>2006</td>
<td>38/48 (79.2)</td>
<td>10/48 (20.8)</td>
<td>0/48</td>
</tr>
<tr>
<td>Stable angina</td>
<td>2002</td>
<td>78/235 (33.2)</td>
<td>98/235 (41.7)</td>
<td>28/235 (12.0)</td>
</tr>
<tr>
<td>Supraventricular arrhythmias</td>
<td>2003</td>
<td>61/147 (41.5)</td>
<td>77/147 (52.4)</td>
<td>9/147 (6.1)</td>
</tr>
<tr>
<td>Ventricular arrhythmias and sudden cardiac death</td>
<td>2006</td>
<td>103/217 (47.5)</td>
<td>100/217 (46.1)</td>
<td>14/217 (6.5)</td>
</tr>
</tbody>
</table>

**Interventional guidelines**

| PCI | 2006 | 119/136 (87.3) | 69/136 (50.7) | 28/136 (20.6) | 15/136 (11.0) | 56/136 (41.2) | 65/136 (47.8) | 0/136 |
| CABG | 2004 | 38/122 (31.1) | 50/122 (41.0) | 34/122 (27.9) | 6/122 (4.9) | 45/122 (36.9) | 71/122 (58.2) | 0/122 |

**Diagnostic guidelines**

| Exercise testing | 2003 | 190/333 (57.1) | 83/333 (24.9) | 60/333 (18.0) | 0/333 | 333/333 (100.0) |
| Echocardiography | 2003 | 36/48 (75.0) | 43/48 (93.7) | 18/48 (37.5) | 24/48 (50.0) | 17/48 (35.4) |
| Radionuclide imaging | 2003 | 42/80 (52.5) | 12/80 (15.0) | 24/80 (30.0) | 24/80 (30.0) | 24/80 (30.0) |

**Summary of all guidelines, median (IQR), %**

| Exercise testing | 2006 | 33/121 (27.4) | 29/121 (24.0) | 9/121 (7.4) | 9/121 (7.4) | 33/121 (27.4) |
| Echocardiography | 2003 | 36/48 (75.0) | 43/48 (93.7) | 18/48 (37.5) | 24/48 (50.0) | 17/48 (35.4) |
| Radionuclide imaging | 2003 | 42/80 (52.5) | 12/80 (15.0) | 24/80 (30.0) | 24/80 (30.0) | 24/80 (30.0) |

**Abbreviations:** CABG, coronary artery bypass graft; IQR, interquartile range; PCI, percutaneous coronary intervention; STEMI, ST-elevation myocardial infarction.

aClass I: conditions for which there is evidence and/or general agreement that a given procedure or treatment is useful and effective. Class II: conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment. Class III: conditions for which there is evidence and/or general agreement that the procedure or treatment is not useful/effective and in some cases may be harmful.

bLevel of evidence A: recommendation based on evidence from multiple randomized trials or meta-analyses. Level of evidence B: recommendation based on evidence from a single randomized trial or nonrandomized studies. Level of evidence C: recommendation based on expert opinion, case studies, or standard of care. Summary statistics include only guidelines reporting level of evidence.

©2009 American Medical Association. All rights reserved.

(Reprinted) JAMA, February 25, 2009—Vol 301, No. 8 835

Downloaded From: http://jama.jamanetwork.com/pdfaccess.ashx?url=/data/journals/jama/4453/ on 06/16/2017
EVIDENCE UNDERLYING THE ACC/AHA CLINICAL PRACTICE GUIDELINES

Interventional guidelines

Disease guidelines

Summary of all guidelines, JAMA, 836

Abbreviations: CABG, coronary artery bypass graft; IQR, interquartile range; PCI, percutaneous coronary intervention; STEMI, ST-elevation myocardial infarction.

The increase in number of recommendations included in the ACC/AHA guidelines is likely due to greater complexity of patient management decisions. The result has been longer documents. Recommendations with an absence of supporting evidence often require elaboration in the text to explain their rationale, which may be as extensive as the paragraphs reviewing the results of various clinical trials. Extensive documents including a large proportion of uncertain or non—evidence-based recommendations may make it increasingly difficult, when referring to a guideline, to locate the most important and/or evidence-based information relevant to an individual patient. Thus, they may reduce the implementation of evidence-based recommendations because the length of the documents may interfere with prompt access to guideline information.

