ORIGINAL CONTRIBUTIONS

THE DEVELOPMENT OF EVIDENCE-BASED PREHOSPITAL GUIDELINES USING A GRADE-BASED METHODOLOGY

Kathleen M. Brown, MD, Charles G. Macias, MD, MPH, Peter S. Dayan, MD, MSc, Manish I. Shah, MD, Tasmeen S. Weik, DrPh, MPH, Joseph L. Wright, MD, MPH, Eddy S. Lang, MDCM, CCFP (EM)

Received June 19, 2013 from Pediatrics and Emergency Medicine, George Washington School of Medicine, Washington, DC, and Emergency Department, Children's National Medical Center, Washington, DC. Revision received August 27, 2013; accepted for publication August 30, 2013.

Co-authors: Charles G. Macias, MD, MPH, Department of Pediatrics, Section of Emergency Medicine, and Center for Clinical Effectiveness, Baylor College of Medicine/Texas Children's Hospital, Houston Texas; Peter S. Dayan, MD, MSc, Department of Pediatrics, Columbia University College of Physicians and Surgeons, New York, New York; Manish I. Shah, MD, Department of Pediatrics, Section of Emergency Medicine, Baylor College of Medicine/Texas Children's Hospital, Houston, Texas; Tasmeen S. Weik, DrPh, MPH, Health Resources and Services Administration/Maternal and Child Health Bureau Rockville, Maryland; Joseph L. Wright, MD, MPH, Department of Pediatrics, George Washington University School of Medicine and Public Health, and Child Health Advocacy Institute, Children's National Medical Center, Washington, DC; and Eddy S. Lang, MDCM, CCFP (EM), Department of Emergency Medicine, University of Calgary, and Alberta Health Services, Calgary, Alberta, Canada.

This research was funded through a Cooperative Agreement between Children's National Medical Center and the National Highway Traffic Safety Administration (DTNH22-09-H-00282). Funding was also provided by the EMS for Children Program of the Maternal and Child Health Bureau, Health Resources and Services Administration.

The views contained in this article are those of the authors and not necessarily those of the National Highway Traffic Safety Administration.

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

The authors acknowledge Rinal Patel (MPH) and Jaclynn Haymon (MPA, RN) for their project management support, and Richard Alcorta as the MIEMSS Project Liaison.

Address correspondence to Kathleen M. Brown, MD, Children's National Medical Center, 111 Michigan Avenue NW, Washington, DC 20010, USA. e-mail: Kbrown@cnmc.org

doi: 10.3109/10903127.2013.844871

ABSTRACT

Background. The burgeoning literature in prehospital care creates an opportunity to improve care through evidence-based guidelines (EBGs). Previously, an established process for the creation of such guidelines and adoption and implementation at the local level was lacking. This has led to great variability in the content of prehospital protocols in different jurisdictions across the globe. Recently the Federal Interagency Committee on Emergency Medical Services (FICEMS) and the National EMS Advisory Council (NEMSAC) approved a National Prehospital Evidence-based Guideline Model Process for the development, implementation, and evaluation of EBGs. The Model Process recommends the use of established guideline development tools such as Grading of Recommendations, Assessment, Development, and Evaluation (GRADE). **Objective**. To describe the process of development of three prehospital EBGs using the National Prehospital EBG Model Process (EBG Model Process) and the GRADE EBG development tool. **Methods**. We conducted three unique iterations of the EBG Model Process utilizing the GRADE EBG development tool. The process involved 6 distinct and essential steps, including 1) assembling the expert panel and providing GRADE training; 2) defining the evidence-based guideline (EBG) content area and establishing the specific clinical questions to address in patient, intervention, comparison, and outcome (PICO) format; 3) prioritizing outcomes to facilitate systematic literature searches; 4) creating GRADE tables, or evidence profiles, for each PICO question; 5) vetting and endorsing GRADE evidence tables and drafting recommendations; and 6) synthesizing recommendations into an EMS protocol and visual algorithm. Feedback and suggestions for improvement were solicited from participants and relevant stakeholders in the process. **Results**. We successfully used the process to create three separate prehospital evidence-based guidelines, formatted into decision tree algorithms with levels of evidence and graded recommendations assigned to each decision point. However, the process revealed itself to be resource intensive, and most of the suggestions for improvement would require even more resource utilization. **Conclusions**. The National Prehospital EBG Model Process can be used to create credible, transparent, and usable

prehospital evidence-based guidelines. We suggest that a centralized or regionalized approach be used to create and maintain a full set of prehospital EBGs as a means of optimizing resource use. **Key words:** Clinical practice guidelines; evidence-based medicine; prehospital care

PREHOSPITAL EMERGENCY CARE 2014;18(Suppl 1):3–14

INTRODUCTION AND BACKGROUND

The Importance of Evidence-based Prehospital Patient Care Protocols

The United States (US) emergency medical services (EMS) system is founded on the principle of delegated practice. Ideally, medical oversight establishes a high standard of patient care, which is then carried out by prehospital providers in the field. Medical direction must ensure that all prehospital providers have the resources to provide optimal and presumably evidence-based care. This includes education and training, support from qualified direct medical oversight physicians, and evidence-based treatment protocols to guide providers' actions.

Evidence-based delivery of prehospital care is becoming increasingly complex as a result of an evolving body of relevant clinical research applicable to the prehospital setting. Similarly, prehospital providers are facing a widening scope of practice as evidence for beneficial prehospital interventions mount, and new technologies are being proposed for use in the prehospital setting. Although the evidence base for prehospital care is growing, many still remain concerned about the lack of high quality evidence for many prehospital interventions. Public expectation and a more careful review of the available evidence would argue for the application of the interventions that have been proven to be effective and discontinuation of those proven to be harmful or wasteful. Unfortunately, even those interventions with proven efficacy, such as care of nontraumatic cardiac arrest[1–5] and use of formal trauma systems for severe trauma,[6–10] are not uniformly applied. In addition, interventions in which the predominance of evidence demonstrates a lack of effectiveness and potential for harm are still being applied in some systems.

