Glasgow Coma Scale for Field Triage of Trauma: A Systematic Review
Glasgow Coma Scale for Field Triage of Trauma: A Systematic Review

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. 290-2015-00009-I

Prepared by:
Pacific Northwest Evidence-based Practice Center
Portland, OR

Investigators:
Roger Chou, M.D.
Annette M. Totten, Ph.D.
Miranda Pappas, M.A.
Nancy Carney, Ph.D.
Spencer Dandy, B.S.
Sara Grusing, B.A.
Rongwei Fu, Ph.D.
Ngoc Wasson, M.P.H.
Craig Newgard, M.D., M.P.H., FACEP
Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The National Highway Traffic Safety Administration (NHTSA) requested and provided funding for this report.

The reports and assessments provide organizations with comprehensive, evidence-based information on common medical conditions and new health care technologies and strategies. They also identify research gaps in the selected scientific area, identify methodological and scientific weaknesses, suggest research needs, and move the field forward through an unbiased, evidence-based assessment of the available literature. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review and public comment prior to their release as a final report.

AHRQ expects that the EPC evidence reports and technology assessments, when appropriate, will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

If you have comments on this evidence report, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Andrew Bindman, M.D. Arlene S. Bierman, M.D., M.S.
Director Director
Agency for Healthcare Research and Quality Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H. David W. Niebuhr, M.D., M.P.H., M.Sc.
Director Task Order Officer
Evidence-based Practice Center Program Center for Evidence and Practice Improvement
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality Agency for Healthcare Research and Quality

Cathy Gotschall, Sc.D.
Senior Health Scientist
National Highway Traffic Safety Administration
U.S. Department of Transportation
Acknowledgments

The authors acknowledge Andrew Hamilton, M.L.S., M.S., for conducting literature searches, and Leah Williams, B.S., for providing editorial assistance. Both are located at the Oregon Health & Science University.

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

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 with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who provided input to this report follows:

Eileen Metzger Bulger, M.D., FACS*
Professor
Department of Surgery
University of Washington
Chief of Trauma, Harborview Medical Care
Seattle, WA

E. Brooke Lerner, Ph.D.*
Professor With Tenure
Departments of Emergency Medicine and Pediatrics
Medical College of Wisconsin
Milwaukee, WI

Steven M. Green, M.D., FACEP*
Professor
Emergency Medicine and Pediatrics
Loma Linda University Medical Center and Children’s Hospital
Loma Linda, CA

N. Clay Mann, Ph.D., M.S.*
Professor
Department of Pediatrics
University of Utah School of Medicine
Salt Lake City, UT

Eileen Metzger Bulger, M.D., FACS*

P. Daniel Patterson, Ph.D., M.P.H., M.S., N.R.P.*
Adjunct Assistant Professor
Department of Emergency Medicine
University of Pittsburgh
Pittsburgh, PA

Julie C. Leonard, M.D., M.P.H.*
Associate Professor
Department of Pediatrics
Division of Emergency Medicine
The Ohio State University College of Medicine
Columbus, OH
Peer Reviewers

Prior to publication of the final evidence report, the EPC sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers.

Peer Reviewers 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 with potential nonfinancial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

The list of Peer Reviewers follows:

Blair Bigham, M.D., M.Sc., ACP
Paramedic and EMS Researcher
Physician and Scientist
McMaster University
Hamilton, ON, Canada

Mark Gestring, M.D., FACS
Medical Director, Kessler Trauma Center
Associate Professor of Surgery,
Emergency Medicine and Pediatrics
University of Rochester School of Medicine
Rochester, NY

David Cone, M.D.
Professor of Emergency Medicine
Chief, Section of EMS
Director, EMS Fellowship
Yale University School of Medicine
New Haven, CT

Ritu Sahni, M.D.
Physician, Emergency Medicine
Providence Hospital
Portland, OR
Glasgow Coma Scale for Field Triage of Trauma: A Systematic Review

Structured Abstract

Objectives. To assess the predictive utility, reliability, and ease of use of the total Glasgow Coma Scale (tGCS) versus the motor component of the Glasgow Coma Scale (mGCS) for field triage of trauma, and effects on clinical decisionmaking and clinical outcomes.

Data sources. MEDLINE®, CINAHL®, PsycINFO®, HaPI (Health and Psychosocial Instruments), the Cochrane Databases (January 1995 through June 2016), and reference lists.

Study selection. Studies on predictive utility of the tGCS versus the mGCS or Simplified Motor Scale (SMS), randomized trials and cohort studies on effects of the tGCS versus the mGCS on rates of over- or under-triage, and studies on interrater reliability and ease of use.

Data extraction. One investigator abstracted study characteristics and results; a second checked data for accuracy. Two investigators independently applied prespecified criteria to rate study quality. Data on discrimination were pooled using a random effects model. The strength of evidence was graded using published methods.

Results. Thirty-two studies met inclusion criteria; 24 studies addressed predictive utility and 10 addressed interrater reliability or ease of use. No study assessed comparative effects on over- or under-triage or clinical outcomes. For in-hospital mortality, the tGCS is associated with slightly greater discrimination than the mGCS (pooled mean difference in area under the receiver operating characteristic curve [AUROC], 0.015; 95% confidence interval [CI], 0.009 to 0.022; I² = 85%; 12 studies; strength of evidence [SOE]: Moderate) or the SMS (pooled mean difference in AUROC, 0.030; 95% CI, 0.024 to 0.036; I² = 0%; 5 studies; SOE: Moderate). This means that for every 100 trauma patients, the tGCS correctly discriminates 1 to 3 more cases of in-hospital mortality from cases without in-hospital mortality than the mGCS or SMS. The tGCS is also associated with greater discrimination than the mGCS or SMS for receipt of neurosurgical interventions, severe brain injury, and emergency intubation (differences in AUROC from 0.03 to 0.05; SOE: Moderate). Differences in discrimination between mGCS and SMS were very small.

Findings were robust in sensitivity and subgroup analyses. Differences among the tGCS, mGCS, and SMS in diagnostic accuracy (sensitivity, specificity) using standard thresholds were small, based on limited evidence (SOE: Low). Evidence was insufficient to determine if there were differences between the tGCS and the mGCS in interrater reliability (SOE: Insufficient). Three studies found that the tGCS was associated with lower proportions of correct scores than the mGCS; differences ranged from 6 percent to 27 percent (SOE: Low).

Limitations. Evidence on comparative predictive utility was primarily restricted to effects on discrimination. All predictive utility studies were retrospective, and mGCS and SMS were taken from tGCS rather than independently assessed. Most studies had methodological limitations. We included only English-language studies and were limited in our ability to assess publication bias. Studies on ease of use focused on scoring of video or written patient scenarios.
Conclusions. The tGCS is associated with slightly greater discrimination than the mGCS or SMS for in-hospital mortality, receipt of neurosurgical interventions, severe brain injury, and emergency intubation. The clinical significance of small differences in discrimination is likely to be small and could be offset by factors such as convenience and ease of use. Research is needed to understand how use of the tGCS versus the mGCS or SMS impacts clinical outcomes and risk of over- or under-triage.
Contents

Executive Summary ................................................................. ES-1
Introduction ................................................................................ 1
  Background ............................................................................ 1
    Nature and Burden of Trauma ............................................. 1
    Field Triage of Patients With Trauma ............................... 1
  Rationale for Review ........................................................... 2
    Scope of Review and Key Questions .................................. 3
Methods .................................................................................. 6
  Scope Development .............................................................. 6
  Literature Search Strategy .................................................. 6
  Study Selection ...................................................................... 6
    Population and Conditions of Interest .............................. 7
    Interventions, Comparisons, and Study Designs ............... 7
  Outcomes ............................................................................. 7
    Timing and Setting ............................................................ 8
  Data Abstraction and Data Management .............................. 8
  Assessment of Methodological Risk of Bias of Individual Studies .................................................. 8
  Assessing Research Applicability ......................................... 9
  Data Synthesis and Rating the Body of Evidence ................... 9
  Grading the Body of Evidence for Each Key Question .......... 10
  External Review ................................................................. 11
Results ..................................................................................... 12
  Results of Literature Search ................................................. 12
  Key Question 1. In patients with known or suspected trauma, what is the predictive utility of the total Glasgow Coma Scale compared with the motor component of the Glasgow Coma Scale for predicting in-hospital mortality, morbidity, injury severity score of 16 or greater, head Abbreviated Injury Scale (AIS) score greater than 2 (AIS-2005) or greater than 3 (AIS-1998), presence of intracranial hemorrhage, and utilization indicators of severe injury (e.g., receipt of intracranial monitoring within 48 hours of admission, receipt of a neurosurgical intervention within 12 hours of admission, or early intubation [in the field or immediately upon presentation to the emergency department])? ................................................................. 14
    Key Points .......................................................................... 14
      Detailed Synthesis ............................................................ 16
  Key Question 1a. How does predictive utility vary according to patient age or other patient characteristics (e.g., traumatic brain injury vs. unspecified or other trauma, systolic blood pressure, level of intoxication, type of trauma, 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)? ............................................................................. 47
    Key Points .......................................................................... 47
      Detailed Synthesis ............................................................ 47
  Key Question 2. In patients with known or suspected trauma, what are the comparative effects of the total Glasgow Coma Scale compared with the motor Glasgow Coma Scale on over- and under-triage (e.g., proportion of patients in whom initial emergency medical
services transport is to a higher or lower than appropriate level of care, or proportion
transferred to a lower or higher level of care)? .................................................................50
  Key Points ..........................................................................................................................50
  Detailed Synthesis ..............................................................................................................50
Key Question 2a. How do effects on over- and under-triage vary according to patient age or
other patient characteristics (e.g., traumatic brain injury vs. unspecified or other trauma,
systolic blood pressure, level of intoxication, type of trauma, 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)? ..............................................................51
  Key Points ..........................................................................................................................51
Key Question 3. In patients with known or suspected trauma, what is the comparative
effectiveness of the total Glasgow Coma Scale compared with the motor component of the
Glasgow Coma Scale on clinical outcomes (e.g., in-hospital mortality, morbidity, quality of
life)? ........................................................................................................................................51
  Key Points ..........................................................................................................................51
Key Question 3a. How do effects on clinical outcomes vary according to patient age or other
patient characteristics (e.g., traumatic brain injury vs. unspecified or other trauma, systolic
blood pressure, level of intoxication, type of trauma, 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)? .................................................................................51
  Key Points ..........................................................................................................................51
Key Question 4. In patients with known or suspected trauma, what is the comparative
reliability (e.g., interrater and intrarater kappa) and ease of use (e.g., time to complete,
amount of missing data, user reported satisfaction) of the total Glasgow Coma Scale
compared with the motor component of the Glasgow Coma Scale? ......................................51
  Key Points ..........................................................................................................................51
  Detailed Synthesis .............................................................................................................52
Key Question 4a. How do comparative reliability and ease of use vary according to patient
age or other patient characteristics (e.g., traumatic brain injury vs. unspecified or other
trauma, systolic blood pressure, level of intoxication, type of trauma, 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)? ..............................................................56
  Key Points ..........................................................................................................................56
  Detailed Synthesis .............................................................................................................57

Discussion ..................................................................................................................................59
Key Findings and Strength of Evidence ..................................................................................59
Findings in Relationship to What Is Already Known .................................................................61
Applicability ..............................................................................................................................62
Implications for Clinical and Policy Decisionmaking ...............................................................63
Limitations of the Review Process ..........................................................................................63
Gaps in the Evidence Base .........................................................................................................65
Future Research Needs .............................................................................................................65
Conclusions ................................................................................................................................66
Tables
Table A. Pooled AUROC results of head-to-head studies .................................................................ES-7
Table 1. Characteristics of head-to-head studies ...........................................................................17
Table 2. Proportion of patients experiencing outcomes in head-to-head studies on predictive utility ..............................................................................................................................................21
No table of figures entries found.
Table 3. Summary of discrimination (AUROC) of head-to-head studies ........................................24
Table 4. Pooled AUROC results of head-to-head studies .............................................................27
Table 5. Summary of diagnostic accuracy outcomes for head-to-head studies ............................29
Table 6. Reliability and ease of use findings for studies comparing the total Glasgow Coma Scale versus the motor component only ...........................................................................................................53

Figures
Figure 1. Analytic framework..........................................................................................................5
Figure 2. Literature flow diagram ..................................................................................................13
Figure 3. Pooled AUROC of in-hospital mortality for the total Glasgow Coma Scale versus the motor component only ................................................................................................................23
Figure 4. Pooled AUROC of in-hospital mortality for the total Glasgow Coma Scale versus the Simplified Motor Scale ................................................................................................................33
Figure 5. Pooled AUROC of neurological intervention for the total Glasgow Coma Scale versus the motor component only .............................................................................................................35
Figure 6. Pooled AUROC of neurological intervention for the total Glasgow Coma Scale versus the Simplified Motor Scale .............................................................................................................37
Figure 7. Pooled AUROC of severe brain injury for the total Glasgow Coma Scale versus the motor component only ..................................................................................................................39
Figure 8. Pooled AUROC of severe brain injury for the total Glasgow Coma Scale versus the Simplified Motor Scale ..................................................................................................................41
Figure 9. Pooled AUROC of intubation for the total Glasgow Coma Scale versus the motor component only ..............................................................................................................................43
Figure 10. Pooled AUROC of intubation for the total Glasgow Coma Scale versus the Simplified Motor Scale ....................................................................................................................................45

Appendixes
Appendix A. Search Strategies
Appendix B. Inclusion and Exclusion Criteria
Appendix C. List of Included Studies
Appendix D. List of Excluded Studies
Appendix E. Risk of Bias Criteria
Appendix F. Strength of Evidence Domains and Definitions
Appendix G. Strength of Evidence Table
Appendix H. Head-to-Head Studies for Predictive Utility
Appendix I. Indirect Studies for Predictive Utility
Appendix J. Studies of Reliability and Ease of Use
Appendix K. Quality Assessment of Studies of Predictive Utility
Appendix L. Quality Assessments of Studies of Reliability and Ease of Use
Executive Summary

Background

Unintentional injuries are the leading cause of death among people in the United States ages 1 to 44, and the third leading cause among people ages 45 to 64. Among all age groups, motor vehicle crashes are the first or second leading cause of unintentional injury death. In 2011, there were approximately 40,000,000 emergency department (ED) visits for injuries; of these approximately 2.5 million were due to trauma complications and unspecified injuries. Approximately 18 percent of patients seen in the ED for an injury were transported by emergency medical services (EMS) personnel. Traumatic brain injury (TBI) is an important subset of trauma. Among an estimated 1.7 million annual cases of TBI, there are 52,000 deaths and 275,000 hospitalizations. TBI is a contributing factor to about one-third of injury-related deaths in the United States. From 2001 to 2010, the rate of TBI-related ED visits increased from 421 to 716 per 100,000, though the rate of deaths declined from 18.5 to 17.1 per 100,000 people.

Field Triage of Patients With Trauma

Field triage by EMS is a critical aspect of trauma systems, as it helps to identify potentially seriously injured patients and inform transport decisions. Appropriate decisions regarding transport are crucial because management of severely injured patients in a Level I or a Level II trauma center has been shown to be associated with improved clinical outcomes. On the other hand, unnecessarily triaging patients to high-level trauma care who do not require it may represent an inefficient use of staff and resources.

EMS personnel must rapidly triage individuals who have undergone trauma in challenging environments. Therefore, EMS personnel must have assessment tools that are easy to use, reliable, and accurate. A key component of field triage for patients with suspected serious injury is level of consciousness assessment. The Glasgow Coma Scale (GCS) is an instrument widely used for assessment of consciousness at the site of injury, in EDs, and in hospitals, and to monitor progress or deterioration during treatment. The GCS consists of three items (components): eye (scored 1 to 4), verbal (scored 1 to 5), and motor (scored 1 to 6). Scores on each of these components are added to obtain the total Glasgow Coma Scale (tGCS) score, ranging from 3 to 15. Lower scores on the tGCS indicate lower levels of consciousness, generally correlating with more severe injury associated with poorer prognosis and requiring more intensive care. For patients with TBI, scores of 3 to 8 are generally considered to denote severe head injury, 9 to 12 moderate, and 13 to 15 mild. The 2011 field triage guidelines from the Centers for Disease Control and Prevention (CDC) National Expert Panel recommend transporting patients with tGCS scores of 13 or less to facilities providing the highest level of trauma care.

In some circumstances (e.g., trauma victims who are intoxicated, intubated, or whose other injuries influence response) it may not be possible to accurately assess the verbal and eye components of the GCS. In these cases, assessments may be primarily based on the motor component of the Glasgow Coma Scale (mGCS) alone. The mGCS has been proposed for assessment of trauma patients even when the tGCS can be obtained, since only one item is assessed, potentially increasing ease of use in the field. mGCS scores of 5 or less are considered an indication of patients with severe injury. The Simplified Motor Score (SMS)
has been proposed as a streamlined alternative to the mGCS, which is assessed on a three-point scale (scored 0 to 2, with a score of 0 corresponding to 1 to 4 on the mGCS, 1 corresponding to 5 on the mGCS, and 2 corresponding to 6 on the mGCS). Decisions regarding the use of the tGCS versus more simplified versions of the tGCS should be based on their relative performance. The ultimate goal of selecting one risk prediction instrument over another is to improve clinical outcomes (e.g., in-hospital mortality). However, information on clinical outcomes is often lacking, and decisions about use must often be based on how instruments perform on intermediate outcomes. Intermediate outcomes include measures of over- or under-triage (i.e., the degree to which patients are unnecessarily transported to a Level I or II trauma center [over-triage] or not transported to a Level I or II trauma center when needed[under-triage]) or predictive utility, as assessed using measures of discrimination (ability of an instrument to distinguish patients with the disease from those without), calibration (how well predicted risk correlates with actual risk), standard measures of diagnostic accuracy (e.g., sensitivity, specificity, and predictive values), or adjusted risk estimates (e.g., odds ratio, relative risk, hazards ratio). Other factors that could inform selection of field triage risk assessment instruments include intra- and interrater reliability and ease of use (e.g., time to administer the instrument and amount of missing data).

