Chief Complaint-Based Performance Measures: A New Focus For Acute Care Quality Measurement

Richard T Griffey, MD, MPH[1], **Jesse M. Pines, MD**[2], **Heather L. Farley, MD**[3], **Michael P Phelan, MD**[4], **Christopher Beach, MD**[5], **Jeremiah D Schuur, MD, MHA**[6], and **Arjun K. Venkatesh, MD, MBA**[7]

[1] Division of Emergency Medicine and Institute for Public Health, Washington University School of Medicine, St. Louis, MO

[2] Departments of Emergency Medicine and Health Policy, The George Washington University School of Medicine, Washington D.C.

[3] Department of Emergency Medicine, Institute for Patient Safety, Cleveland Clinic, Cleveland, OH

[4] Department of Emergency Medicine, Christiana Care Health System, Wilmington, DE

[5] Department of Emergency Medicine, Northwestern Feinberg School of Medicine, Chicago, IL

[6] Department of Emergency Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

[7] Department of Emergency Medicine, Yale University School of Medicine, New Haven, CT

Abstract

Performance measures are increasingly important to guide meaningful quality improvement efforts and value-based reimbursement. Populations included in most current hospital performance measures are defined by recorded diagnoses using International Disease Classification (ICD)-9 codes in administrative claims data. While the diagnosis-centric approach allows the assessment of disease-specific quality, it fails to measure one of the primary functions of emergency department (ED) care which involves diagnosing, risk-stratifying, and treating patients' potentially life-threatening conditions based on symptoms (i.e. chief complaints). In this paper we propose chief complaint-based quality measures as a means to enhance the evaluation of quality and value in emergency care. We discuss the potential benefits of chief-complaint based measures, describe opportunities to mitigate challenges, propose an example measure set, and present several recommendations to advance this paradigm in ED-based performance measurement.

Corresponding Author: Richard T. Griffey, MD, MPH, FACEP Associate Chief Director of Quality and Safety Division of Emergency Medicine Associate Professor Washington University School of Medicine Washington University Institute for Public Health Campus Box 8072, 660 S. Euclid Ave. Barnes-Jewish Hospital, St. Louis, MO 63117 314.747.4899 314.362.0478 [Fax] griffeyr@wustl.edu.

Introduction

Performance measurement in medical care is becoming increasingly important, particularly with provisions of the Affordable Care Act (ACA) that support the public reporting of measures both to help inform patients and to determine provider payments. To date most performance measures are disease-specific, often reflecting clinical guidelines (e.g. proportion of patients with ST-segment elevation myocardial infarction with percutaneous intervention within 90 minutes) and specified with International Classification of Diseases, Ninth Edition (ICD-9) diagnosis codes available in administrative claims data generated for billing purposes. Diagnosis-based measures are advantageous in many areas of quality measurement. Diagnoses are readily available in administrative claims data and can be easily used to define a discrete population – the denominator of cases in measure calculation for whom an evidence-based treatment, practice or outcome can be measured.

However, this measurement approach does not capture key aspects of the quality of emergency care, which involves the sorting of undifferentiated complaints into diagnoses and making decisions based on limited information. Diagnosis and risk stratification of these complaints have become important functions of emergency physicians (EPs) as EDs increasingly comprise the primary site of acute, unscheduled care.[1] Diagnosis-based performance measures are inherently less able to capture the value that is provided in this task. For example, the ability of EPs to accurately and efficiently risk-stratify and treat patients with chief complaints (e.g. chest pain, fever, and headaches), after the consideration and safe exclusion of diagnoses such as acute myocardial infarction, sepsis and subarachnoid hemorrhage cannot be directly measured using a diagnosis-based approach. Measures that only include patients for whom a specific diagnosis is made also systematically exclude the measurement of quality in patients who were considered for a diagnosis, but for whom it was ruled out.

Chief complaint-based performance measures offer the opportunity to assess quality and reflect the value of care delivered in the ED. In this paper, we provide an overview of the efforts to date to advance chief complaint-based measures, describe some historical challenges, and propose some potential solutions to support the development of chief complaint-based measures. We also include an example of a potential measure set for chest pain organized by the priority areas of the National Quality Strategy and offer some other potential areas for development of chief complaint-based measurement.

A Brief History of Chief Complaint-Based Measures

Chief-complaint measures are not a new concept. In 2005, the Quality Improvement and Patient Safety (QIPS) Section of the American College of Emergency Physicians (ACEP) developed a set of six Chief Complaint-based Quality Measures, intended for use by ED quality directors for local quality improvement.[2] In 2008, the ACEP Quality and Performance Committee developed 10 quality measures for ED care, of which two were chief complaint-based, but were identified using ICD-9 diagnoses. Two of these original measure concepts were advanced by ACEP and ultimately endorsed by the NQF as part of 22 National Voluntary Consensus Standards for Emergency Care.[3] These included, "Rate of HCG testing among women ages 14 to 50 with the chief complains of abdominal pain"

(NQF #502) and "Ultrasound determination of pregnancy location for pregnant patients with abdominal pain" (NQF #0651). While it is possible to create complaint-based measures using diagnosis codes, in many cases this approach is suboptimal because of the poor correlation between chief complaint and ICD-9 diagnosis,[4] and the systematic exclusion of patients for whom another discrete condition is identified.

