Measurement of adherence to clinical standards has become increasingly important to the practice of emergency medicine (EM). In recent years, along with a proliferation of evidence-based practice guidelines and performance measures, there has been a movement to incorporate measurement into reimbursement strategies, many of which affect EM practice. On behalf of the Society for Academic Emergency Medicine (SAEM) Guidelines Committee 2009–2010, the purposes of this document are to: 1) differentiate the processes of guideline and performance measure development, 2) describe how performance measures are currently and will be used in pay-for-performance initiatives, and 3) discuss opportunities for SAEM to affect future guideline and performance measurement development for emergency care. Specific recommendations include that SAEM should: 1) develop programs to sponsor guideline and quality measurement research; 2) increase participation in the process of guideline and quality measure development, endorsement, and maintenance; 3) increase collaboration with other EM organizations to review performance measures proposed by organizations outside of EM that affect emergency medical care; and 4) answer calls for participation in the selection and implementation of performance measures through The Joint Commission and the Centers for Medicare and Medicaid Services (CMS).
pneumonia, acute myocardial infarction, and asthma.\textsuperscript{3} Data are also reported on patient satisfaction at the facility level; however, the satisfaction measures do not specifically identify the emergency care experience. At the level of the emergency medical provider, performance measures can be voluntarily reported through the CMS Physician Quality Reporting Initiative (PQRI).\textsuperscript{4}

Medical organizations, such as specialty societies, hospitals, and government bodies, have developed and distributed guidelines for many years. Performance measure development, however, is newer, and its subsequent deployment for emergency care has been controversial.\textsuperscript{5} Greater scrutiny has been placed on performance measures, likely because they are more prescriptive, are publicly reported, apply to individual physicians and organizations, and can have financial consequences. This scrutiny has raised questions over the evidence base for the measures, whether improving compliance with particular measures actually improves outcomes, and whether disease-specific measures are appropriate given the heterogeneity of the practice of emergency care, which encompasses mostly diseases that are not reviewed by a reporting program.\textsuperscript{6,7}

The purposes of this article are to 1) differentiate clearly the processes of guideline and performance measure development, 2) describe how performance measures are currently and will be used in pay-for-performance initiatives, and 3) discuss opportunities for the Society for Academic Emergency Medicine (SAEM) to affect future guideline and performance measurement development for emergency care. This article has been written on behalf of the 2009–2010 SAEM Guidelines Committee.

### FINDINGS

#### The Development of Clinical Practice Guidelines for Emergency Care

The Institute of Medicine (IOM) has defined clinical practice guidelines as “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.”\textsuperscript{8} Clinical practice guidelines are generally developed by expert panels convened by medical specialty associations (including relevant professional societies and public or private organizations); governmental agencies at the federal, state, or local level; and health care organizations or plans.

Guideline developers integrate existing evidence with expert opinion to determine treatment recommendations. There are several proposed systems to stratify the quality of evidence and specific recommendations. One such system, proposed by the U.S. Preventive Services Task Force, ranks the level of evidence for the overall clinical practice guidelines and then classifies each recommendation based on a risk–benefit analysis and the level of evidence used (Tables 1 and 2).\textsuperscript{9} Most expert panels will approach the development process by dividing topics among committee members who present analyses of their findings and conclusions. To help inform the guideline development process, there are several resources available, including the “Guidelines for Guidelines” series by World Health Organization Advisory Committee;\textsuperscript{10} the Grading of Recommendations Assessment, Development and Evaluation (GRADE);\textsuperscript{11} and Appraisal of Guidelines Research and Evaluation (AGREE).\textsuperscript{12}

