Evidence-based Practice Center Systematic Review Protocol

Project Title: Field Triage Guideline Revision: Glasgow Coma Scale: Systematic Review

## I. Background and Objectives for the Systematic Review

Emergency Medical Services (EMS) providers must rapidly triage care for individuals who have undergone trauma in challenging environments. Field triage helps to determine the severity of injury and identify patients with poorer prognosis and affects how and where patients are transported and treated.[1,2] Therefore, EMS providers must have assessment tools that are easy to use, reliable, accurate, and effective. A key component of field triage for patients with suspected blunt head trauma is level of consciousness assessment, which is associated with prognosis. The Glasgow Coma Scale (GCS)[3,4] is a widely-used instrument to assess consciousness at the site of injury, in emergency departments, and in hospitals to monitor progress or deterioration during treatment.[5] The GCS consists of three categories of responses: eye, verbal, and motor. Total Glasgow Coma Scale (tGCS) scores range from 3 to 15 with lower scores for lower levels of consciousness, generally correlating with more severe injury requiring more intensive care. Scores of 3 to 8 are considered to denote severe head injury, 9 to 12 moderate and 13 to 15 mild.[6] The 2011 field triage guidelines from the Centers for Disease Control and Prevention (CDC) National Expert Panel recommend transferring patients with tGCS scores of ≤13 to facilities providing the highest level of trauma care.[7]

In circumstances when a total score cannot be accurately obtained (e.g., trauma victims who are intoxicated, intubated, or whose other injuries influence response), the motor component alone (mGCS) is sometimes used.[3,8-10] mGCS scores less than or equal to 5 are considered to denote severe injury.[11,12]

There are a number of challenges in evaluating the comparative performance and effectiveness of the tGCS and the mGCS. An important overarching challenge is the need to understand, consider, and address applicability, that is, the extent to which evidence obtained in a certain population and setting can be used to inform more general conclusions about a broader range of situations. Applicability is impacted by the acute and emergent nature of traumatic injuries, which occur in heterogeneous settings that are often chaotic in nature (e.g., motor vehicle crashes, assault, sports, falls), as well as differences in patient populations related to demographics, presence and severity of intoxication, and medical comorbidities. Triage must be performed by first responders or other personnel with different levels of training and certification (e.g., Emergency Medical Responder, Emergency Medical Technician [EMT], EMT-Intermediate, Advanced EMT, and Paramedic[13]). An assessment must work in a wide range of potential scenarios and overcome the difficulties in measuring and controlling for the many variables that may influence the tGCS and mGCS performance and ultimately patient outcomes. For example, the GCS was originally developed for use in patients with suspected traumatic brain injury (TBI). However, in clinical practice, the GCS is often used for field assessment of any traumatic injury, and research studies often do not separate TBI from trauma without brain injury. In addition, trauma is often complicated

Source: www.effectivehealthcare.ahrq.gov
Published online: February 25, 2016

by the co-occurrence of multiple injuries, and by variations in mechanisms of injury. It is unusual for TBI to occur in isolation[14] and as a result, first responders may need to attend to more immediate life support concerns before assessing for TBI. Delay and lower prioritization of TBI evaluation can contribute to variation in assessment, decisionmaking, and treatment.[15] Lack of data from bypassed or delayed TBI assessment may also increase the risk of selection bias in field studies and limit the usefulness of evaluations of the tGCS and mGCS. Performance of the scales also may vary across different types or mechanisms of injury. For example, the nature and prognosis of a TBI sustained from an impact injury (blunt force which may or may not involve fracture or intracranial lesion) may be different from that of a TBI sustained from an acceleration/deceleration injury[16] (diffuse injury resulting from contrecoup forces). Given this, performance of the same field triage assessment instrument may differ. Current evidence about field assessment of TBI frequently relies on extrapolation from studies conducted in emergency departments, as this environment is more controlled and easier to study.[17] However, the performance of the tGCS and mGCS may be different when administered soon after injury (in the field) as opposed to later (after destination decisions have been made or after arrival in the emergency department). Outcomes also could be influenced by time of assessment, as an accurate earlier assessment could lead to a more appropriate destination decision, impacting subsequent treatment decisions.