The *HCG testing for Abdominal Pain* measure was initially given 'time-limited' endorsement by the NQF in 2012, but later failed to achieve permanent endorsement after the NQF Steering Committee's judgment that there was neither evidence of a gap in care nor evidence of the impact of performing pregnancy testing on outcomes. The *Ultrasound determination of Pregnancy* was also given time-limited endorsement and may be reevaluated by the NQF in 2014.

The Physician Quality Reporting System (PQRS), specifies measures for physician-level reporting for Medicare payment incentives and in 2013 included 259 quality measures, of which 241 were reported using ICD-9 [5] diagnosis-based administrative claims or registries. Two of the five measures most frequently reported on by EPs are the chief complaint-based process measures of EKG performance in patients with syncope and chest pain. However, the NQF Measures Applications Partnership (MAP), which makes recommendations to the Centers for Medicare and Medicaid Services (CMS) regarding the selection of quality measures for federal payment programs no longer supports the use of these measures. The MAP considered both "topped out" (i.e that performance variability is lacking and/or gaps in quality have been closed and with limited further potential for improvement) and so no longer address an important area for measurement.[6] Chief complaint-based measures have been developed in other specialties endorsed by the NQF (e.g. #0514: *MRI Lumbar Spine for Low Back Pain*) and are being used in PQRS, but these represent a minority of measures.

For various reasons, many of the current measures used for PQRS including 4 of the 7 measures in the EM cluster (not only chief-complaint based measures) will soon be retired, leaving fewer measures available for physician reporting and creating a need for new measures and an opportunity to consider chief complaint-based performance measures.

Challenges in the development of Chief Complaint-Based Measures

Common language/ consensus standards

A key barrier to the use of chief complaints for performance measurement lies in the lack of standardization of complaint-based nomenclature, and how chief complaints are organized, categorized and assigned. This has implications for many aspects of quality measurement downstream. Defining reliable patient cohorts for chief complaint-based measures is impossible without some level of standardization. Standardized chief complaints would support not only quality measurement initiatives but also other meaningful endeavors such as biosurveillance.[7-12]

Agreement on what terms to use, which presentations fall into which category and at what level of granularity conditions should be described are some of the fundamental concepts that require clear definitions but lack consensus. Standardized language and ontology are the

first steps. Methods are then needed for implementation that minimizes variability in how chief complaints are assigned from ED to ED, nurse to nurse and visit to visit. Researchers in emergency medicine have called for and have attempted to devise standardized chief complaint language for well over a decade. Various attempts have been made to organize and standardize data elements[12-16] and to sort existing language into standardized formats.[17,18] This issue is not limited to the US; publications in Canada and Australia have proposed various schemes to standardize the language of chief complaints.[19,20] A Society for Academic Emergency Medicine (SAEM) consensus panel also made recommendations for standardizing chief complaint terminology. This was followed in 2006 by a multi-stakeholder meeting and publication "Towards Vocabulary Control for Chief Complaint," which made consensus recommendations to this end.[21]

A Path Forward

One of the first detailed clinical content standards created for general use by the public was Data Elements for Emergency Department Systems (DEEDS), published in 1997 by the US Center for Disease Control and Prevention (CDC) and the National Center for Injury Prevention and Control. DEEDS used existing standards developed by organizations such as Health Level 7 International (HL7), The American Society for Testing and Materials, and others as a starting point for data element definitions. Though the original DEEDS did not result in standardized chief complaint language and ontology, changes in the intended use of these standards, the practice of clinical informatics, the scope of emergency care, vendor needs in implementing Emergency Department Information Systems (EDIS), and the availability of newer HL7 v3 standards, have led to an initiative to update DEEDS. The Emergency Care Workgroup of HL7 International, a non-profit standards setting organization for the exchange, integration, sharing and retrieval of electronic health information, has the purview to update DEEDS specifications for data elements to complaint taxonomy.[22]

An alternative to promulgating a standardized chief complaint schema is to use modern computing capabilities to handle free-text chief complaints and those that already exist in EMRs to group these in a standardized fashion.[23,24] Researchers have used existing coding schemes to map free text chief complaints to standardized terminologies. General purpose coding schemes such as the ICD-9 CM, the Standardized Nomenclature of Medicine Clinical Terms (SNOMED CT) and the Unified Medical Language System (UMLS) usually cover diseases, clinical findings and procedures across care sites and provider types, and are organized to relate concepts like chief complaints in a hierarchical fashion. The UMLS also includes tools to normalize free text expressions with different tenses, punctuations, spelling, etc. to allow for further processing, and combines terms from different coding schemes into hierarchical relationships. Coding systems, such as the CDC's reason for visit classification system (RVC) used in the National Ambulatory Medical Care Survey relate more specifically to conditions, but lack a hierarchical structure.[25] Because coding schemes include large sets of terms and categorizations that may be irrelevant to ED care, various efforts have attempted to cull these to smaller lists that capture most ED complaints.[26-28] Determining the right number of chief complaint categories to capture most ED patients cases while providing the best level of specificity is challenging.[29] The inability to create a

mutually exclusive chief complaint set that can capture all ED visits should not hinder the development of chief complaint performance measures, however, as measures typically require methods for reliably capturing patients with common, undifferentiated chief complaints.