The burgeoning literature and resulting opportunities to improve prehospital protocols with evidence-based guidance may be hindered by a lack of established processes to conduct syntheses of available literature around key questions in prehospital care and to develop transparent and well-constructed recommendations linking this evidence to action. Historically, and with few exceptions, prehospital protocols have been developed through locally established procedures and are driven by individual medical directors with variable reliance on research evidence. While some EMS systems have resources that allow for regular careful review of existing protocols and new evidence, many do not. As a result, much of the guidance, and hence performance indicators, risk being driven by expert opinion, as opposed to an unbiased appraisal of the most relevant research. In addition, each EMS system has a unique set of resources and limitations. This may allow individual medical directors to dismiss the results of studies from other systems as not being relevant to their own system. The result of all these factors is that all too often local protocols are based largely on previous practice patterns rather than the most valid and applicable evidence.

Potential Benefits of a Centralized or National Approach to Development of Prehospital Evidence-based Guidelines

Until recently, there has been no established national process for development of evidence-based prehospital patient care protocols. Various jurisdictions, including individual states, regions, and EMS agencies, have been responsible for their own protocol development and implementation. Many prehospital agencies lack the significant resources and expertise needed to conduct a careful review of the available evidence and update protocols on a regular basis. The availability of a nationally accepted set of evidence-based model protocols or evidence-based guidelines (EBGs) would allow individual EMS agencies access to the best available knowledge about the effectiveness of prehospital care practice, but still allow them to adapt this guidance into their own protocols, taking into account local factors that may influence implementation.

In many areas of the US there are multiple EMS agencies operating within a similar geographic area. Each of these agencies may have their own unique protocols. Even jurisdictions that share similar populations and resources may have very different protocols for particular problems. For example, one county in Michigan has 18 different EMS systems with different service models and protocols.[11] In preparation for the pediatric seizure project, the investigative team conducted a review of ten statewide protocols and noted significant variations in the choice of pharmacologic agent for seizure cessation, the route of administration, and the general directions for care of pediatric seizures.

In 2001, the National EMS Research Agenda recommended that EMS professionals apply the evidence from scientific research to improve patient care.[12] Responding to this call to action, in 2006 the Institute of Medicine (IOM) released a report on the Future of Emergency Care in the United States, which called for several specific recommendations to improve prehospital emergency care, including that "the National Highway Traffic Safety Administration (NHTSA) in

partnership with professional organizations convene a panel of individuals with multidisciplinary expertise to develop evidence-based model prehospital care protocols for the treatment, triage, and transport of patients, including children."[13]

The National Prehospital Evidence-based Guideline Model Process

In response to the call for evidence-based model protocols by the National EMS Research Agenda and the IOM, NHTSA convened a national conference on evidence-based guidelines for EMS that was cosponsored by the FICEMS and the NEMSAC. The conference was attended by representatives of EMS stakeholder organizations, who heard presentations by a panel of international experts with extensive, multidisciplinary expertise in EMS, research, and EBGs. Input from conference attendees was used to develop the National Prehospital EBG Model Process for prehospital EBG development, implementation evaluation, and maintenance.[14]

GRADE as a Tool for Prehospital EBG Development

EBGs are systematically developed statements designed to guide health-care practitioners in specific circumstances. They hold the potential to improve care through standardization and adherence to evidence-based guidance. However, the myriad schemes and tools in existence for creating EBGs have served as an obstacle to harmonization in health-care delivery. With over 60 such schemes in existence, the ability to compare across guidelines and define optimal and measurable approaches to care remains elusive.[15]

Most guideline schemas tie the determination of evidence quality to the formulations of recommendations. This may cause those creating the guideline to feel obliged to issue a strong recommendation in favor of a therapeutic intervention supported in randomized trials even if those trials are at risk of bias and the risk benefit or cost considerations are not evident. In addition, guidelines created with some of these schemas often suffer from a lack of explicitness and transparency that helps the user understand how certain judgments around evidence quality or strength of recommendations were determined.

Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) is emerging as a widely accepted process among international guideline development organizations interested in simplifying and providing greater transparency and commonality in EBG creation.[16] The GRADE methodology is particularly well-tuned to an evidence-based medicine approach by stipulating the need to create clear and well-formulated clinical questions that then serve as the focus for the guideline creation. GRADE also strongly suggests the conduct of or reliance upon high quality systematic reviews to inform the EBG development process. GRADE adopts a consistent approach that rates the quality of evidence around predefined and prioritized outcomes of interest into four categories: high, moderate, low, and very low. GRADE conceptualizes evidence quality as the degree of certainty that one has in the estimates of effect reported in the literature. GRADE defines five dimensions of evidence quality that apply to both randomized controlled trials and observational studies: 1) bias resulting from flawed study design or execution, 2) imprecision, 3) inconsistency, 4) indirectness, and 5) publication bias. Concerns in one or more of these domains can result in randomized trials being downgraded to the lowest possible level of quality (i.e., very low).

The GRADE approach also makes it possible for observational studies that begin with a low ranking to climb to high and hence provide potential substrate for strong recommendations. If observational studies report either a convincing effect size or a proportional relationship between intervention and outcome (dose-response) or convincingly account for and adjust confounding factors, they can be upgraded to the highest tier of evidence quality. GRADE also simplifies the process of issuing guidance by defining only two categories of recommendations: strong and weak. Most importantly, GRADE separates the process of defining the quality, and hence the believability and importance of the evidence base, from the process of developing recommendations based on the evidence. Specifically, GRADE provides specific guidance, founded on explicit statements of values and preferences in order to provide context and support for the recommendations that are the endpoint of the process.