A number of factors could impact the performance of field assessment instruments. These include variability in patient populations (e.g., type of trauma, demographic characteristics, presence and severity of intoxication, and medical comorbidities), level of training and certification of administering personnel (e.g., emergency medical responder, emergency medical technician [EMT], advanced EMT, paramedic, physician, or nurse), receipt of field interventions (e.g., medications, intubation), setting (e.g., country, urban vs. rural) or timing of assessment relative to injury occurrence. Evidence about field triage instruments frequently relies on extrapolation from studies conducted in EDs, as this environment is more controlled and easier to study. However, the performance of the tGCS and mGCS may be different when administered soon after injury by EMS personnel in the field as opposed to later by ED personnel, after destination decisions have already been made and patients have been stabilized with initial interventions.

During the development of field triage guidelines and algorithms by the CDC National Expert Panel in 2011, use of the mGCS was considered a way to potentially simplify field triage. The mGCS 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 is now available on the mGCS.

Scope of Review and Key Questions

This report addresses the following Key Questions:

Key Question 1. In patients with known or suspected trauma, what is the predictive utility of the tGCS compared with the mGCS for predicting in-hospital mortality, morbidity, Injury Severity Score of 16 or greater, head Abbreviated Injury Scale (AIS) score greater than 2 (AIS-2005) or greater than 3 (AIS-1998), presence of intracranial hemorrhage, and utilization indicators of severe injury (e.g., receipt of intracranial monitoring within 48 hours of admission, receipt of a neurosurgical intervention within 12 hours after injury)?
of admission, or early intubation [in the field or immediately upon presentation to the ED])?

Key Question 1a. How does predictive utility 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, 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 ED 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 in whom initial EMS transport is to a higher or lower than appropriate level of care, or proportion transferred to a lower or higher level of care)?

Key Question 2a. How do effects on over- and under-triage 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, 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 ED 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., in-hospital mortality, morbidity, quality of life)?

Key Question 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, 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 ED or urban vs. rural location)?

Key Question 4. In patients with known or suspected trauma, what is the comparative reliability (e.g., interrater and intrarater 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?

Key Question 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, 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 ED or urban vs. rural location)?

The research questions used to guide this review were initially developed by the National Highway Traffic Safety Administration and revised with input from a Technical Expert Panel. The Key Questions focus on predictive utility, over- and under-triage, clinical outcomes of the tGCS versus the mGCS or the SMS, as well as reliability and ease of use. We included studies of children and adults with known or suspected trauma, with assessment using the tGCS, the mGCS, or the SMS. For studies evaluating measures of diagnostic accuracy (sensitivity, specificity, predictive values), we focused on studies that used standard cutoff scores (≤13 for tGCS and ≤5 for mGCS), but also included studies that used alternative cutoffs or modifications of the tGCS and mGCS. For all Key Questions, we included cohort studies and randomized trials that directly compared the tGCS with the mGCS or SMS. For Key Question 4 (reliability and ease of use), we also included cross-sectional studies and studies that assessed one of these scales, and for Key Question 1a (predictive utility) we included studies that assessed one of these scales if they addressed one of the subpopulations specified in the Key Questions not addressed well in the head-to-head studies.

For Key Question 1, we included measures of predictive utility for in-hospital mortality, morbidity, markers of severe injury, or utilization indicators of severe injury, 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), or risk reclassification rates. For Key Question 2, we included studies that reported the proportion of patients who were over- or under-triaged (e.g., the proportion transferred to a higher or lower level of care); for Key Question 3, we included studies that reported clinical outcomes; and for Key Question 4, we included outcomes that assessed reliability (e.g., interrater and intrarater kappa) or ease of use (e.g., time to complete, measures of missing data, user reported satisfaction).

For all Key Questions we included prospective and retrospective studies in which the tGCS, mGCS, or SMS was administered soon after injury (conducted in the field/out-of-hospital setting by EMS personnel) or immediately upon arrival to the ED, or that were based on trauma registry data collected in the field or in the ED.

**Methods**

**Literature Search Strategy**

This review includes studies 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 1990s. In addition, the first studies to compare the predictive utility of the mGCS versus the tGCS were published in 1998 and 2003.

The Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, CINAHL®, PsycINFO®, HaPI (Health and Psychosocial Instruments), and Ovid MEDLINE® (January 1995 through June 2016) were searched for relevant studies and systematic reviews. Investigators also manually reviewed reference lists of relevant studies and searched for
unpublished studies in ClinicalTrials.gov. All citations were independently reviewed by two investigators to determine eligibility for inclusion.

Risk of Bias Assessment of Individual Studies

A single investigator abstracted details about study design, patient population, comparison groups, setting, screening method, analysis, followup, and results. A second investigator reviewed data abstraction for accuracy. Two investigators independently rated the quality of studies (good, fair, poor) using prespecified criteria developed for evaluation of studies on prognosis27 and diagnosis.28 Investigators did not review, assess, or screen papers that they authored. Discrepancies were resolved by consensus.

Data Synthesis

We applied a “best evidence” approach in which higher quality evidence (based on study design, risk of bias, and use of head-to-head vs. indirect comparisons) is prioritized. We did not exclude studies rated high risk of bias a priori, but performed sensitivity analyses to determine how their exclusion would impact conclusions. Within each Key Question, we qualitatively synthesized overall findings and assessed how potential modifiers of effects (e.g., patient characteristics, characteristics of the people administering the instrument, threshold used for the tGCS or mGCS, timing, or setting) impacted results, as well as study design characteristics (type of study, risk of bias). We performed meta-analysis to calculate pooled differences in the area under the receiver operating characteristic curve (AUROC) using the DerSimonian-Laird random effects model with Stata/IC 13.1 (StataCorp LP, College Station, TX). We also performed analyses using the Profile Likelihood method. The AUROC value represents the probability that a patient who experiences an outcome will have a higher score on the triage scale than a person who does not experience the outcome. We defined a small difference in the AUROC a priori as a difference of less than 0.05, moderate as a difference of 0.05 to 0.10, and large as a difference of greater than 0.10. Stratified and sensitivity analyses were performed on the potential modifiers of effects.

We evaluated 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, and did not combine direct and indirect evidence.

Strength of the Body of Evidence

For all comparisons and outcomes we assessed the strength of evidence using the approach described in the Methods Guide for Effectiveness and Comparative Effectiveness Reviews,29 based on the overall risk of bias (graded low, moderate, or high); the consistency of results across studies (graded consistent, inconsistent, or unable to determine when only one study was available); the directness of the evidence linking the intervention and health outcomes (graded direct or indirect); the precision of the estimate of effect, based on the number and size of studies and confidence intervals (CIs) for the estimates (graded precise or imprecise); and reporting bias (suspected or undetected). Assessments of reporting bias were based on whether studies defined and reported primary outcomes and whether we identified relevant unpublished studies.
Results

Results of Literature Searches

Database searches resulted in 4,412 potentially relevant citations. After dual review of abstracts and titles, 698 articles were selected for full-text review. After dual review of full-text articles, 32 studies were included.

Key Question 1. Predictive Utility

Twenty-four studies evaluated predictive utility.\textsuperscript{17-20,30-49} Differences between the tGCS, mGCS, and SMS in discrimination (AUROC) for in-hospital mortality, neurosurgical intervention, severe brain injury, and emergency intubation were <0.05. Results were similar in subgroups stratified by age (child vs. mixed populations of adults and children), type of trauma (TBI vs. mixed trauma), field versus ED assessment, and other subgroup and sensitivity analyses. Main findings are summarized in Table A and below.

- In-hospital mortality
  - For the tGCS versus the mGCS, the pooled AUROC was 0.877 (95% CI 0.847 to 0.906) versus 0.855 (95% CI 0.822 to 0.888), with a pooled mean difference of 0.015 (95% CI 0.009 to 0.022; $I^2=85\%$), based on 12 studies (strength of evidence [SOE]: Moderate).
  - For the tGCS (cutoff of $\leq13$) versus the mGCS (cutoff of $\leq5$), differences in sensitivity ranged from 0 percent to 3 percent; difference in specificity ranged from 0 percent to 5 percent in favor of the mGCS, though the CIs overlapped in most studies. (SOE: Low).
  - For the tGCS versus the SMS, the pooled AUROC was 0.884 (95% CI 0.852 to 0.916) versus 0.840 (95% CI 0.802 to 0.878), for a mean difference of 0.030 (95% CI 0.024 to 0.036, $I^2=0\%$), based on five studies (SOE: Moderate).
  - One study found the out-of-hospital tGCS (cutoff of $\leq13$) associated with slightly higher sensitivity versus the SMS (cutoff of $\leq1$) (75%, 95% CI 73 to 76 vs. 72%, 95% CI 70 to 74) and slightly lower specificity (88%, 95% CI 87 to 88 vs. 89%, 95% CI 89 to 87) (SOE: Low).
  - For the mGCS versus the SMS, the mean difference in the pooled AUROC was 0.014 (95% CI 0.006 to 0.021, $I^2=0\%$), based on four studies (SOE: Moderate).

- Neurosurgical intervention
  - For the tGCS versus the mGCS, the pooled AUROC was 0.786 (95% CI 0.729 to 0.842) versus 0.754 (95% CI 0.688 to 0.819), with a mean difference of 0.031 (95% CI 0.018 to 0.044; $I^2=72\%$), based on seven studies (SOE: Moderate).
  - Two studies found inconsistent results for sensitivity (51%, 95% CI 50 to 52 vs. 46%, 95% CI 45 to 48 and 63%, 95% CI 38 to 84 vs. 68%, 95% CI 43 to 87) and no clear differences in specificity (86%, 95% CI 86 to 86 vs. 88%, 95% CI 88 to 88 and 82%, 95% CI 80 to 84 vs. 83%, 95% CI 81 to 85) between out-of-hospital tGCS (cutoff of $\leq13$) versus the mGCS (cutoff of $\leq5$) in accuracy for identifying people undergoing craniotomy (SOE: Insufficient for sensitivity, low for specificity).
  - For the tGCS versus the SMS, the pooled AUROC was 0.809 (95% CI 0.766 to 0.853) versus 0.769 (95% CI 0.711 to 0.827), with a mean difference of 0.032 (95% CI 0.025 to 0.039, $I^2=0\%$), based on five studies (SOE: Moderate).
One study found the out-of-hospital tGCS (cutoff of ≤13) associated with higher sensitivity than the SMS (cutoff of ≤1) for identifying patients who underwent neurosurgical intervention (60%, 95% CI 56 to 63 vs. 53%, 95% CI 49 to 56) and slightly lower specificity (85%, 95% CI 84 to 85 vs. 86%, 95% CI 86 to 87) (SOE: Low).

For the mGCS versus the SMS, the mean difference in the pooled AUROC was 0.002 (95% CI -0.005 to 0.010, I²=0%), based on four studies (SOE: Moderate).

**Severe brain injury**

For the tGCS versus the mGCS, the pooled AUROC was 0.791 (95% CI 0.734 to 0.827) versus 0.720 (95% CI 0.666 to 0.774), with a mean difference of 0.050 (95% CI 0.034 to 0.065; I²=57%), based on five studies (SOE: Moderate).

One study found no difference between out-of-hospital tGCS (cutoff of ≤13) versus the mGCS (cutoff of ≤5) in sensitivity (62%, 95% CI 55 to 68 vs. 61%, 95% CI 54 to 67) or specificity (85%, 95% CI 83 to 88 vs. 89%, 95% CI 88 to 91) for identifying people with severe head injury (defined as head Abbreviated Injury Scale score of ≥4) (SOE: Low).

For the tGCS versus the SMS, the mean difference 0.000, 95% CI -0.008 to 0.007, I²=0%), based on four studies (SOE: Moderate).

**Emergency intubation**

For the tGCS versus the mGCS, the pooled AUROC was 0.865 (95% CI 0.830 to 0.901) versus 0.822 (95% CI 0.775 to 0.870), with a mean difference of 0.034 (95% CI 0.020 to 0.482; I²=88%), based on six studies (SOE: Moderate).

One study found very small differences between out-of-hospital tGCS (cutoff of ≤13) versus the mGCS (cutoff of ≤5) in accuracy for identifying people who underwent emergency intubation (sensitivity 84%, 95% CI 83 to 84 vs. 81%, 95% CI 81 to 82; and specificity 90%, 95% CI 90 to 90 vs. 92.0%, 95% CI 92 to 92) (SOE: Low).

For the tGCS versus the SMS, the pooled AUROC was 0.843 (95% CI 0.823 to 0.864) versus 0.783 (95% CI 0.747 to 0.819), with a mean difference of 0.040 (95% CI 0.030 to 0.050, I²=55%), based on five studies (SOE: Moderate).

One study found the out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity than the SMS (cutoff of ≤1) for identifying people who underwent emergency intubation (76%, 95% CI 74 to 77 vs. 73%, 95% CI 71 to 74) and slightly lower specificity (89%, 95% CI 89 to 89 vs. 91%, 95% CI 90 to 91) (SOE: Low).

For the mGCS versus the SMS, there was no difference in the AUROC (mean difference 0.000, 95% CI -0.007 to 0.007, I²=0%), based on four studies (SOE: Moderate).

**Trauma center need**

ES-7
Two studies found small differences between the tGCS versus the mGCS in the AUROC (0.617 vs. 0.609 and 0.641 vs. 0.603), sensitivity (30% vs. 27% and 28% vs. 25%), and specificity (93% vs. 95% and 94% vs. 95%) for trauma center care need (defined as Injury Severity Score [ISS] of >15, intensive care unit [ICU] admission >24 hours, need for urgent surgery, or death in the ED) (SOE: Low).

- **Severe injury**
  - Two studies found the tGCS was better able to discriminate those with major injury (defined as an ISS score of >15) from those without major injury (AUROC 0.72, 95% CI 0.71 to 0.72 vs. 0.68, 95% CI 0.68 to 0.69 and 0.65, 95% CI 0.65 to 0.65 vs. 0.61, 95% CI 0.60 to 0.61) (SOE: Low).
  - One study found the out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity and slightly lower specificity than the mGCS (cutoff of ≤5) for identifying people with major injury (defined as an ISS score of >15) from those without major injury (sensitivity 31%, 95% CI 31 to 32 vs. 28%, 95% CI 28 to 28 and specificity 91%, 95% CI 91 to 91 vs. 93%, 95% CI 93 to 93) (SOE: Low).

- **Age:** Effects on discrimination between the tGCS versus the mGCS were similar in studies that enrolled children and those that enrolled mixed populations of adults and children (SOE: Low).

- **Type of trauma:** Effects on discrimination between the tGCS versus the mGCS were similar in studies that evaluated patients with TBI and those that enrolled mixed trauma patients (SOE: Low).

- **Out-of-hospital versus ED assessment:** One study of adults found no differences between out-of-hospital and ED GCS scores on discrimination for in-hospital mortality or neurosurgical intervention but another study of adults or children found out-of-hospital GCS scores associated with higher discrimination for in-hospital mortality than ED scores (AUROC 0.754 vs. 0.635, p-value not reported). Effects on discrimination between the tGCS and the mGCS were similar in studies that evaluated out-of-hospital GCS scores and those that used ED scores (SOE: Insufficient).

- No study evaluated how intoxication status, blood pressure, intubation status, receipt of field intubation, or level/training of field assessors impacts comparative predictive utility of the tGCS versus the mGCS or SMS.

### Table A. Pooled AUROC results of head-to-head studies

<table>
<thead>
<tr>
<th>Outcome and analysis</th>
<th>tGCS vs. mGCS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>I² (%)</th>
<th>tGCS vs. SMS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>I² (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-hospital mortality, overall</td>
<td>0.015 (0.009 to 0.022)</td>
<td>12</td>
<td>85%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Adults or mixed</td>
<td>0.019 (0.012 to 0.025)</td>
<td>10</td>
<td>75%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Children</td>
<td>0.006 (0.002 to 0.011)</td>
<td>2</td>
<td>0%</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Excluding NTDB studies</td>
<td>0.017 (0.008 to 0.025)</td>
<td>10</td>
<td>68%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Excluding studies with potential overlap*</td>
<td>0.016 (0.008 to 0.024)</td>
<td>9</td>
<td>88%</td>
<td>0.031 (0.023 to 0.039)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>Out-of-hospital GCS</td>
<td>0.016 (0.007 to 0.024)</td>
<td>7</td>
<td>91%</td>
<td>0.031 (0.023 to 0.039)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>ED GCS</td>
<td>0.020 (0.006 to 0.034)</td>
<td>3</td>
<td>23%</td>
<td>0.030 (0.020 to 0.039)</td>
<td>2</td>
<td>0%</td>
</tr>
</tbody>
</table>