The ultimate goal of selecting one risk prediction instrument over another is to improve clinical outcomes. However, information on clinical outcomes is often lacking and decisions about their use must often be based on how they perform on intermediate measures such as predictive utility. Assessment of predictive utility is complicated by the need to understand issues related to instrument development and patient evaluation over time (longitudinally).[18] Intervening factors (e.g., types of treatments received) can impact the outcomes that the instruments are attempting to predict, and decisions regarding the level of care may be based in part on the risk assessment score as well as what types of treatments are ultimately received, potentially impacting predicted outcomes. This represents a challenge since standard tools for assessing risk of bias in diagnostic studies do not address important issues specific to risk prediction studies. Understanding the performance of risk prediction instruments optimally requires consideration of discrimination (ability to distinguish persons with the disease from those without), calibration (how well predicted risk correlates with actual risk), and risk reclassification rates (the proportion of patients correctly reclassified using the risk prediction instrument into clinically relevant risk group), in addition to standard measures of diagnostic accuracy.[19]

Intra- and inter-rater reliability and ease of use are also important characteristics of risk assessment instruments,[20,21] including GCS,[4] but these can be difficult to measure. First responders may score similar patients differently[22] depending on the environment or nature of the situation, and EMS providers with different levels of training may vary in how they administer the GCS. Variation is likely, given that reliability can be low even within and between highly trained emergency personnel, including physicians and nurses.[4,20,23] Data on ease of use or usability include time to administer the instrument and the amount of missing data. Other measures may be challenging to obtain as it would be difficult to use some standard testing methods (e.g., cognitive interviewing

techniques such as "think aloud") in emergency situations. Alternative methods, such as third party observation or post use recall, could be subject to observation or recall bias.

During the development of field triage guidelines and algorithms by the Centers for Disease Control and Prevention (CDC) National Expert Panel in 2011,[7] use of the mGCS was considered a way to potentially simplify field triage, but was not adopted due in part to lack of evidence about the comparative accuracy and reliability of the mGCS relative to the tGCS. However, more evidence may now be available on the mGCS. The purpose of this project is to conduct a systematic review of the currently available evidence about the comparative predictive utility, reliability, and ease of use of the tGCS and mGCS in field assessment of TBI or unspecified trauma (i.e., with or without TBI), as well as comparative effects on clinical outcomes and early critical resource use, in order to inform the development of evidence-based guidelines for blunt trauma field triage by EMS personnel. This is the first step of a larger Federal effort to systematically examine the evidence base about prehospital triage decisionmaking and transport of patients, and inform future updates to the Field Triage Guidelines.

The aims of this review are to assess the predictive utility, reliability, and ease of use of the tGCS versus the mGCS when administered soon after TBI in the field, as well as comparative effects on clinical decisionmaking and clinical outcomes. This review will provide a synthesis of currently available evidence and gaps in evidence that may be helpful to inform clinical practice and guideline development.

## II. The Key Questions

Key Question 1: In patients with known or suspected trauma, what is the predictive utility of the total Glasgow Coma Scale (tGCS) compared with the motor GCS (mGCS) score for predicting mortality, morbidity, injury severity score ≥16, head AIS score >2 or >3, presence of intracranial hemorrhage, and utilization indicators of severe injury (e.g., receipt of intracranial monitoring within 48 hours of admission, receipt of surgery within 12 hours of admission, or early intubation [in the field or immediately upon presentation to the ED])?

> 1a. How does predictive utility vary according to patient age or other patient characteristics (e.g., traumatic brain injury (TBI) vs. unspecified or other trauma, systolic blood pressure, level of intoxication, type of trauma, or intubation or receipt of medications in the field), the training and background of the person administering the instrument, and the timing/setting of assessment (i.e., in the field vs. upon presentation to the emergency department or urban vs. rural location)?

Key Question 2: In patients with known or suspected trauma, what are the comparative effects of the tGCS compared with the mGCS on over- and under-triage (e.g., proportion of patients mis-classified with regard to measures of injury severity or need for early interventions for severe injury, or early post-admission rates of transfer to a lower or higher level of care due to over- or under-triage)?

2a. How do effects on clinical decisionmaking vary according to patient age or other patient characteristics (e.g., TBI vs. unspecified or other trauma, systolic blood pressure, level of intoxication, type of trauma, or intubation or receipt of medication in the field), the training and background of the person administering the instrument, and the timing/setting of assessment (i.e., in the field vs. upon presentation to the emergency department or urban vs. rural location)?

Key Question 3: In patients with known or suspected trauma, what is the comparative effectiveness of the tGCS compared with the mGCS on clinical outcomes (e.g., mortality, morbidity, quality of life)?

3a. How do effects on clinical outcomes vary according to patient age or other patient characteristics (e.g., TBI vs. unspecified or other trauma, systolic blood pressure, level of intoxication, type of trauma, or intubation or receipt of medication in the field), the training and background of the person administering the instrument, and the timing/setting of assessment (i.e., in the field vs. upon presentation to the emergency department or urban vs. rural location)?