### Table 1

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>Evidence obtained from at least one properly designed randomized controlled trial.</td>
</tr>
<tr>
<td>Level II-1</td>
<td>Evidence obtained from well-designed controlled trials without randomization.</td>
</tr>
<tr>
<td>Level II-2</td>
<td>Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.</td>
</tr>
<tr>
<td>Level II-3</td>
<td>Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled trials might also be regarded as this type of evidence.</td>
</tr>
<tr>
<td>Level III</td>
<td>Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level A</td>
<td>Good scientific evidence suggests that the benefits of the clinical service substantially outweigh the potential risks. Clinicians should discuss the service with eligible patients.</td>
</tr>
<tr>
<td>Level B</td>
<td>At least fair scientific evidence suggests that the benefits of the clinical service outweigh the potential risks. Clinicians should discuss the service with eligible patients.</td>
</tr>
<tr>
<td>Level C</td>
<td>At least fair scientific evidence suggests that there are benefits provided by the clinical service, but the balance between benefits and risks is too close for making general recommendations. Clinicians need not offer it unless there are individual considerations.</td>
</tr>
<tr>
<td>Level D</td>
<td>At least fair scientific evidence suggests that the risks of the clinical service outweigh potential benefits. Clinicians should not routinely offer the service to asymptomatic patients.</td>
</tr>
<tr>
<td>Level I</td>
<td>Scientific evidence is lacking, of poor quality, or conflicting, such that the risk versus benefit balance cannot be assessed. Clinicians should help patients understand the uncertainty surrounding the clinical service.</td>
</tr>
</tbody>
</table>
The GRADE working group provides a process for grading evidence and providing specific recommendations—either “strong” or “weak”—based on the evidence. GRADE considers many elements of studies including: 1) uncertainty in the estimates of likely benefit, and likely risk, inconvenience, and costs; 2) magnitude of treatment effects; 3) precision of estimate of treatment effect; 4) risks associated with therapy; 5) risk of target event; 6) costs; and 7) varying values. The AGREE methodology can be used by diverse stakeholders and provides a general instrument to evaluate new guidelines, existing guidelines, and updates of existing guidelines. It consists of 23 items in six domains, including: 1) scope and purpose, 2) stakeholder involvement, 3) rigor of development, 4) clarity and presentation, 5) applicability, and 6) editorial independence. Each guideline can be assessed by two to four appraisers who rate each of the 23 items on a four-point Likert scale, which are summed for each domain. Finally, an overall recommendation can be assigned by each appraiser.

During the guideline development process, controversial topics are discussed until consensus is achieved, if possible. One method for achieving consensus is the modified Delphi technique, originally developed by the RAND Corporation. The Delphi method is defined by participant anonymity, iterative polling rounds with feedback, and statistical analysis of group results. Once general agreement is obtained, the writing committee cochairs incorporate separate analyses into a single document, which is redistributed to committee members to review and update. The final draft of a clinical guideline is ultimately submitted to the sponsoring organization for approval.

Many clinical guidelines for emergency care have been drafted by the American College of Emergency Physicians. However, other specialty societies, federal agencies, and international guideline development organizations have produced a large volume of clinical guidelines for conditions commonly encountered in the emergency department (ED) and thus affect diagnostic and therapeutic decisions made by emergency practitioners. Historically, organizations outside of EM have not routinely sought participation from EM. This trend is changing, likely as a result of the maturation of EM as an academic specialty and involvement of emergency physicians in other professional societies related to their areas of research. Increasingly, SAEM is asked to review and endorse guidelines from other specialty organizations and, more recently, to appoint representatives to committees responsible for the creation of specialty-endorsed guidelines.

As SAEM moves forward to consider how to participate more closely with this process, it may learn from other organizations. One example of a highly successful collaboration to develop guidelines exists between the American Heart Association (AHA) and the American College of Cardiology (ACC). Together, they publish and frequently update the *Methodology Manual for AHA Guideline Writing Committees*, which may serve as a model for guideline development within other specialties like EM, who seek an expanded role in guideline development. Finally, there is a large literature on guideline implementation available, which is reviewed by the Cochrane Effective Practice and Organisation of Care Group.

**Differentiating a Performance Measure From a Guideline**

Performance measurement (also known as quality measurement) attempts to quantify the quality of care that health care providers or organizations deliver, with the goal of comparing and improving it. The basic principle is: “If you can measure it, you can manage it.” A performance measure typically assesses whether the health care provided meets some prespecified standard. Other measures may represent data as rankings or percentiles compared with other hospitals or providers, such that a constant fraction of the individual or group being measured is consistently not meeting the standard. For example, in patient satisfaction measures, the goal is often to compare rankings rather than to cross a threshold standard.

Performance measures are selected components of care for a specific health care process and are usually derived from guidelines or other community-defined standards. Measures can address any of the IOM-defined domains of health care quality and target anything from individual provider behavior to systems of care. While guidelines typically attempt to provide evidence-based recommendations for the entire range of diagnostic and therapeutic issues in treating a disease or condition across populations, performance measures prescribe specific, measurable elements of care appropriate for one patient population. Performance measures are usually developed from guidelines, yet not all guidelines are appropriate for performance measurement.