Objective of This Report

Under the auspices of the Emergency Medical Services for Children (EMSC) National Resource Center (NRC), we conducted the first test of the EBG model in consultation with NHTSA. This project was funded by HRSA through the Federal EMSC program. It was conducted from February through June of 2009 with an objective to develop a pediatric prehospital protocol. An initial team that included experts in guideline development, prehospital care, and pediatric emergency care was assembled. The team chose pediatric seizure management as the topic for the prehospital guideline. In 2010 NHTSA funded an investigative team to develop and implement two additional prehospital EBGs. The topics for these two guidelines were helicopter transportation of trauma patients (HEMS) and pain management in trauma patients. Both teams

chose to use the GRADE model for development of an EBG.

We describe how we developed three prehospital EBGs. The paper is intended to provide sufficient detail and direction such that a group with sufficient expertise and resources could undertake a similar project in prehospital EBG development adherent to the GRADE framework and consistent with many of the criteria attributed to high quality guidelines in general.[15] We have broken the process of development and initial implementation of an evidence-based guideline into six steps. For each step we present the rationale, methods, lessons learned, and potential solutions.

STEP 1: ASSEMBLING THE EXPERT PANEL AND GRADE TRAINING

Rationale

The investigative team chose to use the GRADE tool for development of the EBGs for both projects. The team felt that GRADE was particularly well suited for the creation of prehospital EBGs for many of the reasons discussed above. Specifically, GRADE allows the guideline development team to consider factors other than study design when determining the strength of recommendations. For example, evidence for prehospital interventions is often derived from studies that were conducted in another environment, such as an emergency department (ED) or inpatient setting. GRADE allows these studies to be "downgraded" based on the degree of "indirectness" when appropriate. Similarly, the prehospital literature contains many observational trials. GRADE allows for these studies to be "upgraded" when they meet appropriate criteria. Finally, GRADE allows for the consideration of the balance between benefits and harm, patient values and preferences, as well as resource considerations when determining the strength of a recommendation. This was felt by the team to be particularly important when considering how a strong or weak recommendation would be applied in the context of an individual EMS system.

The investigative team also chose to assemble a distinct expert panel for each of the EBG topics. It was felt that the guideline development team should include content experts as well as experts on EBG development and use of the GRADE methodology. Content experts included experts on prehospital care and the specific topic of each EBG; e.g., we selected an expert on pain management for the pain management EBG. None of the content experts had significant experience with use of the GRADE tool and therefore it was recognized that they would need training and guidance in its use.

Methods

Assembly of the EBG development panel differed for the pediatric seizure versus helicopter EMS (HEMS) and pain management EBGs based on the structure of the grant funding. For pediatric seizure, the mandate was to develop a pediatric EBG with the exact topic unspecified while for HEMS and pain management, a grant application had preselected the topics. Thus, for seizure, an EBG panel with general expertise in pediatrics, pediatric trauma, pediatric emergency medicine, and prehospital care was selected. For HEMS and pain management, a panel with general expertise in trauma and emergency medicine were included but members with specific expertise in the topic areas were also selected, including members who had specific pediatric expertise. Both panels included clinical practice guideline (CPG) and GRADE methodologists, a health information specialist to assist in literature searching, federal partners to observe the process, an EMS medical director, and a ground-level EMS provider.

A core working group developed the work plan for the project and identified and reached out for commitments from the expert panel members. All members were paid a nominal stipend for their commitment. All members understood that the stipend would not remunerate their actual labor costs but were willing to donate their time to experience the intellectual novelty of the project.

Key literature pertaining to the GRADE process was distributed to panel members, and a teleconference/webinar, led by the GRADE methodologist, was held to review some of the core concepts around the GRADE methodology. Sample GRADE tables and literature appraisal presentations were posted on a shared website for review by the panel members. Panel members were encouraged to consult with and direct questions to the methodologist throughout the entire process.

Lessons Learned

There was considerable overlap between the two investigative teams; therefore, the second team was able to apply some of the lessons learned during the first project (pediatric seizure) to the HEMS and pain management project. Feedback from the panelists' on the first project focused on two areas: the need for more experience/training with use of the GRADE framework, and the significant time commitment required by each panel member, even after they were comfortable with the GRADE tool.

The team attempted to address these concerns when planning the second project. The second team was given more examples of completed GRADE tables and specific guidance in use of GRADE. They were also

given access to a medical librarian who performed a literature review for each of the questions. Panelists also suggested that assigning a single panelist per clinical question (i.e., for each patient, intervention, comparison, and outcome (PICO)) was suboptimal and somewhat isolating, and most expressed an interest in partnering with other panelists on future efforts even if this implied working on more than one PICO question. For HEMS and pain management, two panelists were assigned to each question with the original intent of providing redundancy to assure review completeness. However, all panelists collaborated and divided the tasks between them. Although some redundancy did occur, the relative paucity of evidence for all prehospital care limited the need for duplication. Partnerships increased individual panel member comfort with GRADE.

Potential Solutions

There was a steep learning curve for use of the GRADE tool. A potential solution for future EBG development is to recruit at least some members of the team who have experience with use of the tool. Intensive training should be given to those panel members who have not had experience using GRADE. Partnering two experts on a single PICO question was helpful. It is important that the panelists be made aware of the significant time commitment involved and be compensated appropriately for their time.

Many of the panel members felt that, given the time commitment and level of expertise required, the process would be difficult to replicate at the local level. Many panelists suggested a model in which there is a centralized group of experienced experts charged with prioritizing and developing prehospital EBGs that could then be disseminated and contextualized by individual EMS agencies or systems.

STEP 2: DEFINING THE EBG CONTENT AREA AND ESTABLISHING THE SPECIFIC CLINICAL QUESTIONS THAT NEED TO BE ADDRESSED

Rationale

Formulating clinical questions that are focused, clearly explicit, and consistently structured constitute a cornerstone of the evidence-based methodology. EBGs that rely on the GRADE approach are no exception. In fact, the outcome-focused nature of the GRADE approach makes the process of question formulation especially critical. The comprehensiveness and focused importance of the questions addressed in any clinical practice guideline effort are one of the key metrics by which the entire process is evaluated.