It may be that hybrid approaches using both standardized, structured chief complaints and processed free text could provide the best amount of information while allowing clinical flexibility. Some authors have proposed use of standardized coded complaints with free text modifiers.[27,30] Other studies have processed free text for syndromic surveillance using the Emergency Medical Text Processor (EMT-P) system, which uses the UMLS tools described above and a manually compiled synonym list, to process free text into standardized categories.[29,31] It is not hard to imagine a similar approach substituting medical conditions for syndromes to identify conditions for which measures of evidence-based care would apply. Though applying any such work to performance measurement is at this point theoretical, such methods used to standardize existing chief complaint schemes for other purposes may provide potential approaches for doing so.

The development of chief-complaint standards through efforts by organizations such as HL7 alongside potential federal payment incentives in Meaningful Use Stage 3 regulations could create important regulatory and financial momentum for vendors and providers to establish a common platform from which chief-complaint quality measures could be developed.[22,32] Similarly, through federal regulations, the Office of the National Coordinator (ONC) has the ability to include chief-complaint based measures in physician incentive programs that could rapidly drive EHR vendors to integrate chief complaint vocabularies and measurement specifications into existing platforms.

Beyond claims data: the rise of E-measures

A recent shift among measure developers, users and endorsement bodies is toward the use of the electronic "E-measure" format. E-measures draw from a number of data streams within an electronic health record (EHR), including clinician documentation, laboratory data, orders and time stamps, as well as from other electronic systems such as a picture archiving and communication system (PACS) or computerized risk-stratification tools providing rich information to inform risk adjustment methods and identify patients for whom quality of care should be measured. The goal of E-measures is to leverage the power of EHRs to evaluate quality using better measures that have more standardized definitions, and reduces the burdens of measurement.

Attempting to measure chief-complaint based care using traditional quality measurement approaches such as chart abstraction is both expensive, time-consuming and is nearly impossible to do reliably—the use of e-measures could enable select populations within a chief complaint for which a high-risk diagnosis is suspected to be identified for quality measurement. For example, an e-measure could identify a subset of patients being evaluated for pulmonary embolus by incorporating the results of laboratory tests such as a D-dimer and/or imaging with CTPA in patients who present with a chief complaint coded as dyspnea or chest pain. This might be used to determine the proportion of low or moderate risk patients who undergoing d-dimer testing prior to CT in a more efficient fashion than chart

review. One study found a 50% reduction in abstractor time required for one measure set alone using e-measures.[33] E-measures can also use natural language processing (NLP) to help automate case finding for quality measurement. A comprehensive description of NLP functionality is beyond the scope of this paper and can be found elsewhere.[34] Briefly, NLP can identify similar terms used to describe one condition (e.g. "hypertension" and "high blood pressure" or to comb through imaging reports to identify specific findings that may help in case-finding, avoiding the need for costly chart abstraction.[35,36] This same NLP framework could be applied upstream to help define the denominator of patients with a suspected illness, and has even been proposed for real-time use in EHRs. For example, performance measures could be developed for patients who present with headaches with specific features that suggest the patient is being evaluated for subarachnoid hemorrhage. This would obviate the current process of gathering ICD-9 codes and performing chart review to look for exclusions to a measure. E-measures that flow from a patient complaint or set of complaints, could thus demonstrate value in reflecting steps taken in appropriate/ evidence-based approaches to various complaints such as the efficient use of CT imaging or coordination for follow-up among patients with high risk headaches. However, prior to the robust use of E-measures, a standardized system for chief complaint assignment needs to be agreed upon to ensure reliable measurement between providers and between EDs.

Different physicians may assign related but different diagnoses to the same patient encounter. These encounters may also be assigned billing codes differently by different coders. Since quality measures using administrative codes usually only look at principal diagnoses and rarely secondary diagnoses, such measures may often fail to capture many patients or the necessary exclusions for measurement. For example, a patient presenting with syncope in whom this is found to be related to a specific dysrhythmia may be assigned a diagnosis code for the dysrhythmia but not for syncope, excluding the very cases that are the target of the syncope measure.[37] In addition, physicians or coders may not include secondary diagnoses that do not seem important to list as a diagnoses yet which may be relevant as an exclusion from a performance measure. For example, a patient may carry a principal diagnosis of tension type headache (ICD-9 331.9) without mention in the secondary diagnosis that the patient also had HIV or was presenting with paresthesias. In the case of a proposed quality measure related to CT use for headache (OP-15), this exclusion for the measure would be missed.[38] In this first scenario an E-measure with standardized chief complaint might capture the reason for an EKG as "syncope" or "passing out." Similar language could potentially be captured by NLP applied to the documented patient history in an electronic record, which may be one way to standardize chief complaints. Similarly, the specifications of an E-measure could search for potential measure exclusions to prevent the second scenario. E-measures may be better equipped to gather data relevant to quality than the blunt instrument of an administrative data code.