Performance measures can be grouped into categories of structure, process, and outcome. Structural measures apply to the characteristics of the care environment. Examples include the presence of an electronic health record, the nurse-to-patient ratio, and the accreditation or licensing status of the facility and providers (e.g., whether a hospital is certified as a Joint Commission Primary Stroke Center).

Measures of process evaluate whether specific care was delivered. Examples of process measures include administration of medications or measurement of vital signs deemed appropriate for a certain condition (e.g., aspirin given for acute myocardial infarction, oxygen saturation measurement for pneumonia). One issue with the use of process measures in the ED is that they are influenced by system properties that can be partially out of the control of providers, such as the relationship between ED crowding and time to antibiotics in pneumonia. These issues can create tension between stakeholders when considering which types of measures to develop. National bodies initially adopted process measures, but the recent trend is to adopt outcome measures, such as mortality and patient satisfaction.

Outcome measures evaluate the affect of care and include readmission, mortality, or patient satisfaction. Policymakers and patients have shown greater interest in outcome measures than in either structure or
process measures.\textsuperscript{29} However, outcome measures are challenging because methods of risk adjustment are imperfect, and attributing a patient’s outcome to whether or not a specific treatment was performed is difficult in general.\textsuperscript{30,31} Additionally, since admitted patients spend a relatively small fraction of their overall hospitalization in the ED, it is inaccurate to attribute patient outcomes to ED care alone. Using ED patient satisfaction scores as an outcome has been controversial because of poor response rates and their association with ED crowding, both of which are often out of the control of providers.\textsuperscript{32} As a result of their limitations, specific criteria have been proposed for publicly reported outcomes measures.\textsuperscript{33} However, to date, few outcomes measures have received the scrutiny to be approved for implementation at the national level and none specific for emergency care.\textsuperscript{34}

Performance Measure Development

During performance measure development, stakeholders assess several attributes of potential measures: clinical importance, scientific soundness, usability to clinicians, and feasibility of measurement.\textsuperscript{33,35} The final goal is the creation of a performance measure that addresses an area for which improvement is needed (i.e., there exists a quality gap), accurately reflects quality care, and that, if more closely adhered to, will ultimately lead to improved outcomes.

Performance measures have been developed by diverse organizations for varied purposes. EDs and hospitals develop performance measures for internal quality improvement programs. Specialty societies develop measures for submission to national measure development programs. Governmental bodies and third-party payers develop measures for public reporting or to guide payment policy. While a multitude of specialty societies and government bodies contribute to the development of guidelines, there are currently two primary organizations working on the creation performance measures for national implementation: the National Committee for Quality Assurance (NCQA)\textsuperscript{36,37} and the American Medical Association-Convened Physician Consortium for Performance Improvement (AMA-PCPI).\textsuperscript{38}

The NCQA is a not-for-profit organization dedicated to health care quality improvement. It was formed in the era of managed care with the aim of assessing the quality of health insurance plans available to employers and employees. The NCQA works with employers, policymakers, physicians, patients, and health plans to determine topics of importance, how to measure them, and how to promote improvement.\textsuperscript{36} Health plans can earn the NCQA seal of approval by reporting on performance in more than 40 areas. Additionally, the NCQA has developed “seals of recognition” for primary care and specialty group practices that complete targeted diseases specific quality improvement efforts.\textsuperscript{37}

The AMA-PCPI is composed of over 100 national medical specialty and state medical societies, the American Board of Medical Specialties and its member boards, experts in methodology and data collection, and representatives from the Agency for Healthcare Research and Quality (AHRQ), CMS, and other organizations and agencies involved in quality improvement and performance measurement. Their mission is to enhance “quality of care and patient safety through the development, testing, and maintenance of evidence-based clinical performance measures and measurement resources for physicians.”\textsuperscript{38} The AMA-PCPI initiates measure development internally and facilitates measure development for external organizations. Measures are developed in working groups and then presented to the AMA-PCPI voting members for feedback and final approval. AMA-PCPI performance measures have gained national recognition and are among those endorsed for use by Medicare.\textsuperscript{39} Currently, the AMA-PCPI and the NCQA collaborate on measure development efforts to share resources and reduce the overall cost of this process.