The panels adopted PICO formatting for EBG development.[17] Implicit in the structure of the PICO framework is that a clinical question must specify the patient or patient population of interest, the intervention under consideration, usually experimental or differing from standard practice, as well as a comparison option. Defining the outcomes of interest and categorizing their importance as being critical or otherwise is an especially important element of the GRADE methodology. In selecting and categorizing outcomes, the panel should place particular emphasis on what would be deemed important outcomes at the patient level, as opposed to surrogate outcomes, even if the evidence base is expected to preferentially inform the latter. The selection of outcomes will have implications as to the categorization of the larger evidence base informing each PICO question. For example, if a critical outcome, as determined by the guideline developers, is supported by what is deemed to be very low quality evidence based on GRADE criteria, the entire evidence base will be drawn down to the very low category even if other outcomes are informed by high quality evidence.

PICO questions fall into four primary clinical domains—therapy, diagnosis, prognosis, and harm. Currently, GRADE guidance is limited to therapy/harm and diagnosis-related dimensions of care. GRADE was used to address both therapy/harm and diagnosis questions in these EBGs, and this approach has been used by others. GRADE continues to evolve a conceptual framework that incorporates diagnosis questions.[17]

Methods

Once the topic of the EBGs had been established, the next task facing the working group was to determine what kind of specific guidance the EBG would attempt to provide. At this point, a series of both structured and unstructured questions were developed by the core working groups of these three EBG projects in a manner that would align with what was considered to be the greatest needs of frontline providers as well as medical directors and other decision-makers within EMS systems. The list of PICO questions was revised on a number of occasions as wider input was gathered as to their importance and relevance.

The slate of PICO questions that each EBG would address was achieved through an iterative process that sought to manage overlap between any two similar PICO questions and to prioritize those answerable queries that would have the greatest impact on patient care through evidence-based guidance. While most PICO questions in these EBGs followed a "therapy" framework, a number of diagnosis issues arose. These included the use of specific point of care tests, such as glucometry in seizing children, or the use of various pain scales to guide analgesia administration. In each of these circumstances, the approach taken in question formulation was to develop a gold standard

approach to informed decision making using rigorous scientific methodology and patient-important outcomes as opposed to measures of test accuracy or instrument validation. As a result, despite the expected paucity of relevant evidence, diagnosis questions were formulated in such a way as they would be informed by a randomized controlled trial comparing two competing diagnostic strategies.[18]

After the final list of PICO questions was agreed upon by the panelists, they were assigned either individually or in pairs for further development. This would include additional question refinement with the opportunity to have other panelists provide feedback on modifications to any of the PICO components being proposed.

Lessons Learned

It was necessary to look at existing protocols to ensure that a minimum amount of evidence existed on the topic prior to conducting searches, that current evidence was not being applied in clinical care, and that variability in care existed among various protocols. Since focused and structured clinically relevant questions must drive guideline development, collaboration of key stakeholders is also essential when developing and delegating PICO questions to EBG panelists.

One of the key skills that the panelists on these projects needed to hone early on in the EBG development process was the development and refining of a clinical question optimized to serve as the foundation for all subsequent stages of EBG development. The paramount importance of this stage was evidenced by many the issues that arose during the subsequent steps. Time spent developing and researching PICO questions that did not turn out to be critical decision steps in the final algorithm may represent unnecessary resource utilization in this resource intense project.

Panelists were given the opportunity to search for and evaluate the evidence base pertaining to their question prior to the categorization of outcomes as being critical, important, or relevant. While we do not believe that this influenced the process, we are cognizant that, according to recent GRADE guidance, the process of selecting and prioritizing outcomes within each PICO question should ideally be undertaken prior to a formal analysis of the existing scientific literature.[19]

Potential Solutions

Future prehospital EBG development efforts should emphasize the paramount importance of this step in order to focus resources on the critical questions that will populate the final algorithm. Development of a draft algorithm to be used as reference by the group prior to development of clinical questions and finalizing PICO questions may be helpful.

In general, panelists agreed that working in tandem on PICO questions held certain advantages in terms of comparing notes on related questions and avoiding duplication of efforts. Panelists suggested that an in-person group meeting would have been helpful at this stage to clarify the PICO questions and discuss the categorization of outcomes.

STEP 3: LITERATURE SEARCHES AND PRIORITIZATION OF OUTCOMES

Rationale

A key component of the GRADE approach is the conduct of a systematic literature search to evaluate the quality of evidence in addressing the outcomes defined in the PICO questions.[18–24] Unique to the GRADE approach is considering these defined outcomes and ranking their importance. As noted above, ideally the comprehensive literature searches should be performed after the prioritization of outcomes has been established for each PICO question.[19]

Methods

Panelists were instructed to consider the various outcomes that pertained to their PICO question and categorize these as being critical, important, or otherwise, all predominantly from the perspective of patients and their families, but with consideration of provider and health-care system perspectives as well.

Panel members performed their literature searches to address their assigned PICO questions. A health information specialist assigned to this project with expertise in systematic literature search processes provided technical support to panel members to ensure consistency and resource availability in accessing relevant literature. In the first iteration of the project (pediatric seizure) the medical librarian was available as a resource but did not perform literature searches for the providers unless asked. During the second project (HEMS and pain management EBGs), the medical librarian was asked to perform literature search for each PICO question. The results of these searches were made available to all of the panelists on a shared website.

Panelists reviewed the search conducted by the medical librarian and selected the relevant literature. In some cases, they also supplemented the search with their own search efforts. Published papers available in Medline, OVID, and the Cochrane Clinical Trials Registry were identified. In addition, and for the HEMS project in particular, panelists were encouraged to also consider other sources of evidence, such as the guidelines of national organizations and the "gray literature." Panelists recorded search strategies and cataloged them as part of the supporting documentation

yet formal systematic review methodology was not followed. Support from the core working group, particularly from content experts, was also made available to review and offer suggestions toward search strategies and the process of including relevant literature for appraisal.