Table 3. Distribution of Levels of Evidence Across Classes of Recommendation

<table>
<thead>
<tr>
<th>Class of Recommendation-Level of Evidence, No. (%)a</th>
<th>Guidelines</th>
<th>Year</th>
<th>I-A</th>
<th>I-B</th>
<th>I-C</th>
<th>II-A</th>
<th>II-B</th>
<th>II-C</th>
<th>III-A</th>
<th>III-B</th>
<th>III-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease guidelines: Atrial fibrillation (n = 111)</td>
<td>2006</td>
<td>8 (7.2)</td>
<td>12 (10.8)</td>
<td>21 (18.9)</td>
<td>3 (2.7)</td>
<td>18 (16.2)</td>
<td>34 (30.6)</td>
<td>2 (1.8)</td>
<td>3 (2.7)</td>
<td>10 (9.0)</td>
<td></td>
</tr>
<tr>
<td>Heart failure (n = 129)</td>
<td>2005</td>
<td>20 (15.5)</td>
<td>13 (10.1)</td>
<td>33 (25.6)</td>
<td>9 (7.0)</td>
<td>11 (8.5)</td>
<td>24 (18.6)</td>
<td>5 (3.9)</td>
<td>1 (0.8)</td>
<td>13 (10.1)</td>
<td></td>
</tr>
<tr>
<td>Perioperative evaluation (n = 50)</td>
<td>2007</td>
<td>5 (10.0)</td>
<td>5 (10.0)</td>
<td>3 (6.0)</td>
<td>0</td>
<td>19 (38.0)</td>
<td>8 (16.0)</td>
<td>1 (2.0)</td>
<td>4 (8.0)</td>
<td>5 (10.0)</td>
<td></td>
</tr>
<tr>
<td>Peripheral artery disease (n = 237)</td>
<td>2005</td>
<td>29 (12.2)</td>
<td>88 (37.1)</td>
<td>30 (12.7)</td>
<td>4 (1.7)</td>
<td>43 (18.1)</td>
<td>19 (8.0)</td>
<td>3 (1.3)</td>
<td>9 (3.8)</td>
<td>10 (4.2)</td>
<td></td>
</tr>
<tr>
<td>STEMI (n = 422)</td>
<td>2004</td>
<td>45 (10.7)</td>
<td>95 (22.5)</td>
<td>108 (25.6)</td>
<td>5 (1.2)</td>
<td>50 (11.8)</td>
<td>68 (16.1)</td>
<td>7 (1.7)</td>
<td>21 (5.0)</td>
<td>23 (5.5)</td>
<td></td>
</tr>
<tr>
<td>Secondary prevention (n = 48)</td>
<td>2006</td>
<td>10 (20.8)</td>
<td>26 (54.2)</td>
<td>2 (4.2)</td>
<td>1 (2.1)</td>
<td>7 (14.6)</td>
<td>2 (4.2)</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable angina (n = 235)</td>
<td>2002</td>
<td>12 (5.1)</td>
<td>34 (14.5)</td>
<td>52 (21.9)</td>
<td>1 (0.4)</td>
<td>39 (16.6)</td>
<td>58 (24.7)</td>
<td>2 (0.9)</td>
<td>19 (8.1)</td>
<td>38 (16.2)</td>
<td></td>
</tr>
<tr>
<td>Supraventricular arrhythmias (n = 147)</td>
<td>2003</td>
<td>5 (3.4)</td>
<td>32 (21.8)</td>
<td>24 (16.3)</td>
<td>4 (2.7)</td>
<td>20 (13.6)</td>
<td>53 (36.0)</td>
<td>0</td>
<td>3 (2.0)</td>
<td>6 (4.1)</td>
<td></td>
</tr>
<tr>
<td>Unstable angina (n = 298)</td>
<td>2007</td>
<td>57 (19.1)</td>
<td>82 (27.5)</td>
<td>47 (15.8)</td>
<td>1 (0.3)</td>
<td>32 (10.8)</td>
<td>25 (8.4)</td>
<td>8 (2.7)</td>
<td>5 (1.7)</td>
<td>16 (5.4)</td>
<td></td>
</tr>
<tr>