Research to support measures

The rapid evolution of performance measurement now demands more sophisticated methods that meet a higher standard of scientific acceptability with the expectation that measures can be used for accountability and public reporting.[39] Unlike disease-specific measures, chief-complaint based measures are harder to evaluate using NQF's existing Measure Evaluation

Criteria because of challenges in ensuring the reliability of capturing the appropriate patient population (the measure denominator) both over time and between ED's. In addition, demonstrating the *Scientific Acceptability* of a performance measure requires not only a strong body of evidence showing a relationship between a process or intervention and meaningful patient outcomes and evidence but also that the relevant data can be reliably captured. The NQF Measure Evaluation Criteria should be modified to equitably evaluate chief complaint-based measures. The criteria should specify standard approaches to reliability assessment of measure denominators in addition to the process or outcome being measured.

These standards can be more difficult to meet for chief complaint based measures. With the exception of a limited number of chief complaints such as chest pain, the past thirty years of emergency care research related to treatment has largely been disease-focused for conditions such as traumatic injuries, and stroke. By comparison, complaints such as shortness of breath or abdominal pain, for which patient populations can be quite heterogeneous, have been more commonly included in research related to diagnostic accuracy. Measures related to diagnostic accuracy could be important in reflecting the evolving role of the ED as a diagnostic center. More accurate definition of chief complaints would also allow for the measurement of utilization, such as admission observation, or cardiac stress testing for a given complaint, with a better-defined denominator of patients. Challenges to this approach remain in accurate risk-stratification. There may not be a definitive diagnosis against which to judge accuracy, except in the case of serious diagnoses that may appear in claims data shortly after the initial visit such as missed AMI or SAH.

Given the lack of validated measures in this area and the potential to develop robust measures of utilization, missed diagnosis, and diagnostic accuracy, funders and researchers should expand efforts to design related research in a manner that can translated to quality measurement. Measures will need to leverage the emerging diagnostic evidence base for these conditions to reframe the measure evaluation process. Though many common presentations to the ED may not fit well into a scheme that attempts to measure care in a prospective way, there are some areas where patients presenting with specific complaint groupings should, as a rule, receive specific components of care. Initial approaches to developing chief complaint-based measures should focus on chief complaints for which there is a broad evidence base on prognosis, diagnostic strategies and treatment. Many such chief complaints may be diagnoses themselves (e.g. seizure or epistaxis) (see Table 1), or symptoms that are proximate to important diagnoses such as AMI (e.g. chest pain) with evidence-based steps in care. Another option might be to use discrete questions addressed in existing complaint-based ACEP Clinical Policies for which data have been rigorously reviewed with graded recommendations for care, such as chief complaints (that are also diagnoses) addressed by ACEP Clinical Policies (**Table 1**).[40-43]

The measure endorsement process and reframing measure evaluation criteria

Performance measures are often developed by specialty societies, regulatory, advocacy or payer groups, their contracted agents or other organizations in response to calls for measures. Adoption of measures typically requires endorsement by the NQF, a process that

involves multi-stakeholder review by technical expert panels using specific evaluation criteria. This process has changed in a number of ways that are important to the development of chief complaint based emergency care measures.

Newly-endorsed measures are expected to exist in the public domain, be intended for accountability through public reporting and pay-for-performance programs, include robust reliability and validity testing and be harmonized with similar measures. Measure testing to demonstrate validity and reliability is estimated to cost over $100,000 per measure. These higher expectations for measure development methodology are important, have the potential to price clinical providers or specialty organizations out of the measure development process.

In the past, the NQF has convened work to advance the development of conceptual models for quality measures of regionalized emergency care, crowding, and preparedness, and in 2007 endorsed 22 measures of emergency care. More recently, as part of the National Quality Strategy there has been a move away from unit or department-specific measures by CMS and the AHRQ. There is currently no measure evaluation pathway specific to emergency or acute, unscheduled care. This means that measures relevant to emergency medicine must be advanced through one of the 19 recently-announced measure evaluation pathways, which are focused on specific clinical areas rather than on a care model.

The NQF measure evaluation and endorsement process is primarily accomplished by Steering Committees comprised of multiple medical specialties, payers and consumers who evaluate measures within discrete clinical areas such as "Neurology" or "Infectious Diseases."[44] As a consequence, candidate measures are evaluated alongside a number of other measures, any of which may only impact the specialty of 1 or 2 committee members. It is often the case that most committee members may not work in environments where chief complaint-based measures are relevant and so may be less sensitive to their importance or applicability.

This presents the possibility that endorsement may have as much or more to do with the process and pathway under which a measure is advanced than the quality of the measure itself. For example, the ACEP-developed NQF #502 "Rate of HCG testing among women ages 14 to 50 with the chief complains of abdominal pain," which had already been endorsed by NQF based on data demonstrating wide variation in meeting this guideline, scored highly across all criteria but failed to be re-endorsed in 2011 by a steering committee on Perinatal and Reproductive Health, which included only one EP. When viewed in comparison to disease-specific measures, such as the provision of antenatal steroids for pre-term birth, the Steering Committee may have given lower importance to a measure about patients with abdominal pain, but who did not have a specific disease.