ES-8
<table>
<thead>
<tr>
<th>Outcome and analysis</th>
<th>tGCS Vs. mGCS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>( I^2 )</th>
<th>tGCS Vs. SMS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>( I^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. setting</td>
<td>0.015 (0.008 to 0.022)</td>
<td>10</td>
<td>87%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>TBI patients</td>
<td>0.009 (-0.002 to 0.020)</td>
<td>3</td>
<td>0%</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Low risk of bias</td>
<td>0.017 (0.015 to 0.020)</td>
<td>5</td>
<td>0%</td>
<td>0.030 (0.022 to 0.037)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>Enrollment before 2006</td>
<td>0.018 (0.011 to 0.024)</td>
<td>10</td>
<td>77%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Enrollment 2006 or later</td>
<td>0.006 (0.001 to 0.011)</td>
<td>2</td>
<td>0%</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Neurosurgical intervention, overall</td>
<td>0.032 (0.020 to 0.043)</td>
<td>7</td>
<td>72%</td>
<td>0.032 (0.025 to 0.039)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Adults or mixed</td>
<td>0.031 (0.018 to 0.044)</td>
<td>6</td>
<td>76%</td>
<td>0.032 (0.025 to 0.039)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Children</td>
<td>0.034 (0.009 to 0.059)</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Excluding studies with potential overlap*</td>
<td>0.032 (0.011 to 0.053)</td>
<td>4</td>
<td>79%</td>
<td>0.038 (0.024 to 0.052)</td>
<td>3</td>
<td>19%</td>
</tr>
<tr>
<td>Out-of-hospital GCS</td>
<td>0.032 (0.011 to 0.053)</td>
<td>4</td>
<td>79%</td>
<td>0.038 (0.024 to 0.052)</td>
<td>3</td>
<td>19%</td>
</tr>
<tr>
<td>ED GCS</td>
<td>0.029 (0.020 to 0.039)</td>
<td>2</td>
<td>0%</td>
<td>0.029 (0.020 to 0.038)</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>U.S. setting</td>
<td>0.032 (0.020 to 0.044)</td>
<td>7</td>
<td>72%</td>
<td>0.032 (0.025 to 0.039)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>TBI patients</td>
<td>0.017 (-0.022 to 0.036)</td>
<td>2</td>
<td>66%</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Low risk of bias</td>
<td>0.026 (0.019 to 0.034)</td>
<td>4</td>
<td>0%</td>
<td>0.029 (0.021 to 0.037)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>Enrollment before 2006</td>
<td>0.033 (0.021 to 0.045)</td>
<td>6</td>
<td>74%</td>
<td>0.032 (0.025 to 0.039)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Enrollment 2006 or later</td>
<td>0.019 (-0.009 to 0.047)</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Severe brain injury, overall</td>
<td>0.050 (0.034 to 0.065)</td>
<td>5</td>
<td>57%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>Adults or mixed</td>
<td>0.046 (0.038 to 0.054)</td>
<td>4</td>
<td>0%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>Children</td>
<td>0.121 (0.068 to 0.174)</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Excluding NTDB studies</td>
<td>0.050 (0.034 to 0.065)</td>
<td>5</td>
<td>57%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>Excluding studies with potential overlap*</td>
<td>0.065 (0.020 to 0.111)</td>
<td>3</td>
<td>76%</td>
<td>0.051 (0.034 to 0.068)</td>
<td>3</td>
<td>74%</td>
</tr>
<tr>
<td>Out-of-hospital GCS</td>
<td>0.041 (0.028 to 0.053)</td>
<td>2</td>
<td>0%</td>
<td>0.051 (0.034 to 0.068)</td>
<td>3</td>
<td>74%</td>
</tr>
<tr>
<td>ED GCS</td>
<td>0.060 (0.028 to 0.093)</td>
<td>3</td>
<td>73%</td>
<td>0.044 (0.030 to 0.059)</td>
<td>2</td>
<td>51%</td>
</tr>
<tr>
<td>U.S. setting</td>
<td>0.050 (0.034 to 0.065)</td>
<td>5</td>
<td>57%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>TBI patients</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Low risk of bias</td>
<td>0.046 (0.038 to 0.053)</td>
<td>3</td>
<td>0%</td>
<td>0.044 (0.035 to 0.053)</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>Enrollment before 2006</td>
<td>0.050 (0.034 to 0.065)</td>
<td>5</td>
<td>57%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>Enrollment 2006 or later</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Emergency intubation, overall</td>
<td>0.034 (0.020 to 0.048)</td>
<td>6</td>
<td>88%</td>
<td>0.040 (0.030 to 0.050)</td>
<td>5</td>
<td>55%</td>
</tr>
<tr>
<td>Adults or mixed</td>
<td>0.034 (0.020 to 0.048)</td>
<td>6</td>
<td>88%</td>
<td>0.040 (0.030 to 0.050)</td>
<td>5</td>
<td>55%</td>
</tr>
<tr>
<td>Children</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
### Key Question 2. Over- and Under-Triage Rates

No study evaluated comparative effects of the tGCS versus the mGCS or SMS on over- or under-triage rates.

### Key Question 3. Effectiveness of Clinical Outcomes

No study evaluated comparative effects of the tGCS versus the mGCS or SMS on clinical outcomes.

### Key Question 4. Interrater Reliability and Ease of Use

Ten studies evaluated interrater reliability or ease of use. Evidence on comparative interrater reliability and ease of use was very limited. There were few head-to-head studies, studies had methodological limitations, and studies on ease of use focused on scoring of written or video patient scenarios. No study assessed ease of use as measured by time to complete assessments or assessor satisfaction.

- The interrater reliability of tGCS and mGCS appears to be high, but evidence was insufficient to determine if there were differences between scales (SOE: Insufficient).
  - Evidence was insufficient to assess effects of patient or assessor characteristics on comparative interrater reliability of the tGCS versus the mGCS (SOE: Insufficient).
  - No study evaluated how comparative interrater reliability or ease of use of the tGCS versus the mGCS varies according to assessment setting (SOE: Insufficient).
- Three studies found the tGCS associated with a lower proportion of correct scores than the mGCS (differences in proportion of correct scores ranged from 6% to 27%), though the difference was statistically significant in only one study (SOE: Low).
  - The proportion of correct GCS scores was generally lowest for assessment of patient scenarios with moderate injury severity in three studies, including one study that evaluated the tGCS and the mGCS (SOE: Low).

---

<table>
<thead>
<tr>
<th>Outcome and analysis</th>
<th>tGCS Vs. mGCS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>$I^2$</th>
<th>tGCS Vs. SMS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>$I^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excluding studies with potential overlap*</td>
<td>0.026 (0.015 to 0.037)</td>
<td>4</td>
<td>68%</td>
<td>0.033 (0.025 to 0.040)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>Out-of-hospital GCS</td>
<td>0.026 (0.015 to 0.037)</td>
<td>4</td>
<td>68%</td>
<td>0.033 (0.025 to 0.040)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>ED GCS</td>
<td>0.048 (0.039 to 0.058)</td>
<td>2</td>
<td>0%</td>
<td>0.048 (0.039 to 0.057)</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>U.S. setting</td>
<td>0.034 (0.020 to 0.048)</td>
<td>6</td>
<td>88%</td>
<td>0.040 (0.030 to 0.050)</td>
<td>5</td>
<td>55%</td>
</tr>
<tr>
<td>TBI patients</td>
<td>0.011 (-0.010 to 0.032)</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Low risk of bias studies</td>
<td>0.037 (0.022 to 0.052)</td>
<td>4</td>
<td>79%</td>
<td>0.046 (0.038 to 0.054)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>Enrollment before 2006</td>
<td>0.038 (0.020 to 0.053)</td>
<td>5</td>
<td>91%</td>
<td>0.040 (0.030 to 0.050)</td>
<td>5</td>
<td>55%</td>
</tr>
<tr>
<td>Enrollment 2006 or later</td>
<td>0.018 (0.005 to 0.031)</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

AUROC=area under the receiver operating characteristic curve; CI=confidence interval; ED= emergency department; GCS= Glasgow Coma Scale; mGCS= motor Glasgow Coma Scale; NTDB= National Trauma Data Bank; SMS= Simplified Motor Score; TBI=traumatic brain injury; tGCS= total Glasgow Coma Scale

*When multiple studies published from the same trauma center, analysis restricted to the most recent study using out-of-hospital GCS scores (excluded Gill 2005,28 Haukoos 2007,40 Acker 201430)
o Three studies found that training or use of a scoring aid increased the proportion of correct scores on both the tGCS and mGCS (increase in proportion of correct scores ranged from 32% to 70%) (SOE: Low).

o Evidence was insufficient to determine effects of level of training or professional background on the proportion of correct scores on the tGCS versus the mGCS (SOE: Insufficient).

- One study found agreement between out-of-hospital and ED scores was similar for the tGCS and the mGCS (SOE: Low).

Discussion

Key Findings and Strength of Evidence

Based on head-to-head studies, we found that the tGCS is associated with slightly better predictive utility than the mGCS, based on the AUROC, a measure of discrimination. The tGCS is better able than the mGCS to discriminate people with trauma who undergo neurosurgical intervention, have severe TBI, or undergo emergency intubation from people who do not experience these outcomes. However, the difference in the AUROC on each of these outcomes was small (<0.05). The tGCS was also better than the mGCS at discriminating trauma patients who died during hospitalization from those who survived hospitalization, but the difference in the AUROC was even smaller (0.01) than for nonmortality outcomes. Findings for the tGCS versus the SMS were similar to findings for the tGCS versus the mGCS for nonmortality outcomes, but the SMS performed slightly worse than the mGCS for in-hospital mortality (difference in AUROC 0.03). This means that for every 100 trauma patients, the tGCS is able to correctly discriminate 1 to 3 more cases of in-hospital mortality from cases without in-hospital mortality than the mGCS or the SMS. Although studies varied in how they defined neurosurgical interventions, severe brain injury, and emergency intubation, findings were generally similar across definitions for these outcomes. Findings for discrimination were robust in sensitivity and subgroup analyses based on the age group analyzed (children vs. adults or mixed), study year (before 2006 or 2006 or later), or risk of bias ratings. However, sensitivity and subgroup analyses were limited by small numbers of studies, particularly for nonmortality outcomes.

Evidence on how age, type of trauma, intubation status, intoxication status, receipt of field interventions, timing of GCS assessment, or level of training of people administering the GCS impacted predictive utility was limited. Few studies specifically evaluated children or patients with TBI, though those available reported findings similar to studies that evaluated adults, mixed populations of adults and children, or mixed trauma patients.

Evidence on interrater reliability and ease of use was limited. For assessment of patients with trauma, only one study, with methodological limitations and imprecise estimates, was included that compared interrater reliability of the tGCS, mGCS, and SMS. Studies that addressed ease of use were limited to those that evaluated whether the measures were scored correctly compared with a reference standard (usually expert assessment). Three studies found that the percentage of correct scores was higher for the mGCS than the tGCS, though in only one study was the difference statistically significant. Limited evidence suggests that errors are more frequent when assessing patient scenarios, indicating moderate injury severity (tGCS scores of 9-13).51,53,56 For both scales, use of a scoring aid or training appears to improve the proportion of correct scores.
No study evaluated other measures of ease of use, such as time to complete the assessment or assessor satisfaction.

One study found that agreement between field and ED scores was similar for the tGCS and mGCS. Although differences between field and ED scores were noted for both scales, the study also found that blood pressure readings changed. Therefore, some differences between field and ED scores may accurately reflect changing status of the patient due to receipt of out-of-hospital interventions and evolving clinical status, rather than true lack of agreement.

**Applicability**

Our findings on predictive utility of different GCS scales appear to have broad applicability to field triage in the United States, as they are based on large studies conducted in U.S. trauma settings in mixed populations of adults and children with various types of trauma. We also restricted study inclusion to studies published in 1995 or later, with most studies conducted in the last 5 to 10 years, suggesting a high level of applicability to use in the context of current trauma systems.

Nonetheless, we identified a number of factors that can impact applicability. Despite the broad applicability of the evidence, its applicability to specific patient populations (e.g., specific type of trauma, age, presence and severity of intoxication, presence of medical comorbidities, and presence of other injuries) is less certain. Limited evidence suggests similar results in children versus mixed populations of adults plus children and in patients with TBI versus mixed trauma populations. No study evaluated how predictive utility varied according to the level or training of field training personnel. In fact, no study that used out-of-hospital scores reported information on the training of the people administering the GCS. Another factor that could impact applicability is that the performance of the tGCS and mGCS may be different when administered soon after injury (in the field) as opposed to later (after field stabilization and destination decisions have been made or after patients have arrived in the ED). A number of studies on predictive utility were conducted in ED settings, which is more controlled and easier to study than field settings. Evidence on the predictive utility from studies conducted in the ED may be of limited applicability to field settings. However, we found that predictive utility was similar in studies that utilized out-of-hospital versus ED GCS scores. We also found no clear differences in estimates of predictive utility when we restricted analyses to studies conducted in U.S. settings or to more recent (2006 or later) studies, which may be more applicable to current U.S. practice.

**Research Recommendations**

Head-to-head observational or randomized studies that assess one set of patients with the tGCS and another set with the SMS or mGCS are needed to understand effects on clinical outcomes as well as risk of over- or under-triage. For over- and under-triage, studies should utilize standardized, validated measures to determine the appropriateness of transport and triage decisions. For predictive utility, prospective studies that independently assess patients using the tGCS and the mGCS or SMS would be useful for confirming the findings of the currently available retrospective studies. Studies are needed to better understand the predictive utility in important subpopulations, including children, older patients, patients with specific types of trauma, and patients who have received field interventions prior to assessment. For patients who are intoxicated or intubated, studies that measure how frequently the tGCS reverts to the mGCS due to the inability to assess the other GCS components would be helpful. Studies that evaluate
how the predictive utility of the tGCS compares with the mGCS or SMS varies according to the level or type of training of assessing personnel in the field are also needed. Finally, studies that assess measures of predictive utility other than discrimination (e.g., calibration, adjusted risk estimates, diagnostic accuracy, risk reclassification) would be useful for providing more complete information regarding predictive utility.

Conclusions

The tGCS is associated with slightly greater discrimination than the mGCS or SMS for in-hospital mortality, receipt of neurosurgical interventions, severe brain injury, and emergency intubation, with differences in the AUROC ranging from 0.01 to 0.05. The clinical significance of small differences in discrimination is likely to be small, and could be offset by factors such as convenience and ease of use. Research is needed to understand how use of the tGCS versus the mGCS or SMS impacts clinical outcomes and risk of over- or under-triage.
References


Introduction

Background

Nature and Burden of Trauma

Unintentional injuries are the leading cause of death among people in the United States ages 1 to 44, and the third leading cause among people ages 45 to 64. Among all age groups, motor vehicle crashes are the first or second leading cause of unintentional injury death. In 2011, there were approximately 40,000,000 emergency department (ED) visits for injuries; of these approximately 2.5 million were due to trauma complications and unspecified injuries. Approximately 18 percent of patients seen in the ED for an injury were transported by emergency medical services (EMS) personnel. Traumatic brain injury (TBI) is an important subset of trauma. Among an estimated 1.7 million annual cases of TBI, there are 52,000 deaths and 275,000 hospitalizations. TBI is a contributing factor to about one-third of injury-related deaths in the United States. From 2001 to 2010, the rate of TBI-related ED visits increased from 421 to 716 per 100,000, though the rate of deaths declined from 18.5 to 17.1 per 100,000 people.

Field Triage of Patients With Trauma

Field triage by EMS is a critical aspect of trauma systems, as it helps to identify potentially seriously injured patients and inform transport decisions. Appropriate decisions regarding transport are crucial because management of severely injured patients in a Level I or a Level II trauma center has been shown to be associated with improved clinical outcomes. On the other hand, unnecessarily triaging patients to high-level trauma care may represent an inefficient use of staff and resources.

EMS personnel must rapidly triage individuals who have undergone trauma in challenging environments. Therefore, EMS personnel must have assessment tools that are easy to use, reliable, and accurate. A key component of field triage for patients with suspected serious injury is level of consciousness assessment. The Glasgow Coma Scale (GCS) is an instrument widely used for assessment of consciousness at the site of injury, in EDs, and in hospitals, as well as to monitor progress or deterioration during treatment. The GCS consists of three items (components): eye (scored 1 to 4), verbal (scored 1 to 5), and motor (scored 1 to 6). Scores on each of these components are added to obtain the total Glasgow Coma Scale (tGCS) score, ranging from 3 to 15. Lower scores on the tGCS indicate lower levels of consciousness, generally correlating with more severe injury associated with poorer prognosis and requiring more intensive care. For patients with TBI, scores of 3 to 8 are generally considered to denote severe head injury, 9 to 12 moderate, and 13 to 15 mild. The 2011 field triage guidelines from the Centers for Disease Control and Prevention (CDC) National Expert Panel recommend transporting patients with tGCS scores of 13 or less to facilities providing the highest level of trauma care.

In some circumstances (e.g., trauma victims who are intoxicated, intubated, or whose other injuries influence response) it may not be possible to accurately assess the verbal and eye components of the tGCS. In these cases, assessments may be primarily based on the motor component of the Glasgow Coma Scale (mGCS) alone. In addition, the mGCS has been proposed for assessment of trauma patients even when the tGCS can be obtained, since only one
item is assessed, potentially increasing ease of use in the field.\textsuperscript{18} mGCS scores of 5 or less are considered an indication of patients with severe injury.\textsuperscript{18,19} The Simplified Motor Score (SMS) has been proposed as a streamlined alternative to the mGCS; it is assessed on a three-point scale (scored 0 to 2, with a score of 0 corresponding to 1 to 4 on the mGCS, 1 corresponding to 5 on the mGCS, and 2 corresponding to 6 on the mGCS).\textsuperscript{20}

Decisions regarding the use of the tGCS versus more simplified versions of the tGCS should be based on their relative performance. The ultimate goal of selecting one risk prediction instrument over another is to improve clinical outcomes (e.g., mortality). However, information on clinical outcomes is often lacking and decisions about their use must often be based on how they perform on intermediate outcomes. Intermediate outcomes include measures of over- or under-triage (i.e., the degree to which patients are unnecessarily transported to a Level I or II trauma center [over-triage] or not transported to a Level I or II trauma center when needed [under-triage]) or predictive utility, as assessed using measures of discrimination (ability of an instrument to distinguish people with the disease from those without), calibration (how well predicted risk correlates with actual risk), standard measures of diagnostic accuracy (e.g., sensitivity, specificity, and predictive values), or adjusted risk estimates (e.g., odds ratio, relative risk, hazards ratio).\textsuperscript{21} Other factors that could inform selection of field triage risk assessment instruments include intra- and interrater reliability and ease of use (e.g., time to administer the instrument and amount of missing data).\textsuperscript{12,22,23}

A number of factors could impact the performance of field assessment instruments. These include variability in patient populations (e.g., type of trauma, demographic characteristics, presence and severity of intoxication, and medical comorbidities), level of training and certification of administering personnel (e.g., emergency medical responder, emergency medical technician [EMT], advanced EMT, paramedic, physician, or nurse\textsuperscript{24}), receipt of field interventions (e.g., medications, intubation), setting (e.g., country, urban vs. rural), or timing of assessment relative to injury occurrence. Evidence about field triage instruments frequently relies on extrapolation from studies conducted in EDs, as this environment is more controlled and easier to study.\textsuperscript{25} However, the performance of the tGCS and mGCS may be different when administered soon after injury by EMS personnel in the field as opposed to later by ED personnel, after destination decisions have already been made and patients have been stabilized with initial interventions.

During the development of field triage guidelines and algorithms by the CDC National Expert Panel in 2011,\textsuperscript{4} use of the mGCS was considered a way to potentially simplify field triage. However, the mGCS was not adopted due in part to lack of evidence about the comparative accuracy and reliability of the mGCS relative to the tGCS. More evidence is now available on the mGCS.