It may be evident from the proceeding discussion that developing high-quality clinical practice guidelines and performance measures is an expensive and time-consuming endeavor. Guidelines involve systematic literature review, several rounds of meetings and conferences, and a writing committee to draft the guideline manuscript. Measure development is similarly time-consuming and expensive. This work requires organizational and logistical support and often involves travel. Participant time is the most expensive resource, and participation is often without compensation by individuals who have an academic, professional, or business interest in the measures.

The Role of the National Quality Forum and the AHRQ in Performance Measure Endorsement

The National Quality Forum (NQF) is a not-for-profit organization that aims to improve the quality of health care for all Americans through setting national goals for performance improvement, endorsing standards for measuring and publicly reporting performance, and promoting the attainment of national goals through education and outreach programs.\textsuperscript{40} The NQF has been granted special status by Congress as the official health standards organization for health quality measurement in the United States.

The NQF has developed a formal Consensus Development Process (CDP) to evaluate and endorse performance measures.\textsuperscript{41} The CDP is designed to call for input actively and carefully consider the diverse interests of health care stakeholder groups. The NQF CDP involves nine steps, illustrated in Table 3. The entire process is posted on the NQF website, and NQF seeks public input throughout the course. Once a measure has been endorsed by the NQF, organizations such as CMS and The Joint Commission (TJC) are able to choose measures for implementation based on their own priorities. However, the lack of a formal process between measure endorsement by the NQF and ultimate implementation may represent an opportunity for improvement. In addition, the AHRQ has recently been granted similar status to vet developed measures prior to implementation by Medicaid.

Criteria for Evaluating and Maintaining Performance Measures and Guidelines

At the time of NQF endorsement, an organization such as TJC is designated as the measure steward and a
either reduce or increase the importance of a particular guideline, and guideline recommendations may change altogether. Similarly, performance measures should be regularly assessed for their continued value in improving the quality of care and for any unintended consequences. Ideally, failed performance measures should be rapidly repealed. However, certain performance measures, including the measurement of antibiotic timing in pneumonia, have remained in place despite consistent negative feedback from clinicians, researchers, and specialty societies.5-7

Technical Barriers in Performance Measure Development
The use of guidelines and performance measures has important limitations that SAEM should be cognizant of when participating in measurement efforts. As research and content experts, SAEM has a role to guard against the promulgation of inappropriate measures. The limited evidence base around much of emergency care should cause measure developers hesitation in proposing measures that will be implemented without formal testing. However, there is currently little funding for testing measures, so many go from development to implementation, leaving the testing to be worked out after the fact. In addition, groups that fund measures often begin a project with a goal of developing a certain number of measures for a topic area, rather than aiming to develop an appropriate number of measures that are well supported by clinical evidence. Several early EM workgroups produced several measures, many of which have not been integrated into formal performance measurement programs by outside groups.44,45

There is also a tension between developing and using measures for improvement compared to measures that are intended for judging, rewarding, or punishing. Measures for improvement need not have the same degree of scientific precision, such as risk adjustment between institutions, as improvement can be measured compared to prior performance. By contrast, measures that judge specific elements of quality should meet detailed specifications to ensure accuracy and fairness in comparisons between providers or institutions.33

Potential conflicts of interest are also important, as guidelines and measures can have significant monetary consequences, such as recommending one medicine or class of medicines or defining pay-for-performance schemes. Although organizations such as NQF attempt to address this by public disclosure and multistakeholder processes, it may not be sufficient. For example, a recent imaging efficiency project has representation from Siemens, the American College of Radiology, and other professional societies that have an important economic stake in the outcome of the process.46

Finally, there are important technical issues that limit the ability of even well-constructed measures to truly differentiate performance between institutions, such as sample size, clustering of patients within hospitals, patient severity, and hospital case mix. Although these issues are beginning to be addressed for hospital inpatient measures, there is little research into risk adjustment in emergency care.