Key Question 4: In patients with known or suspected trauma, what is the comparative reliability (e.g., inter-rater and intra-rater kappa) and ease of use (e.g., time to complete, amount of missing data, user reported satisfaction) of the tGCS compared with the mGCS score?

4a. How do comparative reliability and ease of use vary according to patient age or other patient characteristics (e.g., TBI vs. unspecified or other trauma, systolic blood pressure, level of intoxication, type of trauma, or intubation or receipt of medication in the field), the training and background of the person administering the instrument, and the timing/setting of assessment (i.e., in the field vs. upon presentation to the emergency department or urban vs. rural location)?

Key Question 1 addresses the utility of the tGCS compared with the mGCS for predicting clinical outcomes (mortality, morbidity). In addition, Key Question 1 addresses the predictive utility of the tGCS versus the mGCS on markers of injury severity, as indicated by the injury severity score and utilization markers for severe injury (receipt of early surgery or intracranial pressure monitoring), as a marker of need for tertiary trauma care. Key Question 1 does not directly assess the utility of the tGCS compared with the mGCS for predicting the likelihood that a patient receives tertiary trauma care, since the GCS is used to determine who requires tertiary trauma care. Key Question 2 addresses the impact of the tGCS compared with the mGCS on rates of over- or under-triage, as measured by rates of transfer to a lower or higher level of care, an intermediate outcome. Measuring over- and undertriage is a challenge because factors such as geographic proximity and availability of resources may impact triage decisions, in addition to findings on the tGCS or mGCS. Some overtriage may be acceptable in order to prevent undertriage, which may be more likely to result in adverse clinical outcomes. Therefore, results for Key Question 2 must be interpreted with caution. Key Question 3 addresses the impact of the tGCS compared with the mCGS on clinical outcomes. Key Question 4 addresses the reliability and ease of use of the tGCS compared with the mGCS. For each

key question, a subquestion addresses potential modifiers of treatment effect, including patient age or other patient characteristics, the training and background of the person administering the instrument, and the timing/setting of assessment.

## PICOTS

### Populations

- Persons with known or suspected trauma

### Interventions

- GCS motor score (mGCS)[13,24]
    - Focus on studies of the mGCS using a cutoff score of ≤5 to indicate persons who require high level trauma care, but will include studies that use alternative cutoffs or modifications of mGCS
- GCS total score (tGCS)
    - Focus on studies that use a cutoff tGCS score of ≤13 to indicate persons who require high level trauma care, but will include studies that use alternative cutoffs or modifications of tGCS
- Exclude: Studies that evaluate the utility of mGCS or tGCS in combination with other predictors.
- Potential modifiers: age or other patient characteristics (such as TBI vs. unspecified or other trauma, systolic blood pressure, level of intoxication, type of trauma, or intubation or receipt of medication in the field), the training and background of the person administering the instrument, and the timing/setting of assessment (i.e., in the field vs. upon presentation to the emergency department or urban vs. rural location)

### Comparator

- Studies that compare the mGCS vs. tGCS; for KQs 1, 2, and 4 will also include studies that evaluate either of the instruments alone

### Outcomes (specified for each Key Question)

- KQ1: Predictive utility for mortality, morbidity, injury severity score ≥16,[25] or utilization indicators of severe injury[26] (e.g., receipt of intracranial monitoring within 48 hours of admission, receipt of surgery within 12 hours of admission, or receipt of early intubation [in the field or immediately upon arrival to the ED]), as measured by diagnostic accuracy, adjusted risk estimates, measures of discrimination (e.g., the c-index), measures of calibration (e.g., the Hosmer-Lemeshow test), and risk reclassification rates[27]
- KQ 2: Over- or under-triage
    - Proportion of patients who are transferred to a higher or lower level of care[27]
- KQ 3: Clinical outcomes

    - Mortality (prior to hospital arrival, in the emergency department, or after hospital admission)
    - Morbidity, including cognitive impairment, and medical complications related to the brain injury
    - Quality of life, including functional capacity at discharge or follow-up
- KQ 4:
    - Reliability (e.g., inter-rater and intra-rater kappa)
    - Ease of use (e.g., time to complete, measures of missing data, user reported satisfaction)

**Timing**

- tGCS and mGCS administered soon after injury (in the field) or immediately upon arrival in the emergency department

**Setting**

- Prehospital setting (in the field) or immediately upon arrival at the hospital emergency department
- Exclude: Studies conducted in the developing world