**Rationale for Review**

The purpose of this report is to systematically review the currently available evidence on the comparative predictive utility, reliability, and ease of use of the tGCS and mGCS in field assessment of trauma (with or without TBI), as well as comparative effects on clinical outcomes and early critical resource use. This review provides a synthesis of currently available evidence and gaps in evidence that may be helpful to inform clinical practice and guideline development for field triage of trauma by EMS personnel. The review is the first step of a larger Federal effort to systematically examine the evidence base about out-of-hospital triage decisionmaking and transport of trauma patients, and inform future updates to the Field Triage Guidelines.\textsuperscript{4}
Scope of Review and Key Questions

The report addresses the following Key Questions:

Key Question 1. In patients with known or suspected trauma, what is the predictive utility of the tGCS compared with the mGCS for predicting in-hospital mortality, morbidity, Injury Severity Score of 16 or greater, head Abbreviated Injury Scale (AIS) score greater than 2 (AIS-2005) or greater than 3 (AIS-1998), presence of intracranial hemorrhage, and utilization indicators of severe injury (e.g., receipt of intracranial monitoring within 48 hours of admission, receipt of a neurosurgical intervention within 12 hours of admission, or early intubation [in the field or immediately upon presentation to the ED])?

Key Question 1a. How does predictive utility 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, 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 ED 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 in whom initial EMS transport is to a higher or lower than appropriate level of care, or proportion transferred to a lower or higher level of care)?

Key Question 2a. How do effects on over- and under-triage 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, 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 ED 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., in-hospital mortality, morbidity, quality of life)?

Key Question 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, intubation or receipt of medication in the field), the training and
Key Question 4. In patients with known or suspected trauma, what is the comparative reliability (e.g., interrater and intrarater 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?

Key Question 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, 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 ED or urban vs. rural location)?

The research questions used to guide this review were initially developed by the National Highway Traffic Safety Administration and revised with input from a Technical Expert Panel. The Key Questions focus on predictive utility, over- and under-triage, clinical outcomes of the tGCS versus the mGCS or SMS, as well as reliability and ease of use.

Key Question 1 addresses the predictive utility of the tGCS compared with the mGCS for predicting clinical outcomes (in-hospital 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 neurosurgical interventions such as early surgery or intracranial pressure monitoring) and as a marker of need for tertiary trauma care. Key Question 1 does not 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 one of the factors used to determine who is transported to tertiary trauma care. Therefore, receipt of tertiary trauma care does not represent a marker of injury severity independent from the GCS score.

Key Question 2 addresses the impact of the tGCS compared with the mGCS on clinical decisionmaking based on rates of over- and under-triage, an intermediate outcome, as determined by the proportion of patients transported to a higher or lower than appropriate level of care based on a standardized definition for trauma center need. We also considered rates of transfer to higher or lower levels of care as a potential marker of over- or under-triage. Measuring over- and under-triage is a challenge because factors other than findings on field triage assessment scales, including other patient characteristics (e.g., mechanism of injury, hemodynamic instability, respiratory distress, comorbidities), geographic proximity, and availability of resources, also impact triage decisions. Interpretation of transfer rates is particularly challenging since minimally injured over-triaged patients may be discharged directly home rather than transferred to lower-level care. In addition, some over-triage may be acceptable in order to prevent under-triage, which may be more likely to result in adverse clinical outcomes while over-triage may primarily represent inefficient use of resources and increased costs without necessarily
adversely impacting clinical outcomes.\textsuperscript{32} 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. The analytic framework (Figure 1) used to guide this review is shown below.

The analytic framework shows the target populations, interventions, and health outcomes examined, with numbers corresponding to the Key Questions.

**Figure 1. Analytic framework**

EMS=emergency medical services; KQ=Key Question; mGCS=Motor Glasgow Coma Scale; tGCS=total Glasgow Coma Scale; vs.=versus

*Based on tGCS score of \( \leq 13 \) or mGCS score of \( \leq 5 \)
Methods

The methods for this Comparative Effectiveness Review (CER) follow the guidance in the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews.33

Scope Development

The initial Key Questions were provided by the National Highway Traffic Safety Administration (NHTSA). The Key Questions were further developed and the final protocol developed with additional input from NHTSA and a Technical Expert Panel (TEP) convened for this report. The TEP consisted of eight experts in adult and pediatric field triage, emergency medical services (EMS), trauma surgery, general surgery, critical care, and emergency medicine. TEP members disclosed financial and other conflicts of interest prior to participation. The AHRQ Task Order Officer and the investigators reviewed the disclosures and determined that the TEP members had no conflicts of interest that precluded participation.

The final protocol was posted on the AHRQ Web site on February 28, 2016 at: http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2187. The protocol was also registered in the PROSPERO international database of prospectively registered systematic reviews (registration number CRD42016035944).34

Literature Search Strategy

A research librarian conducted searches in the Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, CINAHL®, PsycINFO®, HaPI (Health and Psychosocial Instruments), and Ovid MEDLINE® (January, 1995 to June, 2016), limiting to English-language abstracts. Search strategies are provided in Appendix A. We restricted search start dates to January, 1995 to improve applicability to current U.S. trauma care; only five states had fully implemented trauma systems in the early 1990s.35 In addition, the first studies to compare the predictive utility of the motor component of the Glasgow Coma Scale (mGCS) with the total score of the Glasgow Coma Scale (tGCS) were published in 1998 and 2003.18,19

We also hand-searched the reference lists of relevant studies and searched for unpublished studies in ClinicalTrials.gov.

Study Selection

We developed criteria for inclusion and exclusion of studies based on the Key Questions and the populations, interventions, comparators, outcomes, timing, types of studies, and setting (PICOTS) approach, in accordance with the Methods Guide for Effectiveness and Comparative Effectiveness Reviews.33 Inclusion and exclusion criteria are summarized below and described in more detail in Appendix B. Abstracts were reviewed by two investigators, and all citations deemed potentially appropriate for inclusion by at least one of the reviewers were retrieved for full-text review. Two investigators then independently reviewed all full-text articles for final inclusion. Investigators did not review, assess, or screen papers that they authored. Inclusion was restricted to English-language articles. Discrepancies were resolved by discussion and consensus, with a third investigator if necessary.
A list of the included studies appears in Appendix C; a list of excluded studies and primary reasons for exclusion can be found in Appendix D.

**Population and Conditions of Interest**

For all Key Questions we included studies of children and adults with known or suspected trauma. Although the population of interest was patients with blunt trauma, we included studies of general trauma patients. We excluded studies of individuals without trauma, studies that focused on nonblunt trauma patients, or studies of patients with and without trauma in which the proportion without trauma was more than 10 percent and results were not reported separately for patients with trauma.

**Interventions, Comparisons, and Study Designs**

We focused on studies of the total Glasgow Coma Scale (tGCS), the motor component of the Glasgow Coma Scale (mGCS), and the Simplified Motor Score (SMS). For studies evaluating measures of diagnostic accuracy (sensitivity, specificity, predictive values), we focused on studies that used standard cutoff scores of 13 or less on the tGCS, 5 or less on the mGCS, or 1 or 0 on the SMS to indicate people who require high-level trauma care, but also included studies that used alternative cutoffs or modifications of the tGCS and mGCS. We excluded studies that evaluated the utility of mGCS or tGCS in combination with other predictors in a multi-item risk assessment or triage instrument.

For all Key Questions, we included cohort studies and randomized controlled trials (RCTs) that directly compared the tGCS with the mGCS or SMS. For Key Question 4 (reliability and ease of use), we also included cross-sectional studies and studies that assessed one of these scales. For Key Question 1a (predictive utility) we included studies that assessed one of these scales if they addressed one of the subpopulations specified in the Key Questions (e.g., children, intoxicated, intubated, traumatic brain injury [TBI] patients) not addressed well in the head-to-head studies. We restricted Key Questions 2 and 3 to head-to-head studies because given the large number of other factors that impact these outcomes, it is not possible to isolate the effects of risk assessment scales on over- or under-triage or clinical outcomes from single-arm studies.

**Outcomes**

For Key Question 1, we included measures of predictive utility for in-hospital mortality, morbidity, markers of severe injury (e.g., Abbreviated Injury Scale [AIS] score of ≥4 or Injury Severity Score [ISS] of ≥16), or utilization indicators of severe injury (e.g., receipt of intracranial monitoring within 48 hours of admission, receipt of a neurosurgical intervention within 12 hours of admission, or receipt of early intubation [in the field or immediately upon arrival to the emergency department (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), or risk reclassification rates.

For Key Question 2, we included studies that reported the proportion of patients who were over- or under-triaged (e.g., the proportion of patients transported to a higher or lower than appropriate level of care according to a standardized definition for trauma center need or the proportion transferred to a higher or lower level of care).

For Key Question 3, we included studies that reported clinical outcomes, including mortality (prior to hospital arrival, in the ED, or after hospital admission), measures of morbidity,
including cognitive impairment and medical complications related to trauma, and quality of life, including functional capacity at discharge or followup.

For Key Question 4, we included outcomes that assessed reliability (e.g., interrater and intrarater kappa) or ease of use (e.g., time to complete, measures of missing data, user-reported satisfaction).

Timing and Setting

For all Key Questions we included prospective and retrospective studies in which the tGCS, mGCS, or SMS were administered soon after injury (conducted in the field/out-of-hospital setting by Emergency Medical Services personnel) or immediately upon arrival to the ED, or that were based on trauma registry data collected in the field or in the ED. We excluded studies in which the Glasgow Coma Scale (GCS) was administered after more than 4 hours in the ED or hospital or when it was administered after hospital admission. We also excluded studies conducted in the developing world, which may be less applicable to U.S. trauma care settings.

Data Abstraction and Data Management

A single investigator abstracted information on study design, year, geographic location, patient characteristics (i.e., demographics, type and mechanism of trauma, type of injury, tGCS scores, severity of injury, intoxication status, systolic blood pressure, intubation or receipt of medication in the field), the proportion of patients who experienced outcomes, which triage instrument was used, timing of triage assessment, cutoff scores used (for studies that evaluated sensitivity and specificity), the training and experience of the person administering the GCS, assessment setting (in the field or upon ED presentation), and results relevant to each Key Question. All data abstractions were reviewed by a second investigator for accuracy and discrepancies were resolved through discussion and consensus.

Assessment of Methodological Risk of Bias of Individual Studies

We assessed risk of bias of included studies using predefined criteria. Our methods for assessing risk of bias are based on the recommendations in the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. For Key Question 1 (predictive utility), we applied the Quality in Prognostic Studies (QUIPS) tool for prognostic studies. The QUIPS tool includes domains on study participation, study attrition, prognostic factor measurement, outcomes measurement, study confounding, and statistical analysis and reporting. For Key Question 4 (reliability and ease of use), we assessed risk of bias using criteria adapted from the Quality Assessment of Diagnostic Accuracy Studies (QUADAS). This includes criteria about patient selection, whether raters were blinded to other ratings, how the scores from different assessments were compared and the situation and timing of measurement. Two investigators independently assessed risk of bias for each study. Differences were resolved by discussion; we used a third rater to resolve discrepancies if needed (Appendix E). No study met inclusion criteria for Key Questions 2 or 3.

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. For example, low risk of bias studies on predictive utility select all or a random subset of patients who meet predefined criteria, report low attrition, perform the risk assessment scale in all patients, measure
outcomes accurately, and in all patients, assess and measure important confounders, use appropriate statistical methods, and avoid selective reporting of results.

Studies rated “moderate” 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 moderate risk of bias category is broad, and studies with this rating vary in their strengths and weaknesses. The results of some moderate 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 are at least as likely to reflect flaws in the study design as the true difference between the compared interventions. We did not exclude studies rated high risk of bias a priori, but performed sensitivity analyses in which such studies were excluded.

Assessing Research Applicability

Applicability is defined as the extent to which the effects observed in published studies are likely to reflect the expected results when a specific intervention is applied to the population of interest under “real-world” conditions. It is an indicator of the extent to which research included in a review might be useful for informing clinical decisions in specific situations. Because applicability depends on the perspective of the user of the review, we did not assign a rating for applicability (such as “high” or “low”). Rather, factors important for understanding the applicability of studies were recorded, such as population characteristics (age, type of trauma, intoxication status), setting (United States vs. other country, out-of-hospital vs. ED assessment), and type and level of training of people administering the GCS were recorded and assessed in subgroup and sensitivity analyses. We also recorded the funding source for studies. Most studies on predictive utility reported the area under the receiver operating characteristic curve (AUROC), a measure of discrimination. The AUROC value represents the probability that a patient who experiences an outcome will have a higher score on the triage scale than a person who does not experience the outcome. We did not identify published recommendations on how to interpret the magnitude of differences in the AUROC value. Therefore, we defined a small difference in the AUROC a priori as a difference of less than 0.05, moderate as a difference of 0.05 to 0.10, and large as a difference of greater than 0.10. For measuring interrater or intrarater reliability, we defined a meaningful difference in kappa values as 0.20 or greater.

Data Synthesis and Rating the Body of Evidence

We performed random effects meta-analysis to calculate pooled differences on the AUROC from head-to-head studies of the tGCS versus the mGCS or SMS using the DerSimonian-Laird random effects model with Stata/IC 13.1 (StataCorp LP, College Station, TX). We measured statistical heterogeneity using the I² statistic. The DerSimonian-Laird estimator can result in confidence intervals (CIs) for the pooled estimate that are too narrow, particularly when statistical heterogeneity is present. We also performed analyses using the Profile Likelihood method. The AUROC value represents the probability that a patient who experiences an outcome will have a higher score on the triage scale than a person who does not experience the outcome. We defined a small difference in the AUROC a priori as a difference of less than 0.05, moderate
as a difference of 0.05 to 0.10, and large as a difference of greater than 0.10. Therefore, we repeated analyses using an alternative random effects model, the profile likelihood method, which may provide more accurate confidence limits. Most studies reported estimated AUROCs with associated 95 percent CIs. When a study only reported the point estimate of AUROC without providing a 95 percent CI or a standard error, we imputed the standard error using the average standard error from other studies in the same meta-analysis. In all studies, the mGCS or SMS scores were derived from the tGCS and applied to the same patient population (i.e., the risk assessment scales were not applied independently). To account for this nonindependence, we assumed a correlation of 0.5 when comparing the tGCS with the mGCS, or SMS in the primary analysis. Two studies reported data that allowed us to calculate the correlations between the AUROC for the tGCS and the mGCS or SMS, which ranged from 0.5 to 0.9 depending on the outcome and comparison. Therefore, 0.5 is a conservative assumption for the correlation. A high correlation is expected given mGCS or SMS scores are a subset of tGCS applied on the same population. Sensitivity analyses were conducted assuming correlations of 0.3 and 0.8; results were similar and not separately reported.

Primary analyses were stratified by the age group evaluated in the study (children vs. adults or mixed populations). We performed additional sensitivity and subgroup analyses based on timing of GCS assessment (field vs. ED), study dates (all data collected in 2006 or later or some or all data collected prior to 2006), country (United States vs. other), and risk of bias rating. For the primary analysis, we included studies conducted using the National Trauma Data Bank (NTDB). In 2012, 805 hospitals submitted data to the NTDB. Because populations evaluated in single trauma centers or systems could be included (in part or in full) in the NTDB, we performed a sensitivity analysis in which NTDB studies were excluded. For the primary analysis, we included multiple studies from the same trauma center or system unless there was clearly complete overlap in the populations assessed. In sensitivity analyses, we restricted analyses to studies from each trauma center that utilized field GCS scores; if multiple studies utilized field GCS scores we utilized the study that evaluated more recent data.

Grading the Body of Evidence for Each Key Question

For all comparisons and outcomes we assessed the strength of evidence (Appendix F) using the approach described in the Methods Guide for Effectiveness and Comparative Effectiveness Reviews and Methods Guide for Medical Test Reviews, based on the overall risk of bias (graded low, moderate, or high); the consistency of results across studies (graded consistent, inconsistent, or unable to determine when only one study was available); the directness of the evidence linking the intervention and health outcomes (graded direct or indirect); the precision of the estimate of effect, based on the number and size of studies and CIs for the estimates (graded precise or imprecise); and reporting bias (suspected or undetected). Assessments of reporting bias were based on whether studies defined and reported primary outcomes and whether we identified relevant unpublished studies.

Based on our assessments on the domains described above, we graded the strength of evidence for each Key Question using the four key categories recommended in the Methods Guide for Comparative Effectiveness and Effectiveness Reviews. Randomized controlled trials and cohort studies on predictive utility started as “high” strength of evidence and were graded down based on the presence of deficiencies in the domains. Because observational studies on predictive utility started as high, we did not consider factors for upgrading such as dose-response relationship, magnitude of effects, or impact of plausible confounders. A high grade indicates
high confidence that the evidence reflects the true effect and that further research is very unlikely to change our confidence in the estimate of effect. A “moderate” grade indicates moderate confidence that the evidence reflects the true effect and further research may change our confidence in the estimate of effect and may change the estimate. A “low” grade indicates low confidence that the evidence reflects the true effect and further research is likely to change the confidence in the estimate of effect and is likely to change the estimate. An “insufficient” grade indicates evidence either is unavailable or is too limited to permit any conclusion, due to the availability of only high risk of bias studies, extreme inconsistency, or extreme imprecision.

See Appendix G for the strength of evidence table.

**External Review**

Peer reviewers with expertise in trauma and triage provided written comments on the draft report. The AHRQ Task Order Officer and an Evidence-based Practice Center Associate Editor also provided comments and editorial review. The draft report was posted on the AHRQ Web site for 4 weeks for public comment from June 29, 2016 to July 26, 2016. A disposition of comments report with authors’ responses to the peer and public review comments will be posted after publication of the final CER on the public Web site.
Results

Results of Literature Search

Results of the literature search and selection process are summarized in the literature flow diagram (Figure 2). Database searches resulted in 4,412 potentially relevant citations. After dual review of abstracts and titles, 698 articles were selected for full-text review. After dual review of full-text articles, 32 studies were included. Data extraction and risk of bias assessment tables for included studies by Key Question are available in Appendixes H through L.
HaPI = Health and Psychosocial Instruments; n = sample size

*Cochrane databases include the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews
†Other sources include reference list, experts, etc.
‡Two studies were used for Key Question 1 and Key Question 4
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 component of the GCS (mGCS) score for predicting in-hospital mortality, morbidity, Injury Severity Score (ISS) of 16 or greater, head Abbreviated Injury Scale (AIS) score greater than 2 (AIS-2005) or greater than 3 (AIS-1998), presence of intracranial hemorrhage, and utilization indicators of severe injury (e.g., receipt of intracranial monitoring within 48 hours of admission, receipt of a neurosurgical intervention within 12 hours of admission, or early intubation [in the field or immediately upon presentation to the emergency department (ED)])?