Lessons Learned

Using consistent literature search and appraisal methodology with health information specialist support is necessary to ensure that the literature is acquired and evaluated in a uniform manner. This is a time- and labor-intensive process. It was evident from the first project, involving seizure guideline development, that having 1 month for panelists to conduct their literature searches and appraisals was not sufficient. Therefore, this was increased to 3 months for the pain and helicopter EMS (HEMS) guidelines. In addition, the increased availability and early use of the medical librarian in the second was appreciated by the panelists.

During the process of searching for relevant literature, the EBG panel experienced the challenge of finding prehospital-based (or pediatric-specific prehospital-based studies) studies to address the PICO questions. This is consistent with prior studies defining the gap in pediatric prehospital research, despite research priorities that have been defined.[11,25,26] As a result, the EBG panel expanded its search strategies to include studies conducted in the ED and other settings and, in the case of the pediatric prehospital seizure EBG studies, conducted in the prehospital setting in adults. This also influenced the appraisal of the strength of evidence, as discussed below.

Potential Solutions

Ideally, outcome measures should be prioritized and agreed upon by the group prior to the conduct of the individual literature searches. As noted in step 2, it may be helpful to conduct a group meeting to discuss finalization of PICO questions and prioritization and selection of critical outcomes prior to asking the individual panelists to conduct their searches and evaluation of the literature.

Adequate time and preparation should be allotted to allow panelists to conduct a thorough and inclusive literature search. The employment of a medical librarian to assist the expert panel members in their literature reviews was felt to be an essential element by the panelists. Until the quality of the evidence base for prehospital care improves, it will be necessary to keep literature search strategies broad to encompass studies that may include populations or settings that are not specific to the PICO question.

STEP 4: CREATION OF EVIDENCE PROFILES (GRADE TABLES) FOR EACH PICO QUESTION AND PRESENTATIONS SUMMARIZING THE COMPLETED WORK AND DRAFT RECOMMENDATIONS

The GRADE process emphasizes the use of formatted evidence profiles (EPs) to organize, summarize, and communicate the quality of evidence, the effect size for each outcome, and the overall grading of the evidence. EPs provide consistency, explicitness related to judgments made, and transparency for the guideline end-users. Recommendations are then derived by considering the strength of the evidence, the importance of the outcomes addressed, and the values and preferences of patients, providers, and systems.[27]

The GRADE approach involves upgrading or downgrading evidence quality based on an appraisal of criteria such as the study design, quality, consistency, directness, potential bias, and precision of the findings.[18,20,24] The GRADE approach classifies evidence into 4 categories for a given recommendation (high, moderate, low, or very low) and classifies recommendations into 2 categories (strong or weak).[18,20–24] The quality of evidence, variability in values and preferences, and implications to patients and stakeholders influenced the strength of the recommendation.

Methods

To assess and grade the evidence, individual panelists created EPs for each PICO question. Panelists were given the option of using dedicated software for the purpose of creating the EPs; however, most created their own version based on examples provided. Individual working group members were encouraged to seek guidance from content experts within the panel for questions on the process of GRADE-based evidence appraisal and EP completion.

The EP for each PICO question included an appraisal of each relevant article using GRADE criteria. Panel members ranked the quality of each article into one of the four evidence quality categories, and then upgraded or downgraded the quality rating based on specific and predefined criteria. The initial level for the quality of evidence was determined by the study design, such that a randomized trial would be high quality, an observational study would be low quality, and anything else would be very low.[20] In keeping with the GRADE approach, moderate quality evidence was also a potential rating level for quality. For example, the quality of an observational study was upgraded to moderate if there was a large magnitude of effect, there was a dose-response gradient, or all possible confounders would not have reduced the effect. The quality of an RCT was downgraded if there were

limitations in study design or execution, inconsistency or imprecision of results, indirectness of evidence, or publication bias.[20,21]

Each EP included an assessment of evidence for each outcome with attention to those deemed critical and important.[27,28] For example, in the development of the seizure management guideline, a panelist completed evidence evaluations and profiles separately for seizure cessation within 10 minutes and recurrence of seizure within 1 hour of antiepileptic medication administration. Panelists were asked to provide an effect size and measure of precision for each outcome. Given that the types of PICO questions varied (e.g., diagnosis and therapy), the nature of study designs and reported results also varied. For example, absolute risk differences and relative risks with 95% confidence intervals (CIs) were provided for therapy questions, while risks with 95% CIs were provided for prognosis questions. Within each EP, reviewers provided the number of included studies, their designs, and qualitative assessments of factors that had an impact on the quality of the evidence. Typical factors assessed included study limitations (i.e., threats to validity), directness (e.g., whether population was prehospital patients), consistency of results across studies, precision of effect estimates (mainly confidence intervals), and publication bias.[29–33]

To provide an overall evidence grade, panelists started from the vantage point of considering RCTs as high quality and observational studies as low quality evidence, as recommended by the GRADE originators.[34] The evidence grade (high, moderate, low, or very low) was then raised or lowered based on the assessment of the factors that impacted the evidence quality.[35]

After completing the evidence profiles, panelists worked iteratively with GRADE experts in the working group to create draft recommendations addressing the PICO questions. In addition to the evidence strength, the recommendations were developed as a function of prioritizing multiple outcomes from the perspective of specific values and preferences of patients, providers, and the health-care system. Each panelist developed values and preferences statements in support of their recommendations. Particularly relevant and important to our project, the GRADE methodology allows modifications in the strength of recommendations based on resource limitations.