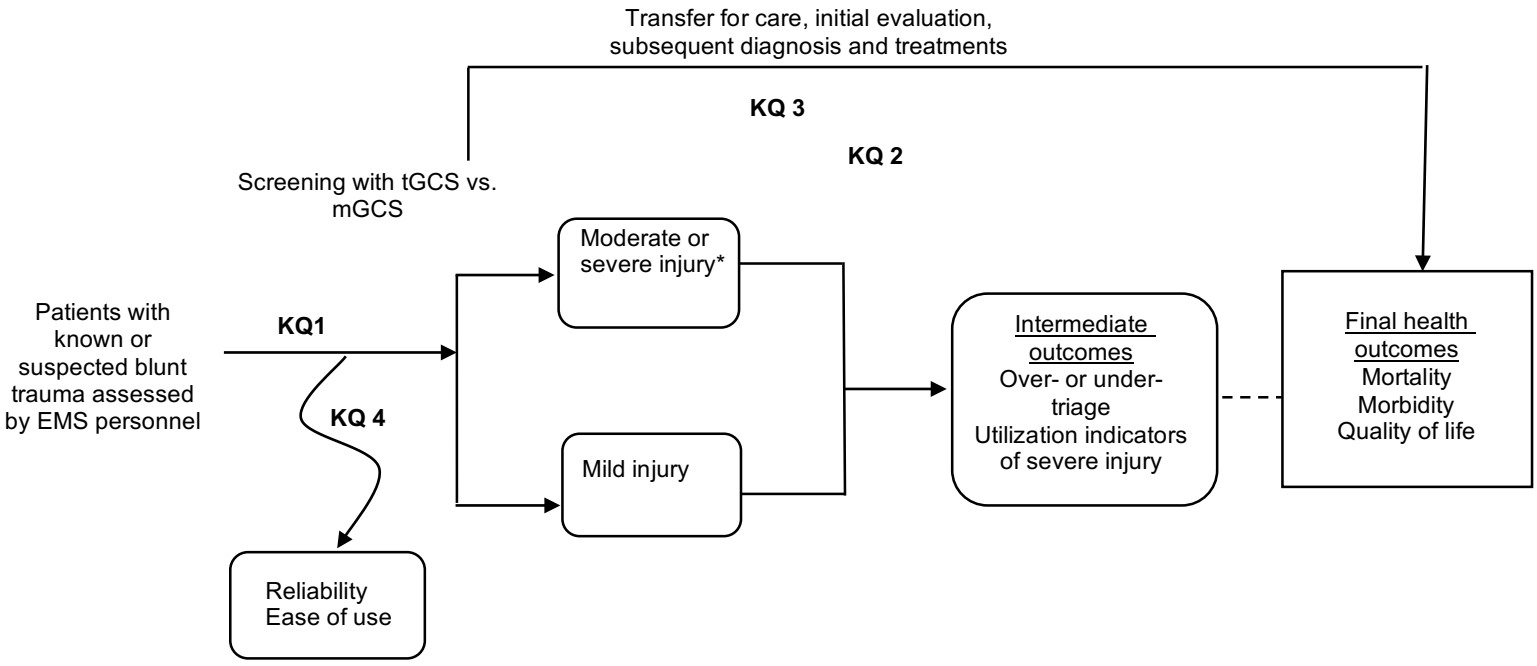
**Study Designs**

- KQ 1, 2, and 4: Randomized trials, cohort, and case-control studies.
- KQ 3: Randomized trials; if few studies then also include cohort and case-control studies

## III. Analytic Framework

The analytic framework illustrates the population, interventions, outcomes, and adverse effects that will guide the literature search and synthesis.

**Figure 1. Analytic framework**




*Based on tGCS of ≤13 or mGCS of ≤5
EMS=emergency medical services; KQ=key question; mGCS=motor score of Glasgow Coma Scale; tGCS=total score of Glasgow Coma Scale

## IV. Methods

### A. Criteria for Inclusion/Exclusion of Studies in the Review

The criteria for inclusion and exclusion of studies will be based on the KQs and the PICOTS defined in the previous section.

Study Design: We will utilize a "best evidence" approach. For KQs 1 (predictive utility), KQ 2 (over- and under-triage) and KQ 4 (reliability and ease of use), studies that directly compare the tGCS versus the mGCS will be prioritized as the top tier evidence. If there is insufficient head-to-head evidence, we will also include studies that evaluate either the tGCS or the mGCS alone, in order to enable indirect comparisons. For KQ 3, randomized trials that directly compare the tGCS versus the mGCS will be prioritized as the top-tier evidence. If there is insufficient RCT

evidence, we will also include comparative cohort and case-control studies.