Emergency Medicine should advocate for the development of an NQF Steering Committee focused on acute, unscheduled care. A standing, multi-stakeholder group that includes clinical and research input from specialties familiar with chief complaint-based care pathways such as family medicine, pediatrics, hospital medicine, and emergency medicine is critical to developing alignment for quality improvement.

Potential areas for development of chief-complaint based measures

The development of chief-complaint based measures is not out of reach, as there are numerous ACEP clinical policies based on strong clinical research for several chief complaints. A perusal of chief-complaint-driven, order entry and electronic documentation templates yields a number of areas for which adherence to evidence-based care pathways could be efficiently measured using EHRs and provide a more comprehensive assessment of emergency care than would use of administrative claims data.

Chest Pain as a Potential Chief Complaint-Based Measure Set

After abdominal pain, chest pain is the most common reason for adult patients to seek care in an ED.[45] Over 5 million patients each year arrive with this complaint and less than 15% have a final diagnosis of Acute Coronary Syndrome, the primary concern for many patients when they seek ED care for chest pain. Additionally, cost of care for these patients has increased with increasing utilization of advanced imaging studies and longer lengths of stay inherent in the comprehensive (and often definitive) assessment and management of chest pain. Given the many other serious, life-threatening diagnoses that EPs consider and evaluate in this group, chest pain is arguably the most important chief-complaint for quality measurement.

To date, a number of chief-complaint based measure concepts have been developed for chest pain. We present these measures as well as their use within public reporting and payment programs in the **Table 2** along with additional measure concepts we propose for consideration, organized by the six priority areas of the National Quality Strategy. We do not expect chief complaint-based measures to replace existing diagnosis-based measures but rather to complement these in relevant measure sets. As expected, measures of chest pain care exist in a continuum with measures of care for acute myocardial infarctions. This diagnosis-driven approach has resulted in measures of suspected ACS, which represents only one potential diagnosis in patients who present with chest pain, but falls short of providing validated measures of emergency care for patients with other diagnoses as pericarditis, pneumothorax, acute aortic dissection, etc. that might map to the chief complaint of chest pain. Measures of appropriate D-dimer or Computed Tomography use for suspected pulmonary embolism are shown in the table as potential example measures to fill this gap.

Many of the measures detailed in the table have achieved NQF endorsement and demonstrate early successes in the development of validated, chief-complaint based measures. However, also notable is the absence of chest pain quality measures across multiple quality domains. Many of the efficiency and outcome measures listed have been available in the public domain, but have not been subject to the scrutiny of the NQF endorsement process. Many of these have also not demonstrated the necessary face validity for widespread adoption due to their reliance on administrative claims data. Future electronically specified measures could not only apply the use of standard chief-complaint language to define the measures but also define numerators and denominators for measures such as the chest-pain rule-out and rule-in rate based on the use of biomarker testing and results available in the EHR.

Process measures of clinical effectiveness dominate chest pain measure sets, but few meet the National Quality Strategy domain of safety and there are no measures of patient experience or patient reported outcomes. Future chest pain measures seeking to improve safety could potentially focus on items such as prevention of contrast-induced nephropathy in chest pain patients, for example. Additionally, measures should build upon the growing evidence base supporting shared-decision making to guide patient-oriented chest pain evaluations[46] and chest pain-specific measures of patient experience that build upon the growing Hospital Care Quality Information from the Consumer Perspective (HCAHPS) program already in use by hospitals. Measures should also be developed modeled on the referral loop closure measure now included as part of the CMS EHR Incentive Program to ensure the timely and accurate communication of ED chest pain evaluations with outpatient providers and vice-versa.

Chief-complaint based measures also have the potential to reflect quality in aligning systems of care and in improving population health. Continuing with chest pain measure sets as an example, measures that support smoking cessation counseling in the ED not only promote evidence based care[47] but drive meaningful population-level, public health impact. That NQF measures related to aspirin prescribing on arrival are approved for use in both PQRS and HOQR (Hospital Outpatient Quality Reporting) payment programs demonstrates the potential for alignment between EPs and hospitals in developing chief complaint-based measures. Similarly measures such as NQF #0290 regarding timeliness of patient transfers can help support the development of regionalized emergency care systems and community level solutions that demonstrate high-performance emergency care. [48] It should be noted that many of the measure concept in Table 2 represent ideas that have not been developed or endorsed as national measures. As all measures include the risk of unintended consequences these would need to be developed in such a way to minimize these risks and to reflect areas over which those being measured have control. For example, measures related to ED revisits may be more reflective of community-based no-ED resources and so might be specified for health systems rather than the ED per se. Such a measure might help drive improved coordination across providers as hospital readmission measures have done.

Acknowledgments

References

1. Pitts SR, Carrier ER, Rich EC, Kellermann AL. Where Americans get acute care: increasingly, it's not at their doctor's office. Health Aff (Millwood). 2010; 29:1620–9. [PubMed: 20820017]

2. Chief Complaint Based Quality Indicators: American College of Emergency Physicians Quality Improvement and Patient Safety Section. 2005.