Consistent with GRADE methodology, recommendation strength ratings (strong or weak) were made with consideration of the balance between desirable and undesirable effects, such that a large difference between a desirable and undesirable effect warranted a strong recommendation either for or against.

Lessons Learned

Perhaps the most important lesson learned regarding using GRADE methodology to summarize and grade evidence and develop recommendations was the amount of training required by panelists to comprehensively understand all processes in order to produce more uniformity in EPs, estimates of effect, and grading. Additionally, the evidence available for specific circumstances did not always lend itself to using the available EP formats. For example, in the protocol to address pain management, the group created an alternative format to assess the evidence regarding the use of pain scales. As pain measurement has no reference standard, the available literature provided data regarding tool validity and reliability but not test characteristics. Finally, few systematic reviews were available to provide overall effect sizes; future efforts would require intense resources to be able to pool existing data from individual studies.

Potential Solutions

As noted in the sections above, intensive training of panelists is essential, particularly for those new to the GRADE tool. In addition to training, panelists need guidance during the process of literature appraisal. Since the available literature may not always lend itself to the standard GRADE format, real-time consultation with a GRADE methodologist may be necessary even for experienced panel members.

The labor-intensive review process was necessary in large part because of the lack of systematic literature reviews that focus on prehospital care. If more prehospital-focused systemic reviews were available, the labor required form the panelists would be significantly diminished.

STEP 5: VETTING AND ENDORSING GRADE EVIDENCE TABLES AND DRAFT RECOMMENDATIONS

Rationale

Our development plan included an in-person meeting of the full panel to vet the work done by the individual panelists or panelist teams, and reach consensus on the draft recommendations that would populate the final decision tree algorithm. The full panel consisted not only of the experts who reviewed and appraised the literature but also local EMS leadership, the GRADE methodologists, and others with topic-specific expertise. We felt that it was essential that all be present for discussion of the literature reviews and draft recommendations in order to ensure that the final recommendations reflected the input and expertise of the full

panel and, in the case of the second project, to facilitate implementation in the Maryland EMS system.

Methods

For each project, a two-day in-person meeting was planned. The meeting of the pediatric seizure panel included the experts who reviewed the literature, which included experts with knowledge of pediatrics, emergency medicine, and prehospital care. In addition, the GRADE methodologist was present. The meeting for the HEMS and pain management team included all those noted above plus members of the Maryland EMS system leadership who provided their expert input. Additionally, as this project included implementation of the final guidelines into the Maryland EMS system we felt it was important that they participate in the process of developing the guidelines.

During the meeting, each of the panel members assigned a PICO question presented their literature review, assessment of the quality of the evidence using GRADE criteria and presentation of draft recommendations as well as supporting values and preferences statements. The group then discussed the information provided and recommendations, resulting in some cases to changes in the assessment of the level of evidence and the strength of the recommendations. We were prepared to use a modified Delphi method to resolve any specific topics of contention. However, this was not necessary as the group was able to come to consensus after discussion.

During the meeting, areas that needed further review were identified. Panelists responsible for the related PICO questions were asked to conduct that review with the assistance of the investigative team and GRADE methodologists. If there were questions about changing the level of evidence or strength of a recommendation, these were vetted by conference calls and email discussions among the full group.

Lessons Learned

The in-person meeting was considered essential by all involved. Including EMS medical directors and field providers in the discussion was also essential and did influence the discussion. One of the advantages of GRADE is the ability to utilize contextual EMS factors to influence the strength of recommendations. For pediatric seizures, contextual factors such as the acceptability of rectal diazepam, the training of EMS providers in administering nasal midazolam, and parental distress at watching a child seize were discussed and significantly impacted the final EBG. For pain management, the perceived reluctance of prehospital providers to begin pain management resulted in upgrading the strength of recommendation to assess pain.

Potential Solutions

The in-person meeting of the full expert panel as well as the participation of EMS providers is essential to the process. Every panelist indicated in their response to the written survey that they felt this part of the process was the most successful element. In fact, panelists expressed a desire to have more in-person meetings earlier in the process (see above). Inclusion of the local EMS leadership was appreciated by that group, influenced the discussion, and facilitated our ability to implement the guideline in a working EMS system.[36]

STEP 6: SYNTHESIZING COLLECTIVE WISDOM INTO AN EMS PROTOCOL AND VISUAL ALGORITHM

Rationale

The creation of decision trees and algorithms was an important and concrete objective for this project since it is the format most EMS agencies use to present their protocols.

Methods

During the in-person meetings, after all presentations had been completed and revisions to the draft work had been discussed and incorporated, the recommendations were transformed into an algorithm that reflects the typical visual format prehospital providers often use. To assist panel members with conceptualizing how their PICO question fit into the larger picture, a visual diagram of an EMS protocol was populated in real time during the meeting and displayed for panelists to comment on and edit. The visualization also identified questions that were either unanswered or not adequately answered as panelists followed the logical progression of decision making in the protocol. Thus, the process was iterative and the first version underwent a number of refinements to achieve a workable and easy to follow management algorithm. After the meeting, some additional areas for literature review were identified and completed. In most cases these did not result in changes to the recommendations. However, when revisions were made to the recommendations panelists were asked to comment and consensus was reached either over email or, when needed, through additional group conference calls.

The draft algorithm with recommendations and level of evidence was presented to the Maryland EMS protocol review committee for consideration during

their annual protocol review process. Feedback from that group indicated that the algorithm format was helpful.

Lessons Learned

The creation of the algorithm formatted "model protocol" was appreciated by the end-users. This step was important to conduct in real time to ensure there was consensus regarding the final GRADE recommendation. Since the clinical questions were chosen and the PICO questions developed before the full literature review and before the creation of the algorithm, in some cases the PICO questions that were assessed did not end up in the final algorithm. As noted above, the creation of the algorithm in real time at the meeting in some cases resulted in the need for additional literature reviews and appraisals.