Our preliminary literature search found few or no systematic reviews addressing the KQs of this review. If systematic reviews are identified, we will consider their inclusion based on relevance and methodological quality, after consultation with the Technical Expert Panel (TEP) and Agency for Healthcare Research and Quality (AHRQ) Task Order Officer (TOO).[28] At a minimum, all systematic reviews will be considered as sources of studies to be reviewed for possible inclusion.

Non-English Language Studies: We will restrict to English-language articles, but will review English language abstracts of non-English language articles to identify studies that would otherwise meet inclusion criteria, in order to assess for the likelihood of language bias.

## B. Searching for the Evidence: Literature Search Strategies and Identification of Relevant Studies to Answer the Key Questions

Publication Date Range: Searches will include articles published since January 1995. This search start date was selected because of changes in trauma care over time; only five states had fully implemented trauma systems in the early 1990's.[29] In addition, the first study comparing the mGCS versus the tGCS was published in 2005.[30]

Literature searches will be updated while the draft report is posted for public comment and out for peer review in order to identify any new publications. Literature identified during the update search will be assessed by following the same process of dual review as all other studies considered for inclusion in the report. If any pertinent new literature is identified for inclusion in the report, it will be incorporated before the final submission of the report.

Literature Databases: MEDLINE, CINAHL, PsycINFO, HAPI (Health & Psychosocial Instruments), and the Cochrane databases will be searched to capture both published and gray literature. The search strategies developed in Ovid MEDLINE are available in Appendix 1.

Scientific Information Packets:

Not applicable for this topic, which addresses non-proprietary assessment scales and does not involve laboratory tests, drugs, or devices.

Hand Searching: Reference lists of included articles will also be reviewed for includable literature.

Contacting Authors: In the event that information regarding methods or results appears to be omitted from the published results of a study, or if we are aware of unpublished data, we will contact the authors to request this information.

Abstract and Article Review Procedures: Initial review of abstracts will be done to determine if a citation is relevant. At this stage the only criteria are that an abstract

presents data about the tGCS and mGCS and that there is an English-language abstract if the article is in a foreign language. To ensure accuracy, all excluded abstracts will be reviewed by a second person. All citations deemed potentially eligible for the review by at least one of the reviewers will be retrieved for full-text review. Each full-text article will be independently reviewed for eligibility by two team members using the inclusion/exclusion criteria outlined above. Any disagreements will be resolved by consensus.

A record of studies excluded at the full-text level with reasons for exclusion will be maintained.

## C. Data Abstraction and Data Management

After studies are selected for inclusion, data will be abstracted into categories that include but are not limited to: a) general information such as study design, year, setting, geographic location, patient characteristics (i.e. TBI vs. unspecified or other trauma, type of injury, total scores, severity of injury, mechanism of injury, intoxication status, systolic blood pressure, type of trauma, or intubation or receipt of medication in the field, and duration since injury; b) characteristics of the tGCS and mGCS and tests used, including timing of administration, cut off scores used, training and experience of the person administering, setting (in the field or upon emergency department presentation or rural vs. urban) and results relevant to each KQ as outlined above.

## D. Assessment of Methodological Risk of Bias of Individual Studies

The quality of included studies will be assessed using predefined criteria. Our methods for assessing risk of bias will be based on the recommendations in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews.[31] For KQ 1, we will apply the QUIPS risk of bias tool for risk prediction instruments.[32] The QUIPS tool includes domains on study participation, study attrition, prognostic factor measurement, outcomes measurement, study funding, and statistical analysis and reporting. For KQ 2-4, we will classify included studies according to the study design (e.g., randomized trial, nonrandomized trial, observational study, etc.), and assess risk of bias using study-design specific criteria adapted from the USPSTF and the Cochrane Back Review group.[33,34] These include criteria related to assessment of selection bias, performance bias, detection bias, attrition bias, and reporting bias.

Two investigators will independently assess risk of bias for each study. Differences will be resolved by discussion and, involvement of a third rater as needed.

The overall risk of bias for each study will be assessed as “low,” “medium,” or “high”.

Studies rated “low” risk of bias have the least risk of bias, and their results are generally considered more valid than studies with the same study design but more flaws. Low risk of bias studies include clear descriptions of the population, setting,

interventions, and comparison groups clear reporting of missing data; apply appropriate means to prevent; and appropriately measure outcomes.

Studies rated "medium" risk of bias are susceptible to some bias, though not enough to necessarily invalidate the results. These studies may not meet all the criteria for "low" risk of bias rating, but do not have flaws likely to cause major bias. The study may also be missing information, making it difficult to assess limitations and potential problems. The "medium" quality category is broad, and studies with this rating will vary in their strengths and weaknesses. The results of some "medium" risk of bias studies are likely to be valid, while others may be only possibly valid.