Key Points

- In-hospital mortality
  - For the tGCS versus the mGCS, the pooled area under the receiver operating characteristic curve (AUROC) was 0.877 (95% confidence interval [CI] 0.847 to 0.906) versus 0.855 (95% CI 0.822 to 0.888), with a pooled mean difference of 0.015 (95% CI 0.009 to 0.022; $I^2=85\%$), based on 12 studies (strength of evidence [SOE]: Moderate).
  - For the tGCS (cutoff of \(\leq 13\)) versus the mGCS (cutoff of \(\leq 5\)), differences in sensitivity ranged from 0 percent to 3 percent; difference in specificity ranged from 0 percent to 5 percent in favor of the mGCS, though the CIs overlapped in most studies (SOE: Low).
  - For the tGCS versus the Simplified Motor Score (SMS), the pooled AUROC was 0.884 (95% CI 0.852 to 0.916) versus 0.840 (95% CI 0.802 to 0.878), for a mean difference of 0.030 (95% CI 0.024 to 0.036, $I^2=0\%$), based on five studies (SOE: Moderate).
  - One study found the out-of-hospital tGCS (cutoff of \(\leq 13\)) associated with slightly higher sensitivity versus the SMS (cutoff of \(\leq 1\)) (75%, 95% CI 73 to 76 vs. 72%, 95% CI 70 to 74) and slightly lower specificity (88%, 95% CI 87 to 88 vs. 89%, 95% CI 89 to 87) (SOE: Low).
  - For the mGCS versus the SMS, the mean difference in the pooled AUROC was 0.014 (95% CI 0.006 to 0.021, $I^2=0\%$), based on four studies (SOE: Moderate).

- Neurosurgical intervention
  - For the tGCS versus the mGCS, the pooled AUROC was 0.786 (95% CI 0.729 to 0.842) versus 0.754 (95% CI 0.688 to 0.819), with a mean difference of 0.032 (95% CI 0.020 to 0.043; $I^2=72\%$), based on seven studies (SOE: Moderate).
  - Two studies found inconsistent results for sensitivity (51%, 95% CI 50 to 52 vs. 46%, 95% CI 45 to 48 and 63%, 95% CI 38 to 84 vs. 68%, 95% CI 43 to 87) and no clear differences in specificity (86%, 95% CI 86 to 86 vs. 88%, 95% CI 88 to 88 and 82%, 95% CI 80 to 84 vs. 83%, 95% CI 81 to 85) between out-of-hospital tGCS (cutoff of \(\leq 13\)) versus the mGCS (cutoff of \(\leq 5\)) in accuracy for identifying people undergoing craniotomy (SOE: Insufficient for sensitivity, low for specificity).
For the tGCS versus the SMS, the pooled AUROC was 0.809 (95% CI 0.766 to 0.853) versus 0.769 (95% CI 0.711 to 0.827), with a mean difference of 0.032 (95% CI 0.025 to 0.039, $I^2=0\%$), based on five studies (SOE: Moderate).

One study found the out-of-hospital tGCS (cutoff of ≤13) associated with higher sensitivity than the SMS (cutoff of ≤1) for identifying patients who underwent neurosurgical intervention (60%, 95% CI 56 to 63 vs. 53%, 95% CI 49 to 56) and slightly lower specificity (85%, 95% CI 84 to 85 vs. 86%, 95% CI 86 to 87) (SOE: Low).

For the mGCS versus the SMS, the mean difference in the pooled AUROC was 0.002 (95% CI -0.005 to 0.010, $I^2=0\%$), based on four studies (SOE: Moderate).

Severe brain injury

For the tGCS versus the mGCS, the pooled AUROC was 0.791 (95% CI 0.734 to 0.827) versus 0.720 (95% CI 0.666 to 0.774), with a mean difference of 0.050 (95% CI 0.034 to 0.065; $I^2=57\%$), based on five studies (SOE: Moderate).

One study found no difference between out-of-hospital tGCS (cutoff of ≤13) versus the mGCS (cutoff of ≤5) in sensitivity (62%, 95% CI 55% to 68% vs. 61%, 95% CI 54% to 67%) or specificity (85%, 95% CI 83 to 88 vs. 89%, 95% CI 88 to 91) for identifying people with severe head injury (defined as head AIS score of ≥4) (SOE: Low).

For the tGCS versus the SMS, the pooled AUROC was 0.763 (95% CI 0.710 to 0.815) versus 0.713 (95% CI 0.654 to 0.771), with a mean difference of 0.048 (95% CI 0.038 to 0.059, $I^2=72\%$), based on five studies (SOE: Moderate).

One study found out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity than the SMS (cutoff of ≤1) for severe brain injury based on presence of head computed tomography (CT) imaging findings (45%, 95% CI 44 to 46 vs. 41%, 95% CI 40 to 42) and similar specificity (89%, 95% CI 89 to 90 vs. 90%, 95% CI 90 to 91) (SOE: Low).

For the mGCS versus the SMS, there was no difference in the AUROC (mean difference 0.000, 95% CI -0.008 to 0.007, $I^2=0\%$), based on four studies (SOE: Moderate).

Emergency intubation

For the tGCS versus the mGCS, the pooled AUROC was 0.865 (95% CI 0.830 to 0.901) versus 0.822 (95% CI 0.775 to 0.870), with a mean difference of 0.034 (95% CI 0.020 to 0.048; $I^2=88\%$), based on six studies (SOE: Moderate).

One study found very small differences between out-of-hospital tGCS (cutoff of ≤13) versus the mGCS (cutoff of ≤5) in accuracy for identifying people who underwent emergency intubation (sensitivity 84%, 95% CI 83 to 84 vs. 81%, 95% CI 81 to 82; and specificity 90%, 95% CI 90 to 90 vs. 92.0%, 95% CI 92 to 92) (SOE: Low).

For the tGCS versus the SMS, the pooled AUROC was 0.843 (95% CI 0.823 to 0.864) versus 0.783 (95% CI 0.747 to 0.819), with a mean difference of 0.040 (95% CI 0.030 to 0.050, $I^2=55\%$), based on five studies (SOE: Moderate).

One study found the out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity than the SMS (cutoff of ≤1) for identifying people who underwent emergency intubation (76%, 95% CI 74 to 77 vs. 73%, 95% CI 71 to 74) and slightly lower specificity (89%, 95% CI 89 to 89 vs. 91%, 95% CI 90 to 91) (SOE: Low).
For the mGCS versus the SMS, there was no difference in the AUROC (mean difference 0.000, 95% CI -0.007 to 0.007, I²=0%), based on four studies (SOE: Moderate).

- Trauma center need
  - Two studies found small differences between the tGCS versus the mGCS in the AUROC (0.617 vs. 0.609 and 0.641 vs. 0.603), sensitivity (30% vs. 27% and 28% vs. 25%), and specificity (93% vs. 95% and 94% vs. 95%) for trauma center care need (defined as ISS score of >15, intensive care unit [ICU] admission >24 hours, need for urgent surgery, or death in the ED) (SOE: Low).

- Severe injury
  - Two studies found the tGCS was better able to discriminate those with major injury (defined as an ISS score of >15) from those without major injury (AUROC 0.72, 95% CI 0.71 to 0.72 vs. 0.68, 95% CI 0.68 to 0.69 and 0.65, 95% CI 0.65 to 0.61, 95% CI 0.60 to 0.61) (SOE: Low).
  - One study found the out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity and slightly lower specificity than the mGCS (cutoff of ≤5) for identifying people with major injury (defined as an ISS score of >15) from those without major injury (sensitivity 31%, 95% CI 31 to 32 vs. 28%, 95% CI 28 to 28 and specificity 91%, 95% CI 91 to 91 vs. 93%, 95% CI 93 to 93). (SOE: Low).

**Detailed Synthesis**

Eighteen studies directly compared the predictive utility of the tGCS versus the mGCS (12 studies) and/or the SMS (6 studies, Table 1, Appendix H). All studies were retrospective analyses in which the mGCS or SMS scores were taken from the tGCS (i.e., the tGCS and mGCS or SMS were not assessed independently). Sample sizes ranged from 96 to 811,143. Fifteen studies were conducted in the United States, two studies in Europe, and one in Canada. Four studies restricted enrollment to children; the other studies enrolled adults or mixed populations of adults and children. Four studies utilized data collected in or after 2006. GCS scores were obtained during out-of-hospital assessment in 11 studies, in the ED in 5 studies, and in mixed (ED or out-hospital) settings in 2 studies. The assessment setting was mixed or unclear in four studies. Four studies focused on patients with TBI and the remainder evaluated mixed trauma populations. Among studies that enrolled mixed trauma patients, the proportion of trauma patients with TBI ranged from 5 to 18 percent in studies that reported this information; none of these studies reported results in subgroup of patients with TBI. No study reported the proportion of intoxicated patients. In two studies, the proportion of patients who underwent out-of-hospital intubation was 0.3 percent and 3.5 percent, it was unclear when the GCS was assessed in intubated patients. Thirteen studies were rated moderate risk of bias and five studies were rated low risk of bias. Eight studies did not report attrition and seven studies reported missing data in more than 20 percent of patients. Most studies on predictive utility focused on measures of discrimination or diagnostic accuracy; adjustment for confounders is generally not performed for either of these measures.
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Settings</th>
<th>Years of Study</th>
<th>Assessment Timing Measures and/or Scores Compared</th>
<th>N</th>
<th>Population Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acker et al., 2014&lt;sup&gt;51&lt;/sup&gt;</td>
<td>USA, Colorado Urban 2 Level 1 pediatric trauma centers 2002 to 2011</td>
<td>ED A: tGCS B: mGCS</td>
<td>2,231</td>
<td></td>
<td>Age (mean, years): 6.9 (SD 5.8) Male: 65% Race: NR TBI: 100% ISS (median): 17 (IQR: 10-26)</td>
</tr>
<tr>
<td>Al-Salamah et al., 2004&lt;sup&gt;52&lt;/sup&gt;</td>
<td>Canada, Ontario Trauma registry 72% urban, 28% suburban or rural 1994 to 2002</td>
<td>Out-of-hospital A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>795</td>
<td></td>
<td>Age (mean, years): 44 (SD 21) Male: 70% Race: NR TBI: NR ISS: NR</td>
</tr>
<tr>
<td>Beskind et al., 2014&lt;sup&gt;17&lt;/sup&gt;</td>
<td>USA, Southern Arizona Urban, University Health Network Level 1 trauma center 2008 to 2010</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>9,816</td>
<td></td>
<td>Age (median, years): 32 (IQR: 20-51) Male: 65.5% Race: NR TBI: NR ISS ≥16: 11.7%</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012&lt;sup&gt;54&lt;/sup&gt;</td>
<td>USA, Ohio Urban, hospitals Trauma and nontrauma centers 2002 to 2007</td>
<td>Out-of-hospital A: tGCS ≤13 B: SMS ≤1</td>
<td>52,412</td>
<td></td>
<td>Age (mean, years): 53 Male: 55.9% White: 79.9% Black: 13.5% Hispanic: 1.5% Other race: 1.7% Race not documented: 3.4% TBI: 15.2% ISS (median): 9 ISS &gt;15: 26.6%</td>
</tr>
<tr>
<td>Cicero and Cross, 2013&lt;sup&gt;55&lt;/sup&gt;</td>
<td>USA Trauma registry* 2007 to 2009</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>104,035</td>
<td></td>
<td>Age (mean, years): 12.6 (SD 5.5) Male: 67% Nonwhite race: 38% TBI: NR ISS (mean): 9.9 (SD 10.3) Major injury (ISS &gt;15): 15%</td>
</tr>
<tr>
<td>Corrigan et al., 2014&lt;sup&gt;56&lt;/sup&gt;</td>
<td>USA Trauma registry* 2007 to 2010</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>77,470</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Settings</td>
<td>Assessment Timing</td>
<td>N</td>
<td>Population Characteristics</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>---------</td>
<td>----------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Davis et al., 2006&lt;sup&gt;57&lt;/sup&gt;</td>
<td>USA, California (San Diego) Urban, other data NR Date NR</td>
<td>Out-of-hospital and ED A: tGCS B: mGCS</td>
<td>12,882</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Eken et al., 2009&lt;sup&gt;58&lt;/sup&gt;</td>
<td>Turkey Tertiary care ED of hospital Level IV trauma center 2006</td>
<td>ED A: tGCS B: mGCS</td>
<td>185</td>
<td>Age (median, years): 59 (range: 18-97) Male: 64% Race: NR TBI: NR ISS: NR</td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005&lt;sup&gt;59&lt;/sup&gt;</td>
<td>USA, California (Loma Linda) Urban, University Level 1 trauma center and children's hospital 1990 to 2002</td>
<td>ED A: tGCS B: mGCS C: SMS</td>
<td>8,412</td>
<td>Age (median, years): 24 (IQR: 15-38) Male: 71.5% Race: NR TBI: 17.1% ISS: NR</td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2006&lt;sup&gt;60&lt;/sup&gt;</td>
<td>USA, California (Loma Linda) Urban, University Level 1 trauma center and children's hospital 1990 to 2002</td>
<td>Out-of-hospital A: tGCS B: mGCS C: SMS</td>
<td>7,233</td>
<td>Age (median, years): 24 (IQR: 16-38) Male: 70% Race: NR TBI: 17% ISS: NR</td>
<td></td>
</tr>
<tr>
<td>Haukoos et al., 2007&lt;sup&gt;61&lt;/sup&gt;</td>
<td>USA, Colorado Urban, Denver Health Medical Center Level 1 trauma center 1995 to 2002</td>
<td>ED A: tGCS B: mGCS C: SMS</td>
<td>21,170</td>
<td>Age (median, years): 32 (IQR: 21-45) Male: 71% Race: NR TBI: 14% ISS score (median): 9 (IQR: 2-14)</td>
<td></td>
</tr>
<tr>
<td>Healey et al., 2003&lt;sup&gt;62&lt;/sup&gt;</td>
<td>USA Trauma registry* 1994 to 2001</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>202,255</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Holmes et al., 2005&lt;sup&gt;63&lt;/sup&gt;</td>
<td>USA, California (Davis) Level 1 trauma center 1998 to 2001</td>
<td>ED A: tGCS B: mGCS</td>
<td>2,043</td>
<td>Ages ≤2 years: 16% Ages &gt;2 years: 84% Male: NR Race: NR TBI: 5% ISS: NR</td>
<td></td>
</tr>
<tr>
<td>Kupas et al., 2016&lt;sup&gt;64&lt;/sup&gt;</td>
<td>USA, Pennsylvania Level I, II, III, or IV trauma centers 1999 to 2013</td>
<td>Out-of-hospital A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>370,392</td>
<td>Age (median, years): 50 Male: 62% White: 79% Black: 14% Asian: 0.9% TBI: NR (88% blunt trauma) ISS &gt;15: 29%</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Settings</td>
<td>Assessment Timing Measures and/or Scores Compared</td>
<td>N</td>
<td>Population Characteristics</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>-----</td>
<td>-----------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Ross et al., 1998</td>
<td>USA, New Jersey&lt;br&gt;Level 1 trauma center&lt;br&gt;1994 to 1996</td>
<td>Out-of-hospital&lt;br&gt;A: tGCS score ≤13&lt;br&gt;B: mGCS score ≤5</td>
<td>1,410</td>
<td>Age (mean, years): 37.1 (range: 13-95)&lt;br&gt;Male: 69%&lt;br&gt;Race: NR&lt;br&gt;TBI: NR&lt;br&gt;ISS (mean): 14.4&lt;br&gt;ISS (median): 13</td>
<td></td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>USA, Colorado&lt;br&gt;Urban, Denver Health Medical Center&lt;br&gt;Level 1 trauma center&lt;br&gt;1999 to 2008</td>
<td>Out-of-hospital&lt;br&gt;A: tGCS ≤13&lt;br&gt;B: mGCS score ≤5&lt;br&gt;B: SMS ≤1</td>
<td>19,408</td>
<td>Age (median, years): 33 (IQR: 22-48)&lt;br&gt;Male: 71%&lt;br&gt;Race: NR&lt;br&gt;TBI: 18%&lt;br&gt;ISS (median): 9 (IQR: 4-17)</td>
<td></td>
</tr>
<tr>
<td>Van de Voorde et al., 2008</td>
<td>Belgium&lt;br&gt;Pediatric trauma registry (PENTA) 2005</td>
<td>Out-of-hospital and ED&lt;br&gt;A: tGCS score ≤13&lt;br&gt;B: mGCS score ≤5</td>
<td>96</td>
<td>Age (mean, years): 8.2 (SD 5.3)&lt;br&gt;Male: 59%&lt;br&gt;Race: NR&lt;br&gt;TBI: NR&lt;br&gt;ISS (median): 16</td>
<td></td>
</tr>
</tbody>
</table>

ED=emergency department; IQR=interquartile range; ISS=Injury Severity Score; mGCS=motor Glasgow Coma Scale; n=number; NR=not reported; NTDB=National Trauma Data Bank; PENTA=pediatric trauma registry; SD=standard deviation; SMS=Simplified Motor Scale; TBI=traumatic brain injury; tGCS=total Glasgow Coma Scale; USA=United States of America

*Patients from the NTDB data set
Four studies were based on analyses of the NTDB database, but evaluated different populations or outcomes. Sample sizes ranged from 77,470 to 811,143. One of the NTDB studies focused on children, one focused on adults, and two evaluated mixed populations. There were also two studies that used data from a trauma center in Loma Linda, California and three studies that used data from the Denver area trauma system in which there could be some overlap in the populations assessed. In sensitivity analyses, we excluded two studies from these trauma systems that focused on GCS scores obtained in the ED, since other studies from these trauma systems evaluated out-of-hospital GCS scores around the same time period.