3. [8/1/2013] National Voluntary Consensus Standards for Emergency Care. 2013. (2013, at http://www.qualityforum.org/Publications/2009/09/National_Voluntary_Consensus_Standards_for_Emergency_Care.aspx.)

4. Raven MC, Lowe RA, Maselli J, Hsia RY. Comparison of presenting complaint vs discharge diagnosis for identifying "nonemergency" emergency department visits. JAMA. 2013; 309:1145–53. [PubMed: 23512061]

5. International Classification of Diseases. 9th editionat http://www.who.int/classifications/icd/en/.)

6. [April 10, 2014] National Quality Forum Measure Applications Partnership. 2013. (at http://www.qualityforum.org/map/.)

7. Barthell EN, Aronsky D, Cochrane DG, Cable G, Stair T. The Frontlines of Medicine Project progress report: standardized communication of emergency department triage data for syndromic surveillance. Ann Emerg Med. 2004; 44:247–52. [PubMed: 15332067]

8. Barthell EN, Cordell WH, Moorhead JC, et al. The Frontlines of Medicine Project: a proposal for the standardized communication of emergency department data for public health uses including syndromic surveillance for biological and chemical terrorism. Ann Emerg Med. 2002; 39:422–9. [PubMed: 11919529]

9. Travers D, Wu S, Scholer M, Westlake M, Waller A, McCalla AL. Evaluation of a chief complaint pre-processor for biosurveillance. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2007:736–40. [PubMed: 18693934]

10. Shapiro AR. Taming variability in free text: application to health surveillance. MMWR Morbidity and mortality weekly report. 2004; 53(Suppl):95–100. [PubMed: 15714636]

11. Mandl KD, Overhage JM, Wagner MM, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. Journal of the American Medical Informatics Association : JAMIA. 2004; 11:141–50. [PubMed: 14633933]

12. Fleischauer AT, Silk BJ, Schumacher M, et al. The validity of chief complaint and discharge diagnosis in emergency department-based syndromic surveillance. Acad Emerg Med. 2004; 11:1262–7. [PubMed: 15576514]

13. Cordell WH, Overhage JM, Waeckerle JF. Strategies for improving information management in emergency medicine to meet clinical, research, and administrative needs. Information Management Work Group. Acad Emerg Med. 1998; 5:162–7. [PubMed: 9492140]

14. Chute CG, Cohn SP, Campbell JR. A framework for comprehensive health terminology systems in the United States: development guidelines, criteria for selection, and public policy implications. ANSI Healthcare Informatics Standards Board Vocabulary Working Group and the Computer-Based Patient Records Institute Working Group on Codes and Structures. Journal of the American Medical Informatics Association : JAMIA. 1998; 5:503–10. [PubMed: 9824798]

15. McClay JC, Campbell J. Improved coding of the primary reason for visit to the emergency department using SNOMED. Proceedings / AMIA Annual Symposium AMIA Symposium. 2002:499–503. [PubMed: 12463874]

16. Husk G, Akhtar S. Chief complaints, emergency department clinical documentation systems, and the challenge of dealing with the patient's own words. Acad Emerg Med. 2007; 14:69–73. [PubMed: 17099190]

17. Travers DA, Haas SW. Evaluation of emergency medical text processor, a system for cleaning chief complaint text data. Acad Emerg Med. 2004; 11:1170–6. [PubMed: 15528581]

18. Thompson DA, Eitel D, Fernandes CM, Pines JM, Amsterdam J, Davidson SJ. Coded Chief Complaints--automated analysis of free-text complaints. Acad Emerg Med. 2006; 13:774–82. [PubMed: 16723726]

19. National Quality Forum Measure Evaluation Criteria. 2012.

20. Innes G, Murray M, Grafstein E. A consensus-based process to define standard national data elements for a Canadian emergency department information system. CJEM. 2001; 3:277–84. [PubMed: 17610770]

21. Haas SW, Travers D, Tintinalli JE, et al. Toward vocabulary control for chief complaint. Acad Emerg Med. 2008; 15:476–82. [PubMed: 18439204]

22. HL7 Version 3 Specification: Data Elements for Emergency Department Systems (DEEDS): HL7 International. 2012.

23. Niiranen ST, Yli-Hietanen JM, Nathanson LA. Toward reflective management of emergency department chief complaint information. IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society. 2008; 12:763–7.

24. Yli-Hietanen J, Niiranen S, Aswell M, Nathanson L. Domain-specific analytical language modeling--the chief complaint as a case study. International journal of medical informatics. 2009; 78:e27–30. [PubMed: 19307149]

25. Schneider, D. National Center for Health Statistics. Centers for Disease Control; Hyattsville, MD: 1979. A Reasonfor Visit Classification for Ambulatory Care..