Studies rated "high" risk of bias have significant flaws that may invalidate the results. They may have a serious or "fatal" flaw or set of flaws in design, analysis, or reporting; large amounts of missing information; or discrepancies in reporting. The results of these studies will be least as likely to reflect flaws in the study design as the true difference between the compared interventions. We will not exclude studies rated high risk of bias a priori, but such studies will be considered to be less reliable than studies at lower risk of bias when synthesizing the evidence, particularly if discrepancies between studies are present.

## E. Data Synthesis

As described earlier, we will apply a "best evidence" approach in which higher quality evidence (based on study design, risk of bias, and use of head-to-head versus indirect comparisons) is prioritized. We will not exclude studies rated high risk of bias *a priori,* but will perform sensitivity analyses to determine how their exclusion would impact conclusions. Within each Key Question, we will qualitatively synthesize overall findings and assess how potential modifiers of effects (e.g. patient characteristics, characteristics of the persons administering the instrument, threshold used for the tGCS or mGCS, timing, or setting) impact results, as well as study design characteristics (type of study, risk of bias). We will perform meta-analysis using random effects models only if evidence is suitable for combining, based on similarities in the populations, interventions, comparisons, and settings evaluated.[35] Meta-analyses will be performed on sensitivity and specificity using a bivariate logistic mixed-effects model and on clinical outcomes using the DerSimonian-Laird model; if statistical heterogeneity is present in analyses of clinical outcomes we will also perform analyses using the Profile Likelihood method. Stratified and sensitivity analyses will be performed on the potential modifiers of effects described above.

For Key Questions 1, 2 and 4, we will include studies that report head to head comparisons of the mGCS and total GCS as well as studies that assess only one of the two instruments. We will evaluate any differences in conclusions based on direct versus indirect comparisons, as assessments of comparative diagnostic accuracy based on direct comparisons can differ from those based on indirect comparisons,[36] and only consider combining direct and indirect evidence if they are suitable for combining, based on assumptions regarding similarity of treatment effects. For Key

Question 3 we will only include head-to-head comparisons, as noncomparative studies on effects of the mGCS or tGCS would be very difficult to interpret.

F. Grading the Strength of Evidence (SOE) for Individual Comparisons and Outcomes

The strength of evidence for each key question will be assessed for each intervention/comparator and outcomes for each Key Question by using the approach described in the AHRQ Methods Guide.[31] To ensure consistency and validity of the evaluation, the grades will be reviewed by the entire team of investigators for:

- Study limitations (low, medium, or high level of study limitations)
- Consistency (consistent, inconsistent, or unknown/not applicable)
- Directness (direct or indirect)
- Precision (precise or imprecise)
- Reporting bias (suspected or undetected) for trials. This will be assessed based on whether primary outcomes are pre-specified and reported and if there is evidence of selective reporting of outcomes; for randomized trials, will also assess for unpublished studies through reviews of clinical trials registries and will assess for differences between protocols and publications in reporting of outcomes.

The strength of evidence will be assigned an overall grade of high, moderate, low, or insufficient according to a four-level scale by evaluating and weighing the combined results of the above domains:

- High-We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable, i.e., another study would not change the conclusions.
- Moderate-We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
- Low-We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
- Insufficient- We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

G. Assessing Applicability

Applicability will be evaluated by considering the characteristics of the population and setting as defined in the PICOTS above.[31] This may include patient characteristics (e.g., demographic characteristics such as age, sex, mechanism of injury, and comorbidities), the sample size of the studies; and timing and location of the field triage (e.g. during transport, at admission to the ER), geographic location (e.g., United States vs. Canada, Europe, or another geographic setting) and the level of training and background of the persons administering the instruments. Studies that are performed in the ER or in which more time has elapsed since the injury will be considered less applicable for the purpose of applying to field triage settings.

Using input from the TEP, we will develop a table or matrix on applicability that shows the components of different scenarios (e.g., patient, injury, environmental characteristics) for which there is direct evidence, situations that differ but for which is it clinically logical to apply the available evidence, and situations with limited or no evidence.