<td>Ventricular arrhythmias and sudden cardiac death (n = 217)</td>
<td>2006</td>
<td>19 (8.8)</td>
<td>32 (14.7)</td>
<td>52 (24.0)</td>
<td>1 (0.5)</td>
<td>35 (16.1)</td>
<td>64 (29.5)</td>
<td>1 (0.5)</td>
<td>2 (0.9)</td>
<td>11 (5.1)</td>
<td></td>
</tr>
<tr>
<td>Summary of disease guidelines, median (IQR), %</td>
<td>May 9, 2009</td>
<td>10.0</td>
<td>18.4</td>
<td>16.3</td>
<td>1.7</td>
<td>16.1</td>
<td>18.6</td>
<td>1.3</td>
<td>2.7</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>Interventional guidelines</td>
<td>PCI (n = 136)</td>
<td>2005</td>
<td>14 (10.3)</td>
<td>13 (18.2)</td>
<td>7 (5.1)</td>
<td>1 (0.7)</td>
<td>34 (25.0)</td>
<td>34 (25.0)</td>
<td>0</td>
<td>4 (2.9)</td>
<td>24 (17.6)</td>
</tr>
<tr>
<td>CABG (n = 84)</td>
<td>2004</td>
<td>12 (14.3)</td>
<td>25 (29.8)</td>
<td>2 (2.4)</td>
<td>4 (4.8)</td>
<td>19 (22.7)</td>
<td>11 (13.1)</td>
<td>0</td>
<td>7 (8.3)</td>
<td>4 (4.8)</td>
<td></td>
</tr>
<tr>
<td>Pacemaker (n = 122)</td>
<td>2008</td>
<td>4 (3.3)</td>
<td>15 (12.3)</td>
<td>19 (15.6)</td>
<td>1 (0.8)</td>
<td>18 (14.8)</td>
<td>31 (25.4)</td>
<td>1 (0.8)</td>
<td>12 (9.8)</td>
<td>21 (17.2)</td>
<td></td>
</tr>
<tr>
<td>Summary of interventional guidelines, median (IQR), %</td>
<td>May 9, 2009</td>
<td>10.3</td>
<td>13.2</td>
<td>5.1</td>
<td>0.8</td>
<td>22.6</td>
<td>25.0</td>
<td>0</td>
<td>8.3</td>
<td>17.2</td>
<td></td>
</tr>
<tr>
<td>Diagnostic guidelines</td>
<td>Exercise testing (n = 71)</td>
<td>2002</td>
<td>0</td>
<td>2 (2.8)</td>
<td>1 (1.4)</td>
<td>0</td>
<td>1 (1.4)</td>
<td>2 (2.8)</td>
<td>0</td>
<td>0</td>
<td>3 (4.2)</td>
</tr>
<tr>
<td>Radionuclide imaging (n = 84)</td>
<td>2003</td>
<td>4 (4.8)</td>
<td>28 (33.3)</td>
<td>4 (4.8)</td>
<td>0</td>
<td>29 (34.5)</td>
<td>14 (16.7)</td>
<td>0</td>
<td>1 (1.2)</td>
<td>4 (4.8)</td>
<td></td>
</tr>
<tr>
<td>Summary of diagnostic guidelines, median (IQR), %</td>
<td>May 9, 2009</td>
<td>2.4</td>
<td>18.1</td>
<td>3.1</td>
<td>18.0</td>
<td>3.1</td>
<td>2.4</td>
<td>18.1</td>
<td>3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary of all guidelines, median (IQR), %</td>
<td>May 9, 2009</td>
<td>9.4</td>
<td>16.6</td>
<td>14.6</td>
<td>1.0</td>
<td>16.2</td>
<td>17.6</td>
<td>0.6</td>
<td>2.8</td>
<td>5.4</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ACC, American College of Cardiology; AHA, American Heart Association; PCI, percutaneous coronary intervention; STEMI, ST-elevation myocardial infarction.