26. Gorelick MH, Alpern ER, Alessandrini EA. A system for grouping presenting complaints: the pediatric emergency reason for visit clusters. Acad Emerg Med. 2005; 12:723–31. [PubMed: 16079425]

27. Aronsky D, Kendall D, Merkley K, James BC, Haug PJ. A comprehensive set of coded chief complaints for the emergency department. Acad Emerg Med. 2001; 8:980–9. [PubMed: 11581085]

28. Thompson DA, Eitel D, Fernandes CM, Pines JM, Amsterdam J, Davidson SJ. Coded Chief Complaints--automated analysis of free-text complaints. Acad Emerg Med. 2006; 13:774–82. [PubMed: 16723726]

29. Lu HM, Zeng D, Trujillo L, Komatsu K, Chen H. Ontology-enhanced automatic chief complaint classification for syndromic surveillance. Journal of biomedical informatics. 2008; 41:340–56. [PubMed: 17928273]

30. Haas SW, Travers D, Tintinalli JE, et al. Toward vocabulary control for chief complaint. Acad Emerg Med. 2008; 15:476–82. [PubMed: 18439204]

31. Travers DA, Haas SW. Unified medical language system coverage of emergency-medicine chief complaints. Acad Emerg Med. 2006; 13:1319–23. [PubMed: 17079790]

32. [5/3/2014] ONC crafting e-quality measure policies for MU3. 2013. (at http://www.govhealthit.com/news/crafting-e-quality-measure-policies-mu3-U43lqcf96RF.)

33. Garrido T, Kumar S, Lekas J, et al. e-Measures: insight into the challenges and opportunities of automating publicly reported quality measures. Journal of the American Medical Informatics Association : JAMIA. 2014; 21:181–4. [PubMed: 23831833]

34. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assn. 2011; 18:544–51.

35. Raja AS, Ip IK, Prevedello LM, et al. Effect of Computerized Clinical Decision Support on the Use and Yield of CT Pulmonary Angiography in the Emergency Department. Radiology. 2012; 262:468–74. [PubMed: 22187633]

36. Murff HJ, FitzHenry F, Matheny ME, et al. Automated Identification of Postoperative Complications Within an Electronic Medical Record Using Natural Language Processing. Jama-J Am Med Assoc. 2011; 306:848–55.

37. Schuur JD, Justice A. Measuring quality of care in syncope: case definition affects reported electrocardiogram use but does not bias reporting. Acad Emerg Med. 2009; 16:40–9. [PubMed: 19145713]

38. Schuur JD, Brown MD, Cheung DS, et al. Assessment of Medicare's imaging efficiency measure for emergency department patients with atraumatic headache. Ann Emerg Med. 2012; 60:280–90. e4. [PubMed: 22364867]

39. Berenson, RAPP.; Krumholz, HM. Achieving the Potential of Health Care Performance Measures: Timely Analysis of Immediate Health Policy Issues: Robert Wood Johnson Foundation. 2013.

40. American College of Emergency Physicians Clinical Policies Subcommittee on S. Huff JS, Melnick ER, et al. Clinical policy: Critical issues in the evaluation and management of adult patients presenting to the emergency department with seizures. Ann Emerg Med. 2014; 63:437–47. e15. [PubMed: 24655445]

41. Edlow JA, Panagos PD, Godwin SA, Thomas TL, Decker WW, American College of Emergency P. Clinical policy: critical issues in the evaluation and management of adult patients presenting to

the emergency department with acute headache. Ann Emerg Med. 2008; 52:407–36. [PubMed: 18809105]

42. Huff JS, Decker WW, Quinn JV, et al. Clinical policy: critical issues in the evaluation and management of adult patients presenting to the emergency department with syncope. Ann Emerg Med. 2007; 49:431–44. [PubMed: 17371707]

43. Fesmire FM, Brady WJ, Hahn S, et al. Clinical policy: indications for reperfusion therapy in emergency department patients with suspected acute myocardial infarction. American College of Emergency Physicians Clinical Policies Subcommittee (Writing Committee) on Reperfusion Therapy in Emergency Department Patients with Suspected Acute Myocardial Infarction. Ann Emerg Med. 2006; 48:358–83. [PubMed: 16997672]

44. Two-Stage Consensus Development Process Redesign. at http://www.qualityforum.org/Measuring_Performance/Improving_NQF_Process/Two-Stage_Consensus_Development_Process_Redesign.aspx.)

45. Bhuiya FA, Pitts SR, McCaig LF. Emergency department visits for chest pain and abdominal pain: United States, 1999-2008. NCHS data brief. 2010:1–8.

46. Hess EP, Knoedler MA, Shah ND, et al. The chest pain choice decision aid: a randomized trial. Circulation Cardiovascular quality and outcomes. 2012; 5:251–9. [PubMed: 22496116]

47. Bernstein SLBP, Cooperman N, Jearld S, Arnsten JH, Moadel A, Gallagher EJ. Efficacy of an emergency department-based multicomponent intervention for smokers with substance use disorders. J Subst Abuse Treat. 2013; 44:139–42. [PubMed: 22763199]

48. [July 2013] National Quality Forum: Regionalized Emergency Medical Care Services (REMCS) Measure Topic Prioritization. 2013. (at http://www.qualityforum.org/Projects/nr/Regionalized_Emergency_Medical_Services/Regionalized_Emergency_Medical_Care_Services_%28REMCS%29.aspx.)

Summary Recommendations

1. Standardized language for chief complaints should be developed based on substantial previous work. EHR vendors should have incentives as part of a Meaningful Use Criteria to ensure reliable coding across EDs.