Table 3
NQF CDP and Measure Evaluation Criteria

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NQF Measure Evaluation Criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Importance to measure and report: Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high-impact aspect of health care where there is variation in or overall poor performance. Candidate measures must be judged to be important to measure and report to be evaluated against the remaining criteria. If not important to measure and report, STOP.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Scientific acceptability of the measure properties: Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Usability: Extent to which intended audiences (e.g., consumers, purchasers, providers, policymakers) can understand the results of the measure and are likely to find them useful for decision-making.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Feasibility: Extent to which the required data are readily available and retrievable without undue burden and can be implemented for performance measurement.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>There are several steps in the ongoing cycle of measure maintenance:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>synthesized</td>
<td>• Annual report from measure steward</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Notification of expiring endorsement status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Completion of measure resubmission form by measure steward</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Member and public comment periods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Reevaluation of measure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ongoing feedback on NQF-endorsed measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Permission granted for republication.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>timeline is set for measure maintenance. Measure maintenance involves updating evidence about the measure’s attributes, a member and public comment period, and a reevaluation of the measure including voting on endorsement. Details of the NQF evaluation criteria and measure maintenance process are detailed in Table 3.54,142</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maintenance of NQF-Endorsed® Performance Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Maintenance of both performance measures and guidelines is a vital part of the process of ensuring that active measures and guidelines are kept current with the latest evidence. It has been estimated that the half-life of a guideline is 5 years.43 New evidence may either reduce or increase the importance of a particular
The Role of Pay-for-reporting and Pay-for-performance

There has been substantial interest from both private and public health care purchasers in the use of financial incentives to improve reporting and quality of care. Respectively, these are referred to as pay-for-reporting and pay-for-performance.

The CMS PQRI is an example of pay-for-reporting. Physicians who report performance for at least 80% of eligible Medicare patients on at least three performance measures are eligible to receive an incentive payment of 2% of their total charges for professional services to their Medicare patients. Of the 186 PQRI-eligible measures for 2009, 10 apply to EM.

The CMS provides pay-for-reporting incentives to hospitals through the National Hospital Quality Measures (so-called “core measures”). Although participation in core measures assessment is voluntary, hospitals are given their full annual payment update only if they report performance on these measures. While the financial impact is directly to the hospital, the pressure to measure and comply extends to the medical staff and providers.

Pay-for-performance is different from pay-for-reporting. Pay-for-reporting results in payments if the organization or physician provides documentation of measure performance, while pay-for-performance provides financial incentives for complying with standards of care that are considered quality targets by payers. Health plans see pay-for-performance as a means to align physician payment with quality of care delivered and narrow the gap between clinical practice and evidence-based guidelines.

There are many different types of pay-for-performance initiatives, with variation in the measures, the targets, and how the rewards are structured. Private payers, such as Blue Cross, were the first to implement pay-for-performance programs. Currently, many payers are increasingly “withholding payments” to hospitals for selected conditions unless hospitals meet prespecified performance target thresholds. Perhaps as a sign of things to come, CMS is currently providing pay-for-performance incentives to top-performing hospitals through the Premier Hospital Quality Incentive Demonstration project.

Pay-for-performance has become increasingly common, although it remains controversial. Many studies have shown that quality has improved in association with pay-for-performance, yet the trials have not been designed to determine whether the payment is responsible for improvement or if other aspects of quality improvement programs—such as common goals, institutional support, data sharing, or public reporting—change behavior. Many clinicians are concerned that pay-for-performance will use performance measures to drive cost control rather than improve quality improvement. In addition, several papers in peer-reviewed journals have actively challenged the validity of using pay-for-performance altogether.

There may also be unintended consequences, such as the overutilization of antibiotics when specific timing standards are required in pneumonia. While pay-for-performance criticisms may be valid and unintended consequences will likely ensue, pay-for-performance will likely play an increasing role in the reimbursement structure of U.S. medicine. In fact, these critiques further highlight the importance of increasing EM’s participation in the development phase of ED-based pay-for-performance initiatives.

The Role of Research in the Development and Refinement of Performance Measures

There are several steps in the process of performance measure development where emergency care researchers can influence development, implementation, and maintenance. The most fundamental method is by conducting well-designed original research to design and test measures. In fact, researchers and quality improvement practitioners can develop and test their own measures. When there is a call for measures from a national body, measures that are developed by emergency care researchers and tested locally in hospitals may be ideally positioned for submission and review by an expert panel, especially if there is an evidence base documenting their importance, scientific acceptability, usability, and feasibility.

Second, emergency physicians can contribute to the review and approval of quality measures by volunteering to serve on expert panels, reviewing quality measures, and offering public comments or comments through membership in organizations. Finally, all nationally approved and implemented measures will undergo measure maintenance on a regular basis. Emergency physicians can influence this process by conducting specific research to address flaws or limitations of measures and then submitting their research directly to vetting organizations like the NQF and AHRQ or implementers like TJC or CMS. For example, the antibiotic timing and blood culture core measures for pneumonia underwent modification in response to research and advocacy by emergency physicians.