## V. References


1. Baxt WG, Moody P. The impact of advanced prehospital emergency care on the mortality of severely brain-injured patients. J Trauma. 1987;27(4):365-9. PMID: 3573084.
2. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. A national evaluation of the effect of trauma-center care on mortality. N Engl J Med. 2006;354(4):366-78. PMID: 16436768.
3. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet. 1974;2(7872):81-4. PMID: 4136544.
4. Teasdale G, Maas A, Lecky F, et al. The Glasgow Coma Scale at 40 years: standing the test of time. Lancet Neurol. 2014;13(8):844-54. PMID: 25030516.
5. Marshall LF. Evoked potentials: a decade later. Crit Care Med. 1991;19(11):1337. PMID: 1935149.
6. Bradshaw D. Report to the Surgeon General: Traumatic Brain Injury Task Force. May 15, 2007. 2008. Available at: https://www.hsdl.org/?view&did=482727. Accessed December 22, 2015.
7. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011. MMWR Recomm Rep. 2012;61(RR-1):1-20. PMID: 22237112.
8. Singh B, Murad MH, Prokop LJ, et al. Meta-analysis of Glasgow coma scale and simplified motor score in predicting traumatic brain injury outcomes. Brain Inj. 2013;27(3):293-300. PMID: 23252405.
9. Brown KM, Hirshon JM, Alcorta R, et al. 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;18 Suppl 1:45-51. PMID: 24134543.
10. Beskind DL, Stolz U, Gross A, et al. A comparison of the prehospital motor component of the Glasgow coma scale (mGCS) to the prehospital total GCS (tGCS) as a prehospital risk adjustment measure for trauma patients. Prehosp Emerg Care. 2014;18(1):68-75. PMID: 24329032.
11. Healey C, Osler TM, Rogers FB, et al. Improving the Glasgow Coma Scale score: motor score alone is a better predictor. J Trauma. 2003;54(4):671-8; discussion 8-80. PMID: 12707528.
12. Ross SE, Leipold C, Terregino C, et al. Efficacy of the motor component of the Glasgow Coma Scale in trauma triage. J Trauma. 1998;45(1):42-4. PMID: 9680010.
13. National Emergency Medical Services Workforce Definitions. (Prepared by the American Institutes for Research through contract DTNH22-080F-00122.). National Highway Traffic Safety Administration. Washington, DC. Available at: www.nhtsa.gov/staticfiles/nti/pdf/811720.pdf. Accessed December 22, 2015.
14. Trentz O, Lenzlinger PM. Impact of head and chest trauma on general condition (Chapter 5). In: Pape H-C, Peitzman AB, Schwab CW, PV G, eds. Damage Control Management in the Polytrauma Patient. New York: Springer; 2010: p. 53.

15. Reith FC, Brennan PM, Maas AI, et al. Lack of Standardization in the Use of the Glasgow Coma Scale: Results of International Surveys. J Neurotrauma. 2016;33(1):89-94. PMID: 25951090.
16. Yokobori S, Bullock MR. Pathbiology of primary traumatic brain injury (Chapter 11). In: Zasler ND, Katz DI, RD Z, eds. Brain Injury Medicine: Principles and Practice. 2nd ed. New York: Demos Medical Publishing; 2013: p.137.
17. Badjatia N, Carney N, Crocco TJ, et al. Guidelines for prehospital management of traumatic brain injury 2nd edition. Prehosp Emerg Care. 2008;12 Suppl 1:S1-52. PMID: 18203044.
18. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338:b375. PMID: 19237405.
19. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clin Chem. 2008;54(1):17-23. PMID: 18024533.
20. Menegazzi JJ, Davis EA, Sucov AN, et al. Reliability of the Glasgow Coma Scale when used by emergency physicians and paramedics. J Trauma. 1993;34(1):46-8. PMID: 8437195.
21. Gill MR, Reiley DG, Green SM. Interrater reliability of Glasgow Coma Scale scores in the emergency department. Ann Emerg Med. 2004;43(2):215-23. PMID: 14747811.
22. Lane P, Sorondo B, Baez AA. Factors influencing the access of severely injured children and elderly patients involved in motor vehicle collisions to trauma center care. Annu Proc Assoc Adv Automot Med. 2001;45:251-67. PMID: 12214354.
23. Holdgate A, Ching N, Angonese L. Variability in agreement between physicians and nurses when measuring the Glasgow Coma Scale in the emergency department limits its clinical usefulness. Emerg Med Australas. 2006;18(4):379-84. PMID: 16842308.
24. Administration NHTS. National EMS scope of practice model. DOT HS. 2007;810:657.
25. Baker SP, O'Neill B, Haddon W, Jr., et al. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma. 1974;14(3):187-96. PMID: 4814394.
26. Baxt WG, Jones G, Fortlage D. The trauma triage rule: a new, resource-based approach to the prehospital identification of major trauma victims. Ann Emerg Med. 1990;19(12):1401-6. PMID: 2240753.
27. Lerner EB, Willenbring BD, Pirrallo RG, et al. A consensus-based criterion standard for trauma center need. J Trauma Acute Care Surg. 2014;76(4):1157-63. PMID: 24662885.
28. Robinson KA, Chou R, Berkman ND, et al. Integrating Bodies of Evidence: Existing Systematic Reviews and Primary Studies. Methods Guide for Comparative Effectiveness Reviews (Prepared by the Scientific Resource Center under Contract No. 290-2012-00004-C). AHRQ Publication No. 15-EHC007-EF. Rockville, MD:February 2015. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm. Accessed December 22, 2015.