2. Multi-stakeholder entities such as HL7 should support the development of chief-complaint value sets to enable EHR-based measures of emergency care. These value sets should be harmonized between measure developers and also be used to guide clinical research upstream of measure development efforts.

3. The emergency medicine research agenda including initial efforts by the newly developed NIH office of Emergency Care Research should emphasize the federal funding of chief-complaint based comparative effectiveness research.

4. The future development of chief-complaint based measures should move beyond easily available process measures to also include more outcomes measures (including patient-oriented and patient-reported outcomes) as part of measure sets designed to assess the quality of chief complaint based care.

5. Emergency medicine should advocate for the development of an NQF Steering Committee focused on acute, unscheduled care. A standing, multi-stakeholder group that includes clinical and research input from specialties familiar with chief complaint-based care pathways such as family medicine, pediatrics, hospital medicine, urgent care, and emergency medicine is critical to developing alignment for quality improvement.

6. The NQF Measure Evaluation Criteria should be modified to equitably evaluate chief complaint-based measures. The criteria should specify standard approaches to reliability assessment of measure denominators in addition to the process or outcome being measured.

7. We also propose that the NQF direct a project around the need for, barriers to, and next steps to develop chief complaint-based measures, as has been done for regionalization of care and other initiatives.

Summary

In summary, there is an increasing recognition of the limitations of diagnosis based performance measures in emergency medicine. This comes at a time of increasing opportunity to pursue chief-complaint based measures that demonstrate the value of emergency care. Barriers to chief complaint based measures remain, including nonstandard language for chief complaints and high costs and methodological hurdles for new measure development and endorsement. Emergency Medicine should pursue opportunities to overcome these barriers, including the development of a standard chief complaint nomenclature, research linking chief complaints and outcomes, and advocacy for policies to encourage vendors towards adoption of common chief complaint language. Without a proactive approach to measurement, emergency clinicians will be unable to describe the value of emergency care to not only payers, but more importantly, to patients.

Table 1

Potential areas for chief complaint based Measures

Headache (clinical policy)
Syncope (clinical policy)
Chest Pain (policy statement)
Seizure (clinical policy)
Adult mild traumatic brain injury (clinical policy)
Vaginal Bleeding (clinical policy in part)
Back Pain
Epistaxis
GI Bleeding
Weak/Dizzy
Visual changes
Wound care
Pediatric Fever
Pediatric Dehydration

Table 2

Current and Proposed Measure Concepts for a Chest Pain Measure Set

Measure	Current Use**	NQF Endorsement
Clinical Process/Effectiveness		
Participation in a Systematic National Database for General Thoracic Surgery	Selective adoption	#0456
EKG within 10 minutes of arrival	Available for use	#0090, #0665
EKG for non-traumatic chest pain	PQRS	#0090
Median time to EKG	HOQR	#0289
ASA on arrival	HOQR, PQRS	#0092, #0142, #0286
% AMI patients arriving in ED 2 hours after onset of symptoms (SCPC-CP1)	Measure Concept	
Time to fibrinolysis	HVBP, HOQR	#0164, #0288
Primary PCI within 90 minutes of hospital arrival for AMI	HBVP	#0163
Undetected missed ACS rate (SCPC CP15)	Measure Concept	
Door-in-Door-Out time for STEMI		
Time to CT surgery evaluation for acute ascending aortic dissection	Measure Concept	
Population and Public Health		
Tobacco use: Screening and Cessation Intervention	Not specified for emergency physician use	#0028
Patient Safety		
Proportion of (AMI or PNA) patients with potentially avoidable complication	Not specified for emergency physician use	#0704, 0708
Iatrogenic Pneumothorax	Requires new specification for emergency physician use	#0346
Prevention of contrast-induced nephropathy	Measure Concept	
Care Coordination		
Communication or results and discharge plan with PMD prior to ED discharge following evaluation for chest pain	Measure Concept	
Referral loop closure with primary cardiologist for ED patients with chest pain	Measure Concept	
Efficient Use of Healthcare Resources		
Avoidable imaging for patients with low-probability for PE	Available for use	#0667
ACS rule out evaluation rate (SCPC-CP6a)	Measure Concept	
ACS rule in evaluation rate (SCPC CP6b)	Measure Concept	
Observation usage for low probability chest pain (SCPC PC 16)	Measure Concept	
Chest pain ED revisitation rate	Measure Concept	
Relative resource use for people with cardiovascular conditions	Not specified for emergency physician use	# 1558
Cardiac stress imaging not meeting appropriate use criteria: Testing in asymptomatic, low risk patient	PQRS	
Patient and Family Engagement		
Use of shared decision making to guide evaluation in patients with low risk chest pain	Measure Concept	
Percent of chest pain patients who received and understood their diagnosis and care plan at ED discharge	Measure Concept	
ED-CAHPS for chest pain patients	Measure Concept	

*National Quality Strategy (NQS) Domains: Patient and Family Engagement, Patient Safety, Care Coordination, Population and Public Health, Efficient use of Healthcare Resources, Clinical Process/Effectiveness

**HOQR: Hospital Outpatient Quality Reporting; HVBP: Hospital Value-Based Purchasing; pQrS: Physician Quality Reporting System