aClass I: conditions for which there is evidence and/or general agreement that a given procedure or treatment is useful and effective. Class II: conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment. Class III: conditions for which there is evidence and/or general agreement that the procedure or treatment is not useful/effective and in some cases may be harmful. Level of evidence A: recommendation based on evidence from multiple randomized trials or meta-analyses. Level of evidence B: recommendation based on evidence from a single randomized trial or nonrandomized studies. Level of evidence C: recommendation based on expert opinion, case studies or standard of care. Summary statistics include only guidelines reporting level of evidence.
To address this problem, the ACC/AHA Task Force on Practice Guidelines has adopted a standard format by placing the recommendations in bolded format at the beginning of the discussion supporting the recommendations and by adding tables and publishing an executive summary. Other guideline writing committees may partially address these issues while still using the current format by separating the summary of clinical trial data from the interpretation of the trial data and the rationale used to justify the recommendations. The data of the present study should serve as a basis to evaluate whether the current format of the guidelines should be altered to achieve a better focus on recommendations supported by objective evidence.

The analysis presented in this article does not address cardiovascular guidelines released by other major societies, such as the European Society of Cardiology (ESC). However, it is likely that the ESC guidelines and other cardiovascular guidelines face similar challenges, particularly concerning the evidence base at the foundation of the recommendations. When guidelines address topics with limited or conflicting information, it would not be unexpected to find variation on specific recommendations between documents released by different societies. Indeed, differences between the recommendations of the ACC/AHA and ESC guidelines have been noted in recent guidelines. 51-72

The presence of a large proportion of recommendations with no supporting data from randomized clinical trials requires careful judgment by guideline authors. In such circumstances, the potential for authors’ conflicts of interest, real or perceived, may be important. Recommendations based only on expert opinion may be prone to conflicts of interest because, just as clinical trialists have conflicts of interests, expert clinicians are also those who are likely to receive honoraria, speakers bureau, consulting fees, or research support from industry. 73,74 It is difficult to quantify the effect of conflicts of interest in a guideline writing process, but this was not the subject of the present study. Certainly, real or not, the perception among guideline readers that financial ties may introduce significant bias in guideline recommendations has been noted in 1 report. 75 A commonly adopted method to deal with conflicts of interest is adding disclosures, although it is not clear what effects such disclosures might have. Disclosing a conflict may make the authors wary about recommending products in which they may have an interest. However, it may also act in the opposite direction by increasing the authors’ confidence in recommending such products once a conflict has been disclosed.

Major guideline-releasing organizations have recognized the importance of having a rigorous policy regarding conflicts of interest; such policies manage and balance potential conflicts rather than eliminating them. The ACC/AHA’s code regulating potential conflicts of interest requires the collection and publication of relationships with industry by guideline-writing groups as well as peer reviewers. Relationships are orally disclosed at every meeting, votes are recorded for all recommendations, and members with significant conflicts abstain from voting, although they can participate in the discussion. In addition, the ACC/AHA task force now requires that 30 to 50% of writing group members have no conflicts of interest, and the guideline writing group must be chaired by someone with no conflicts of interest. Finally, there is no industry funding for guideline development, although the ACC and AHA do receive industry support for distribution of guideline derivative products such as pocket guides.

**Implications for Research**

The findings of this analysis indicate that the current system generating research is inadequate to satisfy the information needs of caregivers and patients in determining benefits and risks of drugs, devices, and procedures. The clinical research system in the United States has been described as a fragmented “nonsystem,” with a lack of common goals, vision, and collaboration. 76 In addition, the current clinical research agenda in the United States is strongly influenced by industry’s natural drive to introduce new products. 76 There is limited sponsorship of trials to address questions of comparative effectiveness or routine clinical practice. The problem of how to generate funding for research addressing practical clinical questions that do not involve a marketable product is currently unresolved. Parties with an interest include patients, health care practitioners, and payers. Frequently, patient advocacy groups are effective in raising funds or influencing congressional funding in this regard. There are examples of public-private partnerships addressing practical questions about technology, such as the Center for Medical Technology Policy. 77 Patients also may have an interest in funding research on practical clinical questions that have direct relevance to reimbursement decisions. 78 Some practical clinical questions have been funded by government agencies such as the Veterans’ Administration and the National Institutes of Health, but the proportion of these budgets available for practical clinical trials appears to be limited. 79,80 A special agency to foster studies of comparative effectiveness is also under consideration. 81 The relative paucity of funding for practical clinical questions and comparative effectiveness studies deserves a prominent place in policy discussions.

In addition to the paucity of funding, the marked inefficiency of the current research system—resulting in high costs and extended duration of many clinical studies—reduces the number of questions that can be addressed. The prohibitive costs and time also may discourage researchers from developing and implementing ideas for investigator-initiated research. In this setting, even the availability of increased funding may not guarantee major achievements in research, as suggested by the
fact that despite the doubling of expenditure in research and development by industry, productivity in terms of new US Food and Drug Administration (FDA) approvals has progressively decreased in the last decade. Improving the research system will require active collaboration among all of the interested parties—ie, academic, professional, and government organizations and industry. One such collaboration initiated under the FDA’s Critical Path Program is the Clinical Trials Transformation Initiative. The mission of this initiative is to identify practices that, through broad adoption, will increase the quality and efficiency of clinical trials.

Key research stakeholders should collaborate in generating a prioritized list of research topics. The ACC/AHA guideline writing group is now assisting in addressing this need by recommending an agenda of research priorities based on important questions that arise in the writing process about where evidence is needed.

A separate issue is the heavy focus of industry on efficacy studies in restricted patient populations necessary to gain FDA approval. Although initially important to document a drug’s efficacy without the confounding of multiple disease states and interacting medications, it is also necessary to study new drugs and devices in the broader population of patients who will receive them in actual practice. These latter studies could be initiated while a drug application is waiting for review by regulatory authorities (phase 3b) or shortly after market approval. These more practical clinical trials would typically address the questions that physicians and third-party payers would have in seeking the proper application of these new treatments in practice.