29. Bazzoli GJ, Madura KJ, Cooper GF, et al. Progress in the development of trauma systems in the United States. Results of a national survey. Jama. 1995;273(5):395-401. PMID: 7823385.
30. Gill M, Windemuth R, Steele R, et al. A comparison of the Glasgow Coma Scale score to simplified alternative scores for the prediction of traumatic brain injury outcomes. Annals of Emergency Medicine. 2005;45(1):37-42. PMID: 15635308.
31. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD:January 2014. Available at: www.effectivehealthcare.ahrq.gov. Accessed December 22, 2015.
32. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280-6. PMID: 23420236.
33. Furlan AD, Malmivaara A, Chou R, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. Spine. 2015;40(21):1660-73. PMID: 26208232.
34. U.S. Preventive Services Task Force. U.S. Preventive Services Task Force Procedure Manual. 2015. Available at: http://www.uspreventiveservicestaskforce.org/Home/GetFile/6/301/procedure-manual_2015/pdf. Accessed December 22, 2015.
35. Trikalinos TA, Balion CM, Coleman CI, et al. Chapter 8: meta-analysis of test performance when there is a "gold standard". J Gen Intern Med. 2012;27 Suppl 1:S56-66. PMID: 22648676.
36. Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med. 2013;158(7):544-54. PMID: 23546566.

## VI. Definition of Terms

Emergency Medical Services: Rapid triage out-of-hospital care and transport for individuals who have undergone trauma.

Field Triage: EMS determinate of the destination for the injured subject.

Patient Care Provider: A licensed health worker whose main role is the provision of EMS services to patients.

## VII. Summary of Protocol Amendments

If we need to amend this protocol, we will give the date of each amendment, describe the change and give the rationale in this section. Changes will not be incorporated into the protocol.

## VIII. Review of Key Questions

The proposed KQs were published in the Effective Health Care Program Statement of Work and have not been posted for public comment as this project did not include topic refinement.

## IX. Key Informants

Not applicable.

## X. Technical Experts

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes and identify particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor do they contribute to the writing of the report. They have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

## XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers. The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published three months after the publication of the evidence report.

Potential Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

## XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than $1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than $1,000 will usually disqualify EPC core team investigators.

## XIII. Role of the Funder

This project was funded under Contract No. HHSA 290-2015-0009-i from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

## Appendix 1. MEDLINE Search Strategies

1 exp Glasgow Coma Scale/ (7453)
2 exp Trauma Severity Indices/ (26003)
3 ((glasgow adj3 coma*) or tgcs or mgcs or gcs).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] (16669)
4 2 and 3 (8367)
5 1 or 4 (8367)
6 exp "wounds and injuries"/ (764490)
7 exp accidents/ (153435)
8 exp violence/ (73131)
9 (tbi or ((head or brain* or cereb* or crani* or skull*) adj3 (injur* or traum* or wound* or damag*))).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] (133113)
10 ((case* or patient* or triag* or unconsciou* or consciou* or call* or "911" or emergenc*) adj5 (injur* or traum* or wound* or damag* or hurt*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] (158827)
11 6 or 7 or 8 or 9 or 10 (1040851)
12 exp Emergencies/ (36151)
13 exp Emergency Medical Services/ (104345)
14 (pre-hospital* or prehospital or paramedic* or emt or ems or emergency medical technician* or ambulance* or ((field* or onsite or on-site or scene* or accident*) adj5 triag*)).mp. (39234)
15 exp Emergency Treatment/ (100137)
16 exp emergency medicine/ (10652)
17 (pre-hospital* or prehospital or paramedic* or emt or ems or emergency medical technician* or ambulance* or ((field* or onsite or on-site or scene* or accident*) adj5 triag*)).mp. (39234)
18 exp accidents/ (153435)
19 (emergency or emergencies or triage or priorit*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] (297798)
20 18 and 19 (13245)
21 12 or 13 or 15 or 16 or 17 or 20 (248295)
22 5 and 11 and 21 (1587)
23 limit 22 to english language (1444)
24 limit 22 to abstracts (1483)
25 23 or 24 (1562)
26 limit 25 to yr="1995 -Current" (1427)