The most commonly evaluated outcome was in-hospital mortality. Other outcomes reported in at least five studies were severe brain injury, receipt of neurosurgical intervention, and intubation (Table 2). The proportion of patients who experienced in-hospital mortality ranged from 3 percent to 18 percent; the proportion with severe brain injury (defined in one study as CT imaging findings of skull fracture, contusion, or hemorrhage, or acute intervention for TBI [neurosurgical procedure, hospitalization >2 days, antiepileptic medications for >7 days]; the others defined severe brain injury based on CT imaging alone) ranged from 5 percent to 39 percent; the proportion who underwent a neurosurgical intervention (defined as craniotomy in two studies and as a composite outcome including various neurosurgical procedures, ventriculostomy, and/or intracerebral pressure monitoring in the others) ranged from 1.5 percent to 10 percent; and the proportion who were intubated (out-of-hospital, ED, or both) ranged from 4 percent to 26 percent.
### Table 2. Proportion of patients experiencing outcomes in head-to-head studies on predictive utility

<table>
<thead>
<tr>
<th>Study, Year</th>
<th>In-Hospital Mortality</th>
<th>Neurosurgical Intervention</th>
<th>Severe Brain Injury</th>
<th>Severe Injury</th>
<th>Intubation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acker et al., 2014&lt;sup&gt;41&lt;/sup&gt;</td>
<td>8.4%</td>
<td>10.4%*</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Al-Salamah et al., 2004&lt;sup&gt;42&lt;/sup&gt;</td>
<td>18%</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>16%†</td>
</tr>
<tr>
<td>Beskind et al., 2014&lt;sup&gt;43&lt;/sup&gt;</td>
<td>2.9%</td>
<td>3.8%</td>
<td>--</td>
<td>--</td>
<td>4.1%‡</td>
</tr>
<tr>
<td>Brown et al., 2014&lt;sup&gt;44&lt;/sup&gt;</td>
<td>4.3%</td>
<td>--</td>
<td>--</td>
<td>39%‡</td>
<td>--</td>
</tr>
<tr>
<td>Caterino, and Raubenolt, 2012&lt;sup&gt;45&lt;/sup&gt;</td>
<td>5.8%</td>
<td>1.5%</td>
<td>15%‡</td>
<td>--</td>
<td>7.6%‡</td>
</tr>
<tr>
<td>Cicero, and Cross, 2013&lt;sup&gt;46&lt;/sup&gt;</td>
<td>3.8%</td>
<td>--</td>
<td>--</td>
<td>21%‡</td>
<td>--</td>
</tr>
<tr>
<td>Corrigan et al., 2014&lt;sup&gt;47&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Davis et al., 2006&lt;sup&gt;48&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Eken et al., 2009&lt;sup&gt;49&lt;/sup&gt;</td>
<td>14%</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Gill et al., 2005&lt;sup&gt;50&lt;/sup&gt;</td>
<td>11%</td>
<td>9.3%</td>
<td>17%‡</td>
<td>--</td>
<td>26%‡</td>
</tr>
<tr>
<td>Gill et al., 2006&lt;sup&gt;51&lt;/sup&gt;</td>
<td>10%</td>
<td>8.8%</td>
<td>17%‡</td>
<td>--</td>
<td>26%‡</td>
</tr>
<tr>
<td>Haukoos et al., 2007&lt;sup&gt;52&lt;/sup&gt;</td>
<td>5.5%</td>
<td>6.6%</td>
<td>14%‡</td>
<td>--</td>
<td>18%‡</td>
</tr>
<tr>
<td>Healey et al., 2003&lt;sup&gt;53&lt;/sup&gt;</td>
<td>6%</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Holmes et al., 2005&lt;sup&gt;54&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
<td>6.3%‡</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Kupas et al., 2016&lt;sup&gt;55&lt;/sup&gt;</td>
<td>5.7%</td>
<td>2.0%*</td>
<td>--</td>
<td>29%**</td>
<td>7.1%‡</td>
</tr>
<tr>
<td>Ross et al., 1998&lt;sup&gt;56&lt;/sup&gt;</td>
<td>6.6%</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>3.5%‡</td>
</tr>
<tr>
<td>Thompson et al., 2011&lt;sup&gt;57&lt;/sup&gt;</td>
<td>5.8%</td>
<td>7.8%</td>
<td>18%‡</td>
<td>--</td>
<td>18%‡</td>
</tr>
<tr>
<td>Van de Voorde et al., 2008&lt;sup&gt;58&lt;/sup&gt;</td>
<td>11%</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

*Craniotomy only
†Intubation in emergency department
‡Intubation in out-of-hospital setting or emergency department
§Studies from National Trauma Data Bank database
∥Injury Severity Score >15, intensive care unit admission ≥24 hours, need for urgent surgery (emergency department disposition to the operating room), or death in the emergency department
¶Skull fracture with underlying brain injury, intracranial hemorrhage, cerebral contusion, or nonspecific intracranial injury
**Injury Severity Score >15
††Traumatic brain injury on computed tomography scan (intracranial hemorrhage, hematoma, contusion, or cerebral edema) or in need of acute intervention (neurosurgical procedure, antiepileptic medication for >7 days, neurologic deficit persisting until discharge, or ≥2 nights of hospitalization for treatment for blunt head injury)
‡‡Intubation in out-of-hospital setting

### In-Hospital Mortality

#### Total Glasgow Coma Scale Versus Motor Glasgow Coma Scale

The tGCS was slightly better than the mGCS at discriminating individuals who experienced in-hospital mortality from those who survived to hospital discharge (Table 3). Based on 12 studies, the pooled AUROC for the tGCS was 0.877 (95% CI 0.847 to 0.906) and for the mGCS
was 0.855 (95% CI 0.822 to 0.888), with a pooled mean difference of 0.015 (95% CI 0.009 to 0.022; I²=85%; Figure 3). Results were unchanged when the analysis was performed using the profile likelihood method. Stratification of studies according to whether they focused on children (2 studies, mean difference in AUROC 0.006, 95% CI 0.002 to 0.011, I²=0%)\textsuperscript{51,55} or evaluated adults or mixed populations (10 studies, mean difference in AUROC 0.019, 95% CI 0.012 to 0.025, I²=75%)\textsuperscript{17,18,20,36,48,52,57-59,61} eliminated statistical heterogeneity, though the pooled mean difference for each subgroup was very similar (Table 4). Results were also similar in subgroup analyses stratified according to use of out-of-hospital (7 studies, mean difference in AUROC 0.016, 95% CI 0.007 to 0.024, I²=91%)\textsuperscript{17,18,36,48,55,57,61} or ED GCS scores (3 studies, mean difference in AUROC 0.020, 95% CI 0.006 to 0.034, I²=23%)\textsuperscript{20,58,59} or years in which data were collected (all data in 2006 or later vs. some or all before 2006), or when analyses were restricted to low risk of bias studies,\textsuperscript{17,18,20,36,59} studies of patients with TBI,\textsuperscript{17,51,57} or studies conducted in the United States (Table 4). Two of the three largest studies (n=202,255 and 101,504; one non-NTDB study had 370,392 patients,\textsuperscript{61} compared with 185 to 21,170 in the other studies) were based on the NTDB database.\textsuperscript{18,55} Estimates from the NTDB studies were very similar (differences in the AUROC 0.02, 95% CI 0.02 to 0.02\textsuperscript{18} and 0.01, 95% CI 0.00 to 0.01\textsuperscript{55}) and excluding the NTDB studies had little effect on estimates (Table 4). When multiple studies were available from a trauma center, restricting the analysis to the most recent study from each trauma center that used out-of-hospital GCS scores (excluding 3 studies\textsuperscript{20,51,59}) also had little effect on estimates.\textsuperscript{20,48}
Figure 3. Pooled AUROC of in-hospital mortality for the total Glasgow Coma Scale versus the motor component only

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC Difference (95% CI)</th>
<th>tGCS (95% CI)</th>
<th>mGCS (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adult or mixed or unclear</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healey et al., 2003</td>
<td>202,255</td>
<td>0.02 (0.01 to 0.02)</td>
<td>0.89 (0.89 to 0.89)</td>
<td>0.87 (0.87 to 0.88)</td>
</tr>
<tr>
<td>Al-Salamah et al., 2004</td>
<td>795</td>
<td>0.01 (-0.02 to 0.04)</td>
<td>0.82 (NR)</td>
<td>0.81 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>6,347</td>
<td>0.01 (-0.02 to 0.04)</td>
<td>0.81 (NR)</td>
<td>0.89 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,160</td>
<td>0.01 (-0.02 to 0.02)</td>
<td>0.89 (0.88 to 0.90)</td>
<td>0.88 (0.87 to 0.89)</td>
</tr>
<tr>
<td>Davis et al., 2006</td>
<td>12,882</td>
<td>0.01 (-0.02 to 0.04)</td>
<td>0.84 (NR)</td>
<td>0.83 (NR)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>0.02 (0.01 to 0.03)</td>
<td>0.92 (0.91 to 0.93)</td>
<td>0.90 (0.89 to 0.91)</td>
</tr>
<tr>
<td>Eken et al., 2009</td>
<td>185</td>
<td>0.07 (-0.01 to 0.16)</td>
<td>0.73 (0.66 to 0.80)</td>
<td>0.66 (0.59 to 0.73)</td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>19,408</td>
<td>0.06 (-0.03 to 0.15)</td>
<td>0.82 (0.74 to 0.90)</td>
<td>0.76 (0.70 to 0.83)</td>
</tr>
<tr>
<td>Beskind et al., 2014</td>
<td>9816</td>
<td>0.01 (-0.02 to 0.04)</td>
<td>0.90 (0.87 to 0.92)</td>
<td>0.89 (0.86 to 0.91)</td>
</tr>
<tr>
<td>Kupas et al., 2016</td>
<td>370,392</td>
<td>0.03 (0.02 to 0.03)</td>
<td>0.83 (0.83 to 0.83)</td>
<td>0.60 (0.80 to 0.81)</td>
</tr>
<tr>
<td><strong>Subtotal (I-squared = 74.6%, p = 0.000)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Children</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cicero and Cross, 2013*</td>
<td>101,504</td>
<td>0.01 (0.00 to 0.01)</td>
<td>0.95 (0.94 to 0.95)</td>
<td>0.94 (0.94 to 0.94)</td>
</tr>
<tr>
<td>Acker et al., 2014</td>
<td>2,231</td>
<td>0.01 (-0.01 to 0.02)</td>
<td>0.95 (0.94 to 0.96)</td>
<td>0.94 (0.93 to 0.96)</td>
</tr>
<tr>
<td><strong>Subtotal (I-squared = 0.0%, p = 0.777)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall (I-squared = 84.7%, p = 0.000)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td>0.01 (0.01 to 0.02)</td>
<td>0.88 (0.85 to 0.91)</td>
<td>0.85 (0.82 to 0.89)</td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td>0.02 (0.01 to 0.02)</td>
<td>0.88 (0.85 to 0.91)</td>
<td>0.85 (0.82 to 0.89)</td>
</tr>
</tbody>
</table>

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; mGCS = motor Glasgow Coma Scale; n = number; NR = not reported; tGCS = total Glasgow Coma Scale

*Patients from the National Trauma Data Bank data set
<table>
<thead>
<tr>
<th>Author, Year Country</th>
<th>Assessment Timing Measures and/or Scores Compared</th>
<th>N</th>
<th>Age</th>
<th>In-Hospital Mortality (95% CI)</th>
<th>Neurosurgical Intervention (95% CI)</th>
<th>Severe Brain Injury (95% CI)</th>
<th>Intubation (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acker et al., 2014 USA</td>
<td>ED A: tGCS B: mGCS</td>
<td>2,231</td>
<td>≤18 years Mean: 6.9 years</td>
<td>A: 0.949 (0.938 to 0.961) B: 0.941 (0.926 to 0.957) p=0.06</td>
<td>A: 0.642 (0.603 to 0.681) B: 0.638 (0.601 to 0.675) p=0.64</td>
<td>A: 0.808 (0.784 to 0.832) B: 0.774 (0.748 to 0.800) p&lt;0.001</td>
<td>NR</td>
</tr>
<tr>
<td>Al-Salamah et al., 2004 Canada</td>
<td>Out-of-hospital A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>795</td>
<td>≥16 years Mean: 44 years</td>
<td>A: 0.82 B: 0.81 p=NR</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Beskind et al., 2014 USA</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>9,816</td>
<td>Mean: 32 years</td>
<td>A: 0.899 (0.874 to 0.923) B: 0.888 (0.864 to 0.913) Mean difference=0.010 (0.002 to 0.018)</td>
<td>A: 0.571 (0.533 to 0.609) B: 0.570 (0.531 to 0.608) Mean difference=0.002 (-0.013 to 0.016)</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012 USA</td>
<td>Out-of-hospital A: tGCS ≤13 B: SMS ≤1</td>
<td>52,412</td>
<td>≥16 years Mean: 53 years</td>
<td>A: 0.85 (0.84 to 0.86) B: 0.82 (0.81 to 0.83)</td>
<td>A: 0.75 (0.73 to 0.77) B: 0.70 (0.68 to 0.72)</td>
<td>A: 0.72 (0.71 to 0.72) B: 0.66 (0.65 to 0.66)</td>
<td>A: 0.86 (0.85 to 0.87) B: 0.83 (0.82 to 0.83)</td>
</tr>
<tr>
<td>Cicero and Cross, 2013 USA</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>104,035</td>
<td>&lt;19 years Mean: 12.6 years</td>
<td>A: 0.946 (0.941 to 0.951) B: 0.940 (0.935 to 0.945)</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Corrigan et al., 2014 USA</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>77,470</td>
<td>≥18 years Mean: 12.6 years</td>
<td>A: 0.886 (NR) B: 0.878 (NR)</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Davis et al., 2006 USA</td>
<td>Out-of-hospital and ED A: tGCS B: mGCS</td>
<td>12,882</td>
<td>NR</td>
<td>A: 0.84 (NR) B: 0.83 (NR)</td>
<td>A: 0.80 (NR) B: 0.78 (NR)</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Eken et al., 2009 Turkey</td>
<td>ED A: tGCS B: mGCS</td>
<td>185</td>
<td>&gt;17 years Mean: 6.9 years</td>
<td>A: 0.735 (0.655 to 0.797) B: 0.662 (0.589 to 0.730)</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Assessment Timing Measures and/or Scores Compared</td>
<td>N</td>
<td>Age</td>
<td>In-Hospital Mortality (95% CI)</td>
<td>Neurosurgical Intervention (95% CI)</td>
<td>Severe Brain Injury (95% CI)</td>
<td>Intubation (95% CI)</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------</td>
<td>-------</td>
<td>--------------</td>
<td>-------------------------------</td>
<td>-----------------------------------</td>
<td>-------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Gill et al., 2005&lt;sup&gt;20&lt;/sup&gt; USA</td>
<td>ED A: tGCS B: mGCS C: SMS</td>
<td>8,412</td>
<td>Median of 24 years</td>
<td>A: 0.906 (NR) B: 0.894 (NR) C: 0.878 (NR)</td>
<td>A: 0.874 (NR) B: 0.848 (NR) C: 0.851 (NR)</td>
<td>A: 0.826 (NR) B: 0.789 (NR) C: 0.791 (NR)</td>
<td>A: 0.865 (NR) B: 0.826 (NR) C: 0.826 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006&lt;sup&gt;26&lt;/sup&gt; USA</td>
<td>Out-of-hospital A: tGCS B: mGCS C: SMS</td>
<td>7,233</td>
<td>Median of 24 years</td>
<td>A: 0.89 (0.88 to 0.90) B: 0.88 (0.87 to 0.89) C: 0.86 (0.86 to 0.89)</td>
<td>A: 0.86 (0.85 to 0.88) B: 0.84 (0.82 to 0.85) C: 0.83 (0.81 to 0.84)</td>
<td>A: 0.83 (0.82 to 0.84) B: 0.79 (0.78 to 0.81) C: 0.79 (0.77 to 0.80)</td>
<td>A: 0.83 (0.81 to 0.84) B: 0.79 (0.78 to 0.80) C: 0.79 (0.77 to 0.80)</td>
</tr>
<tr>
<td>Haukoos et al., 2007&lt;sup&gt;1&lt;/sup&gt; USA</td>
<td>ED A: tGCS B: mGCS C: SMS</td>
<td>21,170</td>
<td>Median of 32 years</td>
<td>A: 0.92 (0.91 to 0.93) B: 0.90 (0.89 to 0.91) C: 0.89 (0.88 to 0.90)</td>
<td>A: 0.83 (0.82 to 0.84) B: 0.80 (0.79 to 0.81) C: 0.80 (0.79 to 0.81)</td>
<td>A: 0.76 (0.75 to 0.77) B: 0.71 (0.70 to 0.72) C: 0.71 (0.70 to 0.72)</td>
<td>A: 0.86 (0.85 to 0.87) B: 0.81 (0.80 to 0.82) C: 0.81 (0.80 to 0.82)</td>
</tr>
<tr>
<td>Healey et al., 2003&lt;sup&gt;18&lt;/sup&gt; USA</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>202,255</td>
<td>NR</td>
<td>A: 0.891 (0.888 to 0.894) B: 0.873 (0.870 to 0.875)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Holmes et al., 2005&lt;sup&gt;40&lt;/sup&gt; USA</td>
<td>ED A: tGCS B: mGCS</td>
<td>2,043</td>
<td>≤2 years</td>
<td>NR</td>
<td>NR</td>
<td>Ages ≤2 years: A: 0.72 (0.56 to 0.87) B: 0.60 (0.48 to 0.72) Ages &gt;2 years: A: 0.82 (0.76 to 0.87) B: 0.71 (0.65 to 0.77) AUROC (95% CI) for TBI in need of acute intervention Ages ≤2 years: A: 0.97 (0.94 to 1.0) B: 0.76 (0.59 to 0.93) Ages &gt;2 years: A: 0.87 (0.83 to 0.92) B: 0.76 (0.71 to 0.81)</td>
<td>NR</td>
</tr>
<tr>
<td>Kupas et al., 2016&lt;sup&gt;61&lt;/sup&gt; USA</td>
<td>Out-of-hospital A: tGCS B: mGCS</td>
<td>370,392</td>
<td>≥18 years</td>
<td>A: 0.831 (0.828 to 0.834) B: 0.803 (0.800 to 0.806) Difference=0.028 (0.026 to 0.030)</td>
<td>A: 0.724 (0.718 to 0.730) B: 0.676 (0.670 to 0.682) Difference=0.048 (0.044 to 0.052)*</td>
<td>A: 0.904 (0.902 to 0.907) B: 0.884 (0.882 to 0.887) Difference=0.020 (0.019 to 0.021)</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year Country</td>
<td>Assessment Timing Measures and/or Scores Compared</td>
<td>N</td>
<td>Age</td>
<td>In-Hospital Mortality (95% CI)</td>
<td>Neurosurgical Intervention (95% CI)</td>
<td>Severe Brain Injury (95% CI)</td>
<td>Intubation (95% CI)</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------------------------------</td>
<td>--------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>-----------------------------------</td>
<td>-------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Thompson et al., 2011 USA</td>
<td>Out-of-hospital a: tGCS ≤13 b: mGCS score ≤5 b: SMS ≤1</td>
<td>19,408</td>
<td>Median: 33 years</td>
<td>A: 0.82 (0.74 to 0.90) b: 0.76 (0.70 to 0.83) c: 0.74 (0.70 to 0.77)</td>
<td>A: 0.70 (0.64 to 0.77) b: 0.66 (0.61 to 0.71) c: 0.66 (0.64 to 0.69)</td>
<td>A: 0.66 (0.60 to 0.71) b: 0.61 (0.57 to 0.65) c: 0.61 (0.58 to 0.64)</td>
<td>A: 0.70 (0.63 to 0.77) b: 0.65 (0.60 to 0.70) c: 0.65 (0.62 to 0.67)</td>
</tr>
</tbody>
</table>