Potential Solutions

Creation of a draft algorithm or decision tree prior to drafting the PICO questions may decrease the likelihood that experts will be assigned to questions that are not included in the final algorithm. However, as noted above in step 2, this could potentially bias the development of the PICO questions.

In addition to including key stakeholders on the expert panel, having adequate and broad representation of medical directors and prehospital providers from disparate potential EMS agencies that might implement the guidelines will enhance consideration of the field implications of guideline implementation.

Gathering Feedback on the Process

Rationale

As these projects represented the first attempts to use both the National Prehospital EBG Model Process with incorporation of GRADE for the development of prehospital EBGs, we felt it was important to collect feedback from those involved with the process.

Methods

Feedback was solicited verbally from the first (pediatric seizure) panel at the in-person meeting, and then subsequently during conference calls as the EBG and algorithms were being finalized. The second panel was asked to fill out a written survey at the conclusion of the in-person meeting and were given the opportunity to provide feedback as the EBG and algorithms were being finalized. The second panel was also encouraged to provide feedback after the meeting during the iterative revision process.

Lessons Learned

We did not ask experts for feedback on the process prior to the in-person meeting; therefore, we can only estimate the amount of time that they dedicated to this process. Reports from the written survey after the second team meeting ranged from 24 to 80 hours per expert. This did not include the considerable amount of time spent by those who needed to do additional work immediately after the meeting and the work needed to be completed by all in order to draft the final manuscripts.

Potential Solutions

Future efforts to trial the National Prehospital EBG Model Process should include feedback from participants at each step of the process so that total amount of effort can be tracked in more detail.

Summary and Future Directions

We have presented a description of the first three initiatives in guideline development using the National Prehospital EBG Model Process and the GRADE development tool to create a prehospital EBG. The specifics of each of those EBGs and the outcome of the implementation process will be described in separate papers. The six-step process that we have outlined resulted in the development of three specific prehospital EBGs that are formatted for use by EMS agencies as a guideline for creation of their local prehospital protocols. Each of the steps described was essential to the creation of a final product that is credible, transparent, and adaptable.

In trialing this process, we have made observations and provided lessons learned that we hope will help guide future efforts to create evidence-based prehospital guidelines. In general, the panelists were positive and supportive of the innovations inherent to the Model Process and the use of GRADE for development of prehospital EBGs. They cited the structured approach for evaluating the literature and the ability to use values and preferences to "contextualize" the strength of the final recommendations as benefits of the GRADE model. Further, given the paucity of available evidence for prehospital interventions, GRADE allowed for the assignment of a level of strength for each recommendation to complement the level of evidence.

The panelists' major concerns included suboptimal preparation and training in the GRADE methodology at the point at which they undertook their project. They acknowledged that their comfort level only increased with experience. Having a GRADE methodologist to consult real time during each step was essential. Lastly, panelists were concerned about the individual and

collective effort required to produce a single EBG. The estimated overall (before and after the in-person meeting) time spent on the project was more than 100 hours for each panelist. This time commitment would be significantly reduced if more relevant systematic reviews were available. The majority of this time was a volunteer commitment. Most of the suggestions for improvement would require even more resource utilization.

Because the process proved to be extremely resource intense, few state and local EMS agencies would have the expertise or resources available to execute the process with fidelity. The use of GRADE in the National Prehospital EBG Model Process allows for the development of national guidelines that can be customized for local values and preferences. We suggest that future efforts be centralized at a national level or regionalized in centers of excellence.