Study Limitations
Our analysis does not account for potential changes over time in the aims of guidelines writing committees, which may have influenced the number of recommendations and the distribution across classes. Moreover, in 1990, the class II level of recommendation was expanded to classes IIa and IIb. With this definition change, standards and thresholds to determine class of recommendation may not have remained constant. Our analysis was designed to evaluate comprehensive guideline documents; therefore, the data included in this article do not reflect recommendations in focused guideline updates that have been recently released for some documents but not yet incorporated into the comprehensive documents (eg, stable angina, percutaneous coronary intervention, STElevation myocardial infarction). These focused updates are driven primarily by results of recent randomized clinical trials but address only a limited number of issues. The change in levels of evidence could be evaluated only in a limited number of guidelines, which are those that are updated more frequently, and may not be representative of the entire cohort of guidelines.

It was beyond the scope of this article to analyze and compare cardiovascular guidelines released by other societies and noncardiovascular guidelines. Finally, this article only addressed ACC/AHA practice guidelines, and the results cannot be directly applied to other types of documents, such as “appropriateness criteria.”

CONCLUSION
Our finding that a large proportion of recommendations in ACC/AHA guidelines are based on lower levels of evidence or expert opinion highlights deficiencies in the sources of definitive data available for the generation of cardiovascular guidelines. To remedy this problem, the medical research community needs to streamline clinical trials, focus on areas of deficient evidence, and expand funding for clinical research. In addition, the process of developing guidelines needs to be improved with information about the impact that recommendations based on lower levels of evidence has on clinical practice. Finally, clinicians need to exercise caution when considering recommendations not supported by solid evidence.

Author Contributions: Dr Tricoci had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Tricoci, Allen, Kramer, Califf, Smith.

Acquisition of data: Tricoci, Allen.

Analysis and interpretation of data: Tricoci, Allen, Kramer, Califf, Smith.

Drafting of the manuscript: Tricoci, Allen, Kramer.

Critical revision of the manuscript for important intellectual content: Tricoci, Allen, Kramer, Califf, Smith.

Statistical analysis: Tricoci, Allen.

Obtained funding: Allen, Kramer.

Administrative, technical, or material support: Califf.

Study supervision: Tricoci, Califf, Smith.

Financial Disclosures: Dr Tricoci reports that as a faculty member of the Duke Clinical Research Institute, an academic clinical research organization, his salary is partially supported by research grants from Schering-Plough. Dr Kramer reports that as a faculty member of the Duke Clinical Research Institute, an academic clinical research organization, and as executive director of the Clinical Trials Transformation Initiative (CTTI), her salary is supported in part by CTTI membership fees paid by the following companies: Amgen Inc, Bayer, bioMerieux, Biotronik, Bristol-Myers Squibb, CR Bard Inc, Eli Lilly, Genentech, GlaxoSmithKline, Hoffman-La Roche Inc, Human Genome Sciences Inc, J&J Medical Devices, J&J Pharmaceutical Research and Development, Novartis, Pfizer, St Jude Medical, The Medicines Company, and Wright Medical. Dr Califf reports that he receives research funding from Novartis Pharmaceutical and Schering-Plough; is a member of the speakers bureaus for Heart.org, Kowa Research Institute, and Novartis Pharmaceutical, and consults for ABC, Amgen, Bayer, Boehringer Ingelheim, Boston Scientific, GSK, Heart.org, Kowa Research Institute, Medtronic, Nitrox LLC, Novartis Pharmaceutical, Roche, Sanofi-Aventis, Schering-Plough, SCiLS, Targacept, University of Florida, and Vvus; and owns equity in Nitrox LLC. No other financial disclosures were reported.

Funding/Support: This project was supported by grant U18HS010548 from the Agency for Healthcare Research and Quality. Dr Smith’s efforts in this work were supported by a Distinguished Scholarship from the Rose Azus Cardiac Research Education Fund, Sharp Hospital Foundation, San Diego, California.

Role of the Sponsors: No sponsor of this research participated in the design and conduct of the study, in the collection, analysis, and interpretation of the data, or in the preparation, review, or approval of the manuscript.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

REFERENCES
EVIDENCE UNDERLYING THE ACC/AHA CLINICAL PRACTICE GUIDELINES


EVIDENCE UNDERLYING THE ACC/AHA CLINICAL PRACTICE GUIDELINES


©2009 American Medical Association. All rights reserved.

(Reprinted) JAMA, February 25, 2009—Vol 301, No 8 841