AUROC=area under the receiver operating characteristic curve; CI=confidence interval; ED=emergency department; mGCS=motor Glasgow Coma Scale; n=number; NR=not reported; SMS=Simplified Motor Scale; TBI=traumatic brain injury; tGCS=total Glasgow Coma Scale; vs.=versus

*Craniotomy only
†Intracranial pressure monitoring only
‡Studies from NTDB database
### Table 4. Pooled AUROC results of head-to-head studies

<table>
<thead>
<tr>
<th>Outcome and Analysis</th>
<th>tGCS vs. mGCS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>( \chi^2 )</th>
<th>tGCS vs. SMS, Difference in AUROC (95% CI)</th>
<th>Number of Studies</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In-hospital mortality, overall</strong></td>
<td>0.015 (0.009 to 0.022)</td>
<td>12</td>
<td>85%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Adults or mixed</td>
<td>0.019 (0.012 to 0.025)</td>
<td>10</td>
<td>75%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Children</td>
<td>0.006 (0.002 to 0.011)</td>
<td>2</td>
<td>0%</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Excluding NTDB studies</td>
<td>0.017 (0.008 to 0.025)</td>
<td>10</td>
<td>68%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Excluding studies with potential overlap*</td>
<td>0.016 (0.008 to 0.024)</td>
<td>9</td>
<td>88%</td>
<td>0.031 (0.023 to 0.039)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Out-of-hospital GCS</strong></td>
<td>0.016 (0.007 to 0.024)</td>
<td>7</td>
<td>91%</td>
<td>0.031 (0.023 to 0.039)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td>ED GCS</td>
<td>0.020 (0.006 to 0.034)</td>
<td>3</td>
<td>23%</td>
<td>0.030 (0.020 to 0.039)</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>U.S. setting</td>
<td>0.015 (0.008 to 0.022)</td>
<td>10</td>
<td>87%</td>
<td>0.030 (0.024 to 0.036)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>TBI patients</td>
<td>0.009 (-0.002 to 0.020)</td>
<td>3</td>
<td>0%</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Low risk of bias studies</td>
<td>0.017 (0.015 to 0.020)</td>
<td>5</td>
<td>0%</td>
<td>0.030 (0.022 to 0.037)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Neurosurgical intervention, overall</strong></td>
<td>0.032 (0.020 to 0.043)</td>
<td>7</td>
<td>72%</td>
<td>0.032 (0.025 to 0.039)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Adults or mixed</td>
<td>0.031 (0.018 to 0.044)</td>
<td>6</td>
<td>76%</td>
<td>0.032 (0.025 to 0.039)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>Children</td>
<td>0.034 (0.009 to 0.059)</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Excluding studies with potential overlap*</td>
<td>0.032 (0.011 to 0.053)</td>
<td>4</td>
<td>79%</td>
<td>0.038 (0.024 to 0.052)</td>
<td>3</td>
<td>19%</td>
</tr>
<tr>
<td><strong>Out-of-hospital GCS</strong></td>
<td>0.032 (0.011 to 0.053)</td>
<td>4</td>
<td>79%</td>
<td>0.038 (0.024 to 0.052)</td>
<td>3</td>
<td>19%</td>
</tr>
<tr>
<td>ED GCS</td>
<td>0.029 (0.020 to 0.039)</td>
<td>2</td>
<td>0%</td>
<td>0.029 (0.020 to 0.038)</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>U.S. setting</td>
<td>0.032 (0.020 to 0.044)</td>
<td>7</td>
<td>72%</td>
<td>0.032 (0.025 to 0.039)</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>TBI patients</td>
<td>0.017 (-0.022 to 0.056)</td>
<td>2</td>
<td>66%</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Low risk of bias studies</td>
<td>0.026 (0.019 to 0.034)</td>
<td>4</td>
<td>0%</td>
<td>0.029 (0.021 to 0.037)</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Severe brain injury, overall</strong></td>
<td>0.050 (0.034 to 0.065)</td>
<td>5</td>
<td>57%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>Adults or mixed</td>
<td>0.046 (0.038 to 0.054)</td>
<td>4</td>
<td>0%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>Children</td>
<td>0.121 (0.068 to 0.174)</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Excluding NTDB studies</td>
<td>0.050 (0.034 to 0.065)</td>
<td>5</td>
<td>57%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>Excluding studies with potential overlap*</td>
<td>0.065 (0.020 to 0.111)</td>
<td>3</td>
<td>76%</td>
<td>0.051 (0.034 to 0.068)</td>
<td>3</td>
<td>74%</td>
</tr>
<tr>
<td><strong>Out-of-hospital GCS</strong></td>
<td>0.041 (0.028 to 0.053)</td>
<td>2</td>
<td>0%</td>
<td>0.051 (0.034 to 0.068)</td>
<td>3</td>
<td>74%</td>
</tr>
<tr>
<td>ED GCS</td>
<td>0.060 (0.028 to 0.093)</td>
<td>3</td>
<td>73%</td>
<td>0.044 (0.030 to 0.059)</td>
<td>2</td>
<td>51%</td>
</tr>
<tr>
<td>U.S. setting</td>
<td>0.050 (0.034 to 0.065)</td>
<td>5</td>
<td>57%</td>
<td>0.048 (0.038 to 0.059)</td>
<td>5</td>
<td>72%</td>
</tr>
<tr>
<td>TBI patients</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
Data on the diagnostic accuracy of the tGCS versus the mGCS were limited (Table 5). In five studies, sensitivity ranged widely from 30 percent to 100 percent for the tGCS (cutoff of ≤13) and from 27 percent to 100 percent for the mGCS (cutoff of ≤5); however, differences in sensitivity were small, ranging from 0 percent to 3 percent (Table 5). Specificity ranged from 68 percent to 93 percent for the tGCS and from 73 percent to 95 percent for the mGCS; difference in specificity ranged from 0 percent to 5 percent in favor of the mGCS, though the CIs overlapped in most studies. Two of the studies were conducted in mixed populations of adults and children and one of the studies focused on children (sensitivity 100%, specificity 74%). The latter study found that the specificity of the tGCS decreased at higher cutoffs (74% for a cutoff of ≤13, 71% for ≤14 and 56% for ≤15) with little change in sensitivity; the other studies did not report specificity at tGCS cutoffs other than 13 or less. One study found that calibration of the tGCS and mGCS was similarly poor based on the Hosmer-Lemeshow test (p<0.01 for both scales).
Table 5. Summary of diagnostic accuracy outcomes for head-to-head studies

<table>
<thead>
<tr>
<th>Author, Year Country</th>
<th>Assessment Timing Measures and/or Scores Compared</th>
<th>N</th>
<th>Age</th>
<th>In-Hospital Mortality</th>
<th>Neurosurgical Intervention</th>
<th>Severe Brain Injury</th>
<th>Intubation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-Salamah et al., 2004 Canada</td>
<td>Out-of-hospital A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>795</td>
<td>≥16 years Mean: 44 years</td>
<td>Sensitivity (95% CI)<em>: 80.28% (72.78 to 86.48) vs. 80.28% (72.78 to 86.48) Specificity (95% CI)</em>: 67.99% (64.26 to 71.56) vs. 73.05% (69.47 to 76.42) PPV (95% CI)<em>: 35.29% (30.08 to 40.78) vs. 39.31% (33.65 to 45.19) NPV (95% CI)</em>: 94.07% (91.54 to 96.02) vs. 94.46% (92.09 to 96.28)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Brown et al., 2014 USA</td>
<td>Out-of-hospital A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>811,143</td>
<td>≥3 years Median: 39 years</td>
<td>Sensitivity: 30.3% vs. 26.7% Specificity: 93.1% vs. 95.1%</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Assessment and/or Scores Compared</td>
<td>N</td>
<td>Age</td>
<td>In-Hospital Mortality</td>
<td>Neurosurgical Intervention</td>
<td>Severe Brain Injury</td>
<td>Intubation</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------</td>
<td>------</td>
<td>-----</td>
<td>-----------------------</td>
<td>---------------------------</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>Out-of-hospital A: tGCS ≤13 B: SMS ≤1</td>
<td>52,412</td>
<td>≥16 years Mean: 53 years</td>
<td>Sensitivity (95% CI)<em>: 75.03% (73.45 to 76.56) vs. 72.20% (70.57 to 73.79) Specificity (95% CI)</em>: 87.63% (87.34 to 87.92) vs. 89.42% (89.14 to 89.69) PPV (95% CI)<em>: 27.20% (26.25 to 28.17) vs. 29.59% (28.55 to 30.64) NPV (95% CI)</em>: 98.28% (98.15 to 98.40) vs. 98.12% (97.99 to 98.25)</td>
<td>Sensitivity (95% CI)<em>: 60.05% (56.53 to 63.50) vs. 52.93% (49.37 to 56.46) Specificity (95% CI)</em>: 84.70% (84.39 to 85.01) vs. 86.40% (86.10 to 86.69) PPV (95% CI)<em>: 5.64% (5.15 to 6.15) vs. 5.59% (5.08 to 6.14) NPV (95% CI)</em>: 99.29% (99.20 to 99.36) vs. 99.18% (99.09 to 99.26)</td>
<td>Sensitivity (95% CI)<em>: 45.40% (44.30 to 46.50) vs. 40.81% (39.72 to 41.89) Specificity (95% CI)</em>: 89.30% (89.01 to 89.59) vs. 90.50% (90.22 to 90.77) PPV (95% CI)<em>: 43.20% (42.13 to 44.27) vs. 43.50% (42.38 to 44.64) NPV (95% CI)</em>: 90.12% (89.84 to 90.40) vs. 89.51% (89.22 to 89.79)</td>
<td>Any emergency intubation Sensitivity (95% CI)<em>: 75.50% (74.13 to 76.83) vs. 72.71% (71.30 to 74.09) Specificity (95% CI)</em>: 88.90% (88.62 to 89.18) vs. 90.60% (90.34 to 90.86) PPV (95% CI)<em>: 35.87% (34.84 to 36.91) vs. 38.88% (37.77 to 40.00) NPV (95% CI)</em>: 97.78% (97.64 to 97.92) vs. 97.58% (97.44 to 97.72) ED intubation Sensitivity (95% CI)<em>: 76.89% (75.43 to 78.31) vs. 74.09% (72.57 to 75.57) Specificity (95% CI)</em>: 88.20% (87.91 to 88.48) vs. 89.83% (89.56 to 90.09) PPV (95% CI)<em>: 30.82% (29.83 to 31.82) vs. 33.22% (32.15 to 34.30) NPV (95% CI)</em>: 98.24% (98.11 to 98.36) vs. 98.07% (97.94 to 98.19)</td>
</tr>
<tr>
<td>Kupas et al., 2016</td>
<td>Out-of-hospital A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>370,392</td>
<td>&gt;18 years Mean: NR</td>
<td>Sensitivity (95% CI): 69.8% (69.2 to 70.4) vs. 67.3% (66.7 to 67.9) Specificity (95% CI): 88.1% (88.0 to 88.2) vs. 90.1% (90.0 to 90.2)</td>
<td>Sensitivity (95% CI): 51.4% (50.2 to 52.5) vs. 46.5% (45.4 to 47.7) Specificity (95% CI): 85.9% (85.8 to 86.0) vs. 87.8% (87.7 to 87.9)</td>
<td>NR</td>
<td>Sensitivity (95% CI): 83.7% (83.3 to 84.2) vs. 81.3% (80.9 to 81.8) Specificity (95% CI): 90.0% (89.9 to 90.1) vs. 92.0% (91.9 to 92.1)</td>
</tr>
<tr>
<td>Author, Year, Country</td>
<td>Assessment and/or Scores Compared</td>
<td>N</td>
<td>Age</td>
<td>In-Hospital Mortality</td>
<td>Neurosurgical Intervention</td>
<td>Severe Brain Injury</td>
<td>Intubation</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------------------</td>
<td>---</td>
<td>-----</td>
<td>-----------------------</td>
<td>---------------------------</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Ross et al., 1998(^{19}) USA</td>
<td>Out-of-hospital A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>1,410</td>
<td>≥13 years Mean: 37 years</td>
<td>Sensitivity (95% CI)<em>: 71.28% (61.02 to 80.14) vs. 72.34% (62.15 to 81.07) Specificity (95% CI)</em>: 84.95% (82.91 to 86.84) vs. 86.02% (84.03 to 87.85) PPV (95% CI)<em>: 25.28% (20.16 to 30.96) vs. 26.98% (21.61 to 32.91) NPV (95% CI)</em>: 97.64% (96.59 to 98.44) vs. 97.75% (96.73 to 98.53)</td>
<td>Sensitivity (95% CI)<em>: 63.16% (38.36 to 83.71) vs. 68.42% (43.45 to 87.42) Specificity (95% CI)</em>: 81.81% (79.68 to 83.81) vs. 82.82% (80.73 to 84.77) PPV (95% CI)<em>: 4.53% (2.36 to 7.78) vs. 5.16% (2.78 to 8.66) NPV (95% CI)</em>: 99.39% (98.74 to 99.75) vs. 99.48% (98.88 to 99.81)</td>
<td>Sensitivity: 61.72% (54.76 to 68.34) vs. 60.77% (53.79 to 67.43) Specificity: 85.47% (83.05 to 87.67) vs. 89.59% (87.73 to 91.26) PLR: 4.25 (3.52 to 5.13) vs. 5.84 (4.79 to 7.12) NLR: 0.45 (0.38 to 0.53) vs. 0.44 (0.37 to 0.52) PPV: 48.68% (42.52 to 54.87) vs. 50.40% (44.05 to 56.73) NPV: 90.91% (88.81 to 92.73) vs. 92.92% (91.29 to 94.33)</td>
<td>NR</td>
</tr>
<tr>
<td>Van de Voorde et al., 2008(^{62}) Belgium</td>
<td>Out-of-hospital and ED A: tGCS score ≤13 B: mGCS score ≤5</td>
<td>96</td>
<td>≤18 years Mean: 8 years</td>
<td>Sensitivity (95% CI): 100% (69.15 to 100) vs. 100% (69.15 to 100) Specificity (95% CI): 74.39% (63.56 to 83.40) vs. 74.36% (63.21 to 83.58)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

CI=confidence interval; ED=emergency department; mGCS=motor Glasgow Coma Scale; n=number; NPV=negative predictive value; NR=not reported; PPV=positive predictive values; SMS=Simplified Motor Scale; tGCS=total Glasgow Coma Scale; vs.=versus

*Calculated
†Study from NTDB database
**Total Glasgow Coma Scale Versus Simplified Motor Scale**

The tGCS was slightly better than the SMS at discriminating patients who experienced in-hospital mortality from patients who survived to hospital discharge. Based on five studies, the pooled AUROC for the tGCS was 0.884 (95% CI 0.852 to 0.916) and for the SMS was 0.840 (95% CI 0.802 to 0.878), with a mean difference of 0.030 (95% CI 0.024 to 0.036, I²=0%; Figure 4). All of the studies were conducted in the United States and none focused on children or patients with TBI. There were no differences in estimates between studies that utilized out-of-hospital GCS scores or ED GCS scores (Table 4). Results were also unchanged when analyses were restricted to low risk of bias studies, or when studies from the same center with potential overlap were excluded. No study was based on the NTDB database.

One study found the out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity versus the SMS (cutoff of ≤1) (75% vs. 72%) and slightly lower specificity (88% vs. 89%; Table 5).
Figure 4. Pooled AUROC of in-hospital mortality for the total Glasgow Coma Scale versus the Simplified Motor Scale

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC Difference (95% CI)</th>
<th>GCS (95% CI)</th>
<th>SMS (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult or mixed or unclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>8,347</td>
<td>0.03 (0.00 to 0.05)</td>
<td>0.91 (NR)</td>
<td>0.88 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,160</td>
<td>0.03 (0.02 to 0.05)</td>
<td>0.89 (0.88 to 0.90)</td>
<td>0.86 (0.86 to 0.89)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.92 (0.91 to 0.93)</td>
<td>0.89 (0.88 to 0.90)</td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>19,408</td>
<td>0.06 (0.01 to 0.15)</td>
<td>0.82 (0.74 to 0.90)</td>
<td>0.74 (0.70 to 0.80)</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>2,043</td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.85 (0.84 to 0.86)</td>
<td>0.82 (0.81 to 0.83)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 0.0%, p = 0.734)</td>
<td></td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.86 (0.85 to 0.92)</td>
<td>0.84 (0.80 to 0.90)</td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AUROC=area under the receiver operating characteristic curve; CI=confidence interval; n=number; NR=not reported; SMS=Simplified Motor Scale; tGCS=total Glasgow Coma Scale
Motor Glasgow Coma Scale Versus Simplified Motor Scale

The mGCS was slightly better than the SMS at discriminating patients who experienced in-hospital mortality from patients who survived to hospital discharge (4 studies, mean difference in AUROC 0.014, 95% CI 0.006 to 0.021, I²=0%).²⁰,³⁶,⁴⁸,⁵⁹ There was no statistical heterogeneity and findings were unchanged in sensitivity and subgroup analyses.