References

1. Valenzuela TD, Roe DJ, Nichol G, Clark LL, Spaite DW, Hardman RG. Outcomes of rapid defibrillation by security officers after cardiac arrest in casinos. N Engl J Med. 2000;343:1206–9.
2. Zheng ZJ, Croft JB, Giles WH, Mensah GA. Sudden cardiac death in the United States, 1989 to 1998. Circulation. 2001;104:2158–63.
3. Eisenberg MS, Horwood BT, Cummins RO, Reynolds-Haertle R, Hearne TR. Cardiac arrest and resuscitation: a tale of 29 cities. Ann Emerg Med. 1990;19:179–86.
4. Nichol G, Stiell IG, Laupacis A, Pham B, De Maio VJ, Wells GA. A cumulative meta-analysis of the effectiveness of defibrillator-capable emergency medical services for victims of out-of-hospital cardiac arrest. Ann Emerg Med. 1999;34: 517–25.
5. Cummins RO, Ornato JP, Thies WH, Peppe P E. Improving survival from sudden cardiac arrest: the "chain of survival" concept: a statement for health professionals from the Advanced Cardiac Life Support Subcommittee and the Emergency Cardiac Care Committee, American Heart Association. Circulation. 1999;1832–47.
6. Bass RR, Gainer PS, Carlini AR. Update on trauma system development in the United States. J Trauma. 1999;47(3 suppl):S15–S21.
7. Mann NC, Mullins RJ, MacKenzie EJ, Jurkovich GJ, Mock CN. Systematic review of published evidence regarding trauma system effectiveness. J Trauma. 1999;47(3 suppl):S25–S33.
8. MacKenzie EJ. Review of evidence regarding trauma system effectiveness resulting from panel studies. J Trauma. 1999;47(3 suppl):S34–S41.
9. Jurkovich GJ, Mock C. Systematic review of trauma system effectiveness based on registry comparisons. J Trauma. 1999;47(3 suppl):S46–S55.
10. Mann NC. Assessing the effectiveness and optimal structure of trauma systems: a consensus among experts. J Trauma. 1999;47(3 suppl):S69–74.
11. Committee on the Future of Emergency Care in the United States Health System. Board of Health Care Services. Emergency Medical Services: At the Crossroads. Washington, DC: National Academy Press; 2006.
12. National Highway Traffic Safety Administration. National EMS Research Agenda. Washington, DC: US Department of Transportation; 2001. Available at www.ems.gov/pdf/EMSResearchAgenda.pdf. Accessed February 4, 2013.
13. Committee on the Future of Emergency Care in the United States Health System. Board on Health Care Services. Emergency Care for Children: Growing Pains. Washington, DC: National Academy Press; 2006.
14. Lang ES, Spaite DW, Oliver ZJ, Gotschall CS, Swor RA, Dawson DE, Hunt RC. A national model for developing, implementing, and evaluating evidence-based guidelines for prehospital care. Acad Emerg Med. Feb 2012;19(2):201–9.
15. Consensus Report, Committee on Standards for Developing Trustworthy Clinical Practice Guidelines; Institute of Medicine. Clinical practice guidelines we can trust. March 23, 2011. www.iom.edu/Reports/2011/Clinical-Practice-Guidelines-We-Can-Trust.aspx
16. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. Apr 26 2008;336(7650):924–6.
17. Goat GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schünemann HJ. GRADE guidelines, 2: framing the question and deciding on important outcomes. J Clin Epidemiol. 2011 Apr;64(4):395–400. Epub 2010 Dec 30.
18. Jaeschke R for the GRADE Working Group. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. BMJ. 2008;337:327–30.
19. www.agreeetrust.org Accessed February 4, 2013.
20. Brozek JL Akl EA, Alonso-Coello P, Lang D, Jaeschke R, Williams JW, Phillips B, Lelgemann M, Lethaby A, Bousquet J, Guyatt GH, Schünemann HJ; GRADE Working Group. Grading quality of evidence and strength of recommendations in clinical practice guidelines, part 1 of 3: an overview of the GRADE approach and grading quality of evidence about interventions. Allergy. 2009; 64: 669–77.
21. Brozek JL Akl EA, Jaeschke R, Lang DM, Bossuyt P, Glasziou P, Helfand M, Ueffing E, Alonso-Coello P, Meerpohl J, Phillips B, Horvath AR, Bousquet J, Guyatt GH, Schünemann HJ; GRADE Working Group. Grading quality of evidence and strength of recommendations in clinical practice guidelines, part 2 of 3: the GRADE approach to grading quality of evidence about diagnostic tests and strategies. Allergy. 2009;64:1109–16.
22. Guyatt GH Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, Vist GE, Schünemann HJ; GRADE Working Group. GRADE: incorporating considerations of resources use into grading recommendations. BMJ. 2008;336:1170–3.
23. Schunemann A, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, Williams JW Jr, Kunz R, Craig J, Montori VM, Bossuyt P, Guyatt GH; GRADE Working Group. GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ. 2008;336:1106–10.
24. Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) Working Group. Grading quality of evidence and strength of recommendations. BMJ. 2004;328: 1490–4.
25. Sayre MR, White LJ, Brown LH, McHenry SD; National EMS Research Strategic Plan Writing Team. The National EMS Research Strategic Plan. Prehosp Emerg Care. 2005;9(3):255–66.
26. Foltin GF Dayan P, Tunik M, Marr M, Leonard J, Brown K, Hoyle J Jr, Lerner EB; Prehospital Working Group of the Pediatric Emergency Care Applied Research Network. Priorities for pediatric prehospital research. Pediatr Emerg Care. 2010;26:773–7.
27. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. GRADE guidelines, 1: introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011 Apr;64(4):383–94. Epub 2010 Dec 31.

28. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schünemann HJ. GRADE guidelines, 2: framing the question and deciding on important outcomes. J Clin Epidemiol. 2011 Apr;64(4):395–400. Epub 2010 Dec 30.
29. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW Jr, Atkins D, Meerpohl J, Schünemann HJ. GRADE guidelines, 4: rating the quality of evidence—study limitations (risk of bias). J Clin Epidemiol. 2011 Apr;64(4):407–15. Epub 2011 Jan 19.
30. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, Williams JW Jr, Meerpohl J, Norris SL, Akl EA, Schünemann HJ. GRADE guidelines, 5: rating the quality of evidence—publication bias. J Clin Epidemiol. 2011 Dec;64(12):1277–82. Epub 2011 Jul 30.
31. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G, Jaeschke R, Williams JW Jr, Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schünemann HJ. GRADE guidelines, 6: rating the quality of evidence—imprecision. J Clin Epidemiol. 2011 Dec;64(12):1283–93. Epub 2011 Aug 11.
32. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA, Norris S, Vist G, Dahm P, Shukla VK, Higgins J, Falck-Ytter Y, Schünemann HJ; GRADE Working Group. GRADE guidelines, 7: rating the quality of evidence—inconsistency. J Clin Epidemiol. 2011 Dec;64(12):1294–302. Epub 2011 Jul 31.
33. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M, Schünemann HJ; GRADE Working Group. GRADE guidelines, 8: rating the quality of evidence—indirectness. J Clin Epidemiol. 2011 Dec;64(12):1303–10. Epub 2011 Jul 30.
34. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, Atkins D, Kunz R, Montori V, Jaeschke R, Rind D, Dahm P, Akl EA, Meerpohl J, Vist G, Berliner E, Norris S, Falck-Ytter Y, Schünemann HJ. GRADE guidelines, 11: making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol. 2013 Feb;66(2):151–7. Epub 2012 Apr 30.
35. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, Atkins D, Kunz R, Brozek J, Montori V, Jaeschke R, Rind D, Dahm P, Meerpohl J, Vist G, Berliner E, Norris S, Falck-Ytter Y, Murad MH, Schünemann HJ; GRADE Working Group. GRADE guidelines, 9: rating up the quality of evidence. J Clin Epidemiol. 2011 Dec;64(12):1311–6. Epub 2011 Jul 30.
36. Brown KM, Hirshorn JM, Alcorta RA, Weik TS, Lawner B, Ho S, Wright J. The implementation and evaluation of an evidence-based statewide prehospital pain management protocol developed using the National Prehospital Evidence-based Guideline Model Process for Emergency Medical Services. Prehosp Emerg Care. 2014; in press.