The Current Role of SAEM in Guideline and Performance Measure Development

SAEM does not currently have a process to develop its own guidelines, but has been invited to appoint representatives to committees and organizations that draft guidelines or endorse guidelines written by other societies. With the growth in guideline development and importance, SAEM saw the need to become more proactive in the development of guidelines that affect EM. In 2007, SAEM established the Guidelines Committee to

1) identify medical organizations responsible for guidelines and to educate them regarding the mission and expertise of SAEM; 2) establish working relationships with, and functional strategies for, collaboration with these groups; and 3) develop a mechanism to evaluate guidelines sent to SAEM for review or endorsement.

Since its founding, the Guidelines Committee has created materials for SAEM representatives that articulate standardized operating procedures and expectations for these relationships. It has also recommended topic
experts to represent SAEM on outside committees working on the creation or revision of treatment guidelines and has reviewed several guidelines, suggested revisions, and made recommendations to the SAEM Board regarding decisions on whether or not to endorse them.

More recently, the SAEM Guidelines Committee has been tasked with developing a mechanism for regular review of proposed performance measures for emergency care and providing an evidence-based analysis for the sponsoring agency on behalf of SAEM. The SAEM Guidelines Committee has surveyed the landscape to understand how performance measures are developed and where in the process SAEM can exert influence to best represent the membership.

**Future Directions for Guidelines, Quality Measurement, and Pay-for-performance in Emergency Care**

Similar to all medical specialties, EM will continue to be affected by the nationally driven health care performance improvement effort. There are unique aspects of the SAEM membership that bring added value to the process of guideline and performance measure development. Notably, the membership of SAEM encompasses the leading emergency care researchers, the leadership of academic EDs, EM experts in guideline development and implementation, and expert clinician educators who can provide valuable, if not crucial input to the guideline development process. Through implementation of performance improvement activities within academic and community hospitals, and publication of findings, SAEM members can validate performance measures and identify unintended consequences, demonstrating the crucial role emergency care researchers have in the quality improvement dialogue. SAEM members will also continue their role as performance improvement educators both within and outside of the specialty. Finally, SAEM members will ultimately provide the training and opportunities to develop the next generation of emergency care performance improvement experts to continue this evolution.

In addition to its role as topic experts and researchers, SAEM is also in the unique position to foster the development of emergency research network participation. Either through existing networks, such as Neurological Emergencies Treatment Trials and Resuscitation Outcomes Consortium, or through newly created performance measures or pay-for-performance networks, SAEM stands to lay the infrastructure for the development and evaluation of policies related to guideline and performance measure development and pay-for-performance.

**RECOMMENDATIONS**

The following are specific recommendations for SAEM’s participation in the process (Figure 1):

1. **SAEM Should Be a Sponsor of Guideline and Quality Measurement Research**

   Research findings should ideally validate guideline and quality measurement recommendations, along with identifying barriers to implementation and developing successful facilitation strategies. SAEM should consider actively supporting research that addresses emergency care guidelines and quality measures, concurrently confirming the value of active guidelines and measures as well as developing the evidence and content for future guidelines and measures. In conjunction with the Research Committee, the Guidelines Committee would be a logical group for making recommendations to the SAEM Board of Directors regarding prioritization of this support. Support for high-priority research could be demonstrated in several ways, including specific tracks for presentation at the annual meeting (thus encouraging the membership to focus in specific areas), as well as the identification and formation of topic-specific interest groups supported by SAEM infrastructure (e.g., providing administrative support, arranging meeting rooms at annual meetings, providing list server function) at the direction of the Board of Directors. Through discussions with Academic Emergency Medicine, SAEM could promote a call for papers regarding guidelines and performance measures to be included in a specific issue. SAEM should also consider interfacing with other potential funding sources (such as the Emergency Medicine Foundation or the AHA) for areas labeled as SAEM priorities. On certain topics, this approach could yield additional support.

---

**Figure 1.** Proposed steps in performance measure development process where SAEM can provide content experts and research expertise to influence the final product. AHRQ = Agency for Healthcare Research and Quality; AMA-PCPI = American Medical Association-Convened Physician Consortium for Performance Improvement; CMS = Centers for Medicare and Medicaid Services; CME = Continuing Medical Education; NCQA = National Committee for Quality Assurance; NOF = National Quality Forum.
2. SAEM Should Increase Participation in the Process of Guideline and Quality Measure Development, Endorsement, Implementation, and Maintenance
Involvement in the development and approval of candidate performance measures will help SAEM participate in the development of credible, meaningful, evidence-based performance measures for emergency care. The SAEM membership includes topic experts and expert clinician-researchers who have contributed to the science of performance measurement. Their expertise could be valuable in developing new measures or revising existing ones.

In our opinion, the two organizations with which SAEM may have the greatest potential to influence performance measure development are the AMA-PCPI and the NQF. SAEM is ineligible to become a voting member of the AMA-PCPI (since only members of the AMA House of Delegates, the Council of Medical Specialty Societies, and the American Board of Medical Specialties’ member boards are eligible). However, SAEM can still play a significant role. According to the AMA-PCPI, “any organization or individual committed to health care quality improvement and/or patient safety, and participates in the development, review, dissemination, or implementation of performance measures and measurement resources” is eligible to join the AMA-PCPI as a nonvoting member.

Therefore, SAEM can participate in AMA-PCPI working groups to provide evidence-based critiques and/or expert opinion representing academic EM. SAEM is, on the other hand, eligible to become a voting member of the NQF in the Health Professionals Council; members of the NQF Health Professionals Council provide the perspective of practicing health care providers regarding performance measurement and public reporting. Membership in the AMA-PCPI will provide SAEM the opportunity to nominate members to working groups developing measure sets that will affect emergency care, while membership in the NQF will give a voice to SAEM and a vote on whether measures should receive endorsement for national implementation. The SAEM Board of Directors has approved joining the NQF. SAEM should also participate in national organizations developing performance measures such as TJC or CMS by nominating SAEM members to serve on committees and workgroups. In addition, SAEM should work to encourage clinical researchers to interface with quality specialists to translate their research findings into guidelines and performance measures. Specifically, researchers can provide evidence that might lead to the adjustment of a measure (such as demonstrating unintended consequences) and can bring their findings to the attention of the NQF during the maintenance phase.

3. SAEM Should Increase Collaboration With Other EM Organizations to Review Performance Measures Proposed by Organizations Outside of EM That Affect Emergency Medical Care
Presenting non-ED measure developers with a unified voice from the ED advocacy and research communities would provide a powerful statement of support or condemnation where appropriate. This, in turn, could lead to increased visibility and future emergency physician participation in measure development and implementation.

4. SAEM Should Answer Calls for Participation in the Selection and Implementation of Performance Measures Through TJC and CMS
The Guidelines Committee could be tasked with actively monitoring calls for participation in the selection and implementation of performance measures by these regulatory bodies. The membership would be notified, and topic experts and guideline development experts from SAEM sought to provide representation and/or opinion on the candidate measures with regard to their validity, feasibility, unintended outcomes, and appropriate postimplementation monitoring.

SUMMARY
In this article, we have described the different processes involved in guideline and performance measures development and implementation, along with their current status within the emergency medicine community. Multiple opportunities exist for SAEM and its members to affect guideline and performance measure development, implementation, and maintenance. As well-developed guidelines and performance measures may improve patient care, SAEM’s involvement in these processes furthers its mission. Finally, future directions suggest increased opportunities for emergency medicine involvement and highlight the importance of SAEM’s unique contribution to the process.

The authors acknowledge the SAEM Guidelines Committee.

References


Appendix A

SAEM Guidelines Committee 2009–2010
Charles V. Pollack, Jr., MA, MD, Chair
Rita Cydulka, MD
Deborah Diercks, MD
Christopher Fee, MD
James Hoekstra, MD
Lance Becker, MD
Charles Emerman, MD
David Karras, MD
Jeff Kline, MD
W. Frank Peacock, MD
David Talan, MD
Ed Jauch, MD
Richard Summers, MD  
Jeremiah Schuur, MD  
Jesse Pines, MD  
Charlene Irvin, MD  
Brent Asplin, MD  
Anthony Ferrogiaro, MD

Greg Fermann, MD  
Benjamin Leacock, MD  
Maryann Mazer, MD  
David Nestler, MD  
Ellen Weber, MD  
Judd Hollander, MD