Neurosurgical Intervention

Total Glasgow Coma Scale Versus Motor Glasgow Coma Scale

The tGCS was slightly better than the mGCS at discriminating patients who went on to receive a neurosurgical intervention from those who did not. Based on seven studies, the pooled AUROC for the tGCS was 0.786 (95% CI 0.729 to 0.842) and for the mGCS was 0.754 (95% CI 0.688 to 0.819) for a mean difference of 0.032 (95% CI 0.020 to 0.043; I²=72%; Figure 5).¹⁷,²⁰,³⁶,⁴⁸,⁵¹,⁵⁹,⁶¹ Results were similar when the analysis was performed using the profile likelihood method. Results were similar in one study⁵¹ that restricted enrollment to children (mean difference in AUROC 0.034, 95% CI 0.009 to 0.059) and the other six studies, which evaluated mixed populations of adults and children (mean difference in AUROC 0.031, 95% CI 0.018 to 0.044, I²=76%). Two studies that reported results for craniotomy only⁵¹,⁶¹ reported results similar to studies that evaluated craniotomy plus other neurosurgical interventions. Results were also similar when studies were stratified according to whether they used out-of-hospital or ED GCS scores, or when analyses were restricted to studies conducted in the United States, studies that focused on TBI patients, or low risk of bias studies (Table 4). No study was based on the NTDB database.

Two studies found inconsistent results for sensitivity (51%, 95% CI 50 to 52 vs. 46%, 95% CI 45 to 48 and 63%, 95% CI 38 to 84 vs. 68%, 95% CI 43 to 87) and no clear differences in specificity (86%, 95% CI, 86 to 86 vs. 88%, 95% CI 88 to 88 and 82%, 95% CI 80 to 84 vs. 83%, 95% CI 81 to 85) between out-of-hospital tGCS (cutoff of ≤13) versus the mGCS (cutoff of ≤5) in accuracy for identifying patients undergoing craniotomy (Table 5).¹⁹,⁶¹
Figure 5. Pooled AUROC of neurological intervention for the total Glasgow Coma Scale versus the motor component only

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC</th>
<th>Difference (95% CI)</th>
<th>tGCS (95% CI)</th>
<th>mGCS (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult or mixed or unclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>8,432</td>
<td>0.03</td>
<td>(0.00 to 0.05)</td>
<td>0.87 (NR)</td>
<td>0.85 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,233</td>
<td>0.02</td>
<td>(0.01 to 0.03)</td>
<td>0.86 (0.85 to 0.88)</td>
<td>0.84 (0.82 to 0.85)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>0.03</td>
<td>(0.02 to 0.04)</td>
<td>0.83 (0.82 to 0.84)</td>
<td>0.80 (0.79 to 0.81)</td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>19,408</td>
<td>0.04</td>
<td>(-0.02 to 0.10)</td>
<td>0.70 (0.64 to 0.77)</td>
<td>0.66 (0.61 to 0.71)</td>
</tr>
<tr>
<td>Beskind et al., 2014</td>
<td>9,816</td>
<td>0.02</td>
<td>(-0.01 to 0.05)</td>
<td>0.69 (0.66 to 0.72)</td>
<td>0.67 (0.64 to 0.70)</td>
</tr>
<tr>
<td>Kupas et al., 2016</td>
<td>361,022</td>
<td>0.05</td>
<td>(0.04 to 0.05)</td>
<td>0.72 (0.72 to 0.73)</td>
<td>0.68 (0.67 to 0.68)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 76.1%, p = 0.001)</td>
<td></td>
<td>0.03</td>
<td>(0.02 to 0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acker et al., 2014*</td>
<td>2,043</td>
<td>0.03</td>
<td>(0.01 to 0.06)</td>
<td>0.81 (0.78 to 0.83)</td>
<td>0.77 (0.75 to 0.80)</td>
</tr>
<tr>
<td>Subtotal (I-squared = NA, p = NA)</td>
<td></td>
<td>0.03</td>
<td>(0.01 to 0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall (I-squared = 71.6%, p = 0.002)</td>
<td></td>
<td>0.03</td>
<td>(0.02 to 0.04)</td>
<td>0.79 (0.73 to 0.84)</td>
<td>0.75 (0.69 to 0.82)</td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td>0.03</td>
<td>(0.02 to 0.04)</td>
<td>0.79 (0.73 to 0.84)</td>
<td>0.75 (0.69 to 0.82)</td>
</tr>
</tbody>
</table>

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; mGCS = motor Glasgow Coma Scale; n = number; NA = not applicable; NR = not reported; tGCS = total Glasgow Coma Scale

*Intracranial pressure monitoring only
The tGCS was slightly better than the SMS at discriminating patients who underwent a neurosurgical intervention from patients who did not undergo a neurosurgical intervention; results were very similar to the comparison of the tGCS with the mGCS. Based on five studies, the pooled AUROC for the tGCS was 0.809 (95% CI 0.766 to 0.853) and for the mGCS was 0.769 (95% CI 0.711 to 0.827), with a mean difference of 0.032 (95% CI 0.025 to 0.039, $I^2=0%$; Figure 6). All of the studies were conducted in the United States and none focused on children or patients with TBI. There were no differences in estimates between studies that utilized out-of-hospital GCS scores or ED GCS scores (Table 4). Results were also unchanged when analyses were restricted to low risk of bias studies, or when studies with potential overlap were excluded. No study utilized data from the NTDB.
Figure 6. Pooled AUROC of neurological intervention for the total Glasgow Coma Scale versus the Simplified Motor Scale

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC Difference (95% CI)</th>
<th>tGCS (95% CI)</th>
<th>SMS (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adult or mixed or unclear</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>8,432</td>
<td>0.02 (-0.00 to 0.05)</td>
<td>0.87 (NR)</td>
<td>0.85 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,233</td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.86 (0.85 to 0.88)</td>
<td>0.83 (0.81 to 0.84)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.83 (0.82 to 0.84)</td>
<td>0.80 (0.79 to 0.81)</td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>19,408</td>
<td>0.04 (-0.02 to 0.10)</td>
<td>0.70 (0.64 to 0.77)</td>
<td>0.66 (0.64 to 0.69)</td>
</tr>
<tr>
<td>Caterino and Raubenoth, 2012</td>
<td>9,816</td>
<td>0.05 (0.03 to 0.07)</td>
<td>0.75 (0.73 to 0.77)</td>
<td>0.70 (0.68 to 0.72)</td>
</tr>
<tr>
<td><strong>Subtotal (I-squared = 0.0%, p = 0.414)</strong></td>
<td></td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.81 (0.77 to 0.85)</td>
<td>0.77 (0.71 to 0.83)</td>
</tr>
<tr>
<td><strong>Profile likelihood method</strong></td>
<td></td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.81 (0.74 to 0.87)</td>
<td>0.77 (0.69 to 0.85)</td>
</tr>
</tbody>
</table>

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; n = number; NR = not reported; SMS = Simplified Motor Scale; tGCS = total Glasgow Coma Scale
One study found the out-of-hospital tGCS (cutoff of ≤13) associated with higher sensitivity than the SMS (cutoff of ≤1) for identifying patients who underwent neurosurgical intervention (60%, 95% CI 56 to 63 vs. 53%, 95% CI 49 to 56) and slightly lower specificity (85%, 95% CI 84 to 85 vs. 86%, 95% CI 86 to 87; Table 5).54

**Motor Glasgow Coma Scale Versus Simplified Motor Scale**

There was no difference between the mGCS and the SMS in ability to discriminate patients who would undergo a neurosurgical intervention from those who would not undergo a neurosurgical intervention (4 studies, mean difference in AUROC 0.002, 95% CI -0.005 to 0.010, I²=0%).20,36,48,59 There was no statistical heterogeneity and findings were unchanged in sensitivity and stratified analyses.

**Severe Brain Injury**

**Total Glasgow Coma Scale Versus Motor Glasgow Coma Scale**

The tGCS was slightly better than the mGCS at discriminating patients found to have a severe brain injury from those without severe brain injury. Based on five studies, the pooled AUROC for the tGCS was 0.791 (95% CI 0.734 to 0.827) and for the mGCS was 0.720 (95% CI 0.666 to 0.774), with a mean difference of 0.050 (95% CI 0.034 to 0.065; I²=57%; Figure 7).20,36,48,55,59,60 Results were similar when the analysis was performed using the profile likelihood method. The mean difference in AUROC was slightly higher in one study of children (0.121, 95% CI 0.068 to 0.174)55,60 than in four studies of mixed populations of adults and children (0.046, 95% CI 0.038 to 0.054, I²=0%),20,36,48,59 but there was no statistically significant interaction with age group (p=0.07). Differences in how severe brain injury was defined could explain some of the differences in estimates. The study in children used a composite outcome of head CT imaging findings or need for intervention.60 All of the studies of mixed populations of adults and children defined severe brain injury on the basis of CT imaging findings.
Figure 7. Pooled AUROC of severe brain injury for the total Glasgow Coma Scale versus the motor component only

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Difference (95% CI)</td>
</tr>
<tr>
<td>Adult or mixed or unclear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>84,332</td>
<td>0.04 (0.01 to 0.07)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,233</td>
<td>0.04 (0.03 to 0.05)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>0.05 (0.04 to 0.06)</td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>12,885</td>
<td>0.05 (0.00 to 0.10)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 0.0%, p = 0.624)</td>
<td></td>
<td>0.05 (0.04 to 0.05)</td>
</tr>
<tr>
<td>Children</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holmes et al., 2005</td>
<td>2,043</td>
<td>0.12 (0.07 to 0.17)</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>0.12 (0.07 to 0.17)</td>
</tr>
<tr>
<td>Overall (I-squared = 57.0%, p = 0.054)</td>
<td></td>
<td>0.05 (0.03 to 0.06)</td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td>0.05 (0.03 to 0.05)</td>
</tr>
</tbody>
</table>

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; mGCS = motor Glasgow Coma Scale; n = number; NR = not reported; tGCS = total Glasgow Coma Scale
Results were similar when studies were stratified according to whether they used out-of-hospital or ED GCS scores, or when analyses were restricted to studies conducted in the United States, studies that focused on TBI patients, or low risk of bias studies (Table 4). There were also no differences when we excluded studies with potential overlap.\textsuperscript{20,59} No study was based on data from the NTDB.

One study found no difference between out-of-hospital tGCS (cutoff of $\leq 13$) versus the mGCS (cutoff of $\leq 5$) in sensitivity (62\%, 95\% CI 55 to 68 vs. 61\%, 95\% CI 54 to 67) or specificity (85\%, 95\% CI 83 to 88 vs. 89\%, 95\% CI 88 to 91) for identifying patients with severe head injury (defined as head AIS score of $\geq 4$)\textsuperscript{19} (Table 5).

**Total Glasgow Coma Scale Versus Simplified Motor Scale**

The tGCS was slightly better than the SMS at discriminating patients found to have severe brain injury from those without a severe brain injury; results were very similar to the comparison of tGCS versus mGCS. Based on five studies, the pooled AUROC for the tGCS was 0.763 (95\% CI 0.710 to 0.815) and for the mGCS was 0.713 (95\% CI 0.654 to 0.771), with a mean difference of 0.048 (95\% CI 0.038 to 0.059, $I^2=72\%$; Figure 8).\textsuperscript{20,36,48,54,59} Although statistical heterogeneity was present, the estimates from individual studies were similar (mean difference in AUROC ranged from 0.035 to 0.060), and all studies favored the tGCS. All of the studies defined severe brain injury similarly, based on head CT imaging findings. All of the studies were conducted in the United States and none focused on children or patients with TBI. There were no differences in estimates between studies that utilized out-of-hospital GCS scores or ED GCS scores (Table 4). Results were also unchanged when analyses were restricted to low risk of bias studies, or when studies with potential overlap\textsuperscript{20,59} were excluded. No study utilized the NTDB database. Results were unchanged using the profile likelihood method.
Figure 8. Pooled AUROC of severe brain injury for the total Glasgow Coma Scale versus the Simplified Motor Scale

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC Difference (95% CI)</th>
<th>tGCS (95% CI)</th>
<th>SMS (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult or mixed or unclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>84,332</td>
<td>0.04 (0.02 to 0.05)</td>
<td>0.83 (NR)</td>
<td>0.79 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,233</td>
<td>0.04 (0.03 to 0.05)</td>
<td>0.83 (0.82 to 0.84)</td>
<td>0.79 (0.77 to 0.80)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>0.05 (0.04 to 0.06)</td>
<td>0.76 (0.75 to 0.77)</td>
<td>0.71 (0.70 to 0.72)</td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>12,885</td>
<td>0.05 (0.00 to 0.10)</td>
<td>0.66 (0.60 to 0.7)</td>
<td>0.61 (0.58 to 0.64)</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>2,043</td>
<td>0.06 (0.06 to 0.06)</td>
<td>0.72 (0.71 to 0.72)</td>
<td>0.66 (0.65 to 0.66)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 72.4%, p = 0.006)</td>
<td></td>
<td>0.05 (0.04 to 0.06)</td>
<td>0.76 (0.71 to 0.81)</td>
<td>0.71 (0.65 to 0.77)</td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.76 (0.69 to 0.83)</td>
<td>0.71 (0.64 to 0.79)</td>
</tr>
</tbody>
</table>

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; n = number; NR = not reported; SMS = Simplified Motor Scale; tGCS = total Glasgow Coma Scale
One study found out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity than the SMS (cutoff of ≤1) for severe brain injury (45%, 95% CI 44 to 46 vs. 41%, 95% CI 40 to 42), based on presence of head CT imaging findings, and similar specificity (89%, 95% CI 89 to 90 vs. 90%, 95% CI 90 to 91; Table 5).54

**Motor Glasgow Coma Scale Versus Simplified Motor Scale**

There was no difference between the mGCS versus the SMS in ability to discriminate patients who would undergo a neurosurgical intervention from those who would not undergo a neurosurgical intervention (4 studies, mean difference in AUROC 0.000, 95% CI -0.008 to 0.007, I²=0%).20,36,48,59 There was no statistical heterogeneity and findings were unchanged in sensitivity and stratified analyses.

**Emergency Intubation**

**Total Glasgow Coma Scale Versus Motor Glasgow Coma Scale**

The tGCS was slightly better than the mGCS at discriminating patients who underwent emergency intubation from those who did not undergo intubation. Based on six studies, the pooled AUROC for the tGCS was 0.865 (95% CI 0.830 to 0.901) and for the mGCS was 0.822 (95% CI 0.775 to 0.870), with a mean difference of 0.034 (95% CI 0.020 to 0.048; I²=88%; Figure 9).17,20,36,48,59,61 Although statistical heterogeneity was present, estimates were similar across studies (mean difference in AUROC ranged from 0.018 to 0.050) and all studies favored the tGCS. There was no clear difference in estimates between two studies that focused on out-of-hospital intubation (mean difference in AUROC 0.039 and 0.040)20,36 and four studies that evaluated any emergency intubation (ED or out-of-hospital) (mean difference in AUROC 0.018 to 0.050).17,48,59,61 All of the studies evaluated mixed populations of adults and children and were conducted in the United States. There were no differences when studies were stratified according to use of out-of-hospital or ED GCS scores, when analyses were restricted to low risk of bias studies, or when we excluded studies with potential overlap in populations. One study reported subgroup findings for trauma patients with TBI;17 as in the analysis of patients with any trauma, results favored the tGCS (mean difference in AUROC 0.011, 95% CI -0.010 to 0.032). No study was based on data from the NTDB.
Figure 9. Pooled AUROC of intubation for the total Glasgow Coma Scale versus the motor component only

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference (95% CI)</td>
<td>tGCS (95% CI)</td>
<td>mGCS (95% CI)</td>
</tr>
<tr>
<td>Adult or mixed or unclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>8,412</td>
<td>-</td>
<td>0.04 (0.02 to 0.06)</td>
<td>0.86 (NR)</td>
<td>0.83 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,209</td>
<td>-</td>
<td>0.04 (0.03 to 0.05)</td>
<td>0.83 (0.81 to 0.84)</td>
<td>0.79 (0.78 to 0.80)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>-</td>
<td>0.05 (0.04 to 0.06)</td>
<td>0.86 (0.85 to 0.87)</td>
<td>0.81 (0.80 to 0.82)</td>
</tr>
<tr>
<td>Thompson et al., 2011</td>
<td>19,408</td>
<td>-</td>
<td>0.05 (-0.01 to 0.11)</td>
<td>0.70 (0.63 to 0.77)</td>
<td>0.65 (0.60 to 0.70)</td>
</tr>
<tr>
<td>Beskind et al., 2014</td>
<td>9,816</td>
<td>-</td>
<td>0.02 (0.00 to 0.03)</td>
<td>0.97 (0.95 to 0.98)</td>
<td>0.95 (0.93 to 0.96)</td>
</tr>
<tr>
<td>Kupas et al., 2016</td>
<td>370,352</td>
<td>-</td>
<td>0.02 (0.02 to 0.02)</td>
<td>0.90 (0.90 to 0.91)</td>
<td>0.88 (0.88 to 0.89)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 88.5%, p = 0.000)</td>
<td></td>
<td>0.03 (0.02 to 0.05)</td>
<td>0.86 (0.83 to 0.90)</td>
<td>0.82 (0.77 to 0.87)</td>
<td></td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td>0.03 (0.02 to 0.05)</td>
<td>0.86 (0.78 to 0.93)</td>
<td>0.82 (0.73 to 0.90)</td>
<td></td>
</tr>
</tbody>
</table>

AUROC=area under the receiver operating characteristic curve; CI=confidence interval; mGCS=motor Glasgow Coma Scale; n=number; NR=not reported; tGCS=total Glasgow Coma Scale
Total Glasgow Coma Scale Versus Simplified Motor Scale

The tGCS was slightly better than the SMS at discriminating patients who underwent emergency intubation from patients who did not undergo intubation; results were very similar to the comparison of tGCS versus mGCS. Based on five studies, the pooled AUROC for the tGCS was 0.843 (95% CI 0.823 to 0.864) and for the SMS was 0.783 (95% CI 0.747 to 0.819), with a mean difference of 0.040 (95% CI 0.030 to 0.050, I²=55%, Figure 10).20,36,48,54,59 Although statistical heterogeneity was present, the estimates from individual studies were similar (mean difference in AUROC ranged from 0.030 to 0.050), and all studies favored the tGCS. All of the studies defined severe brain injury similarly, based on head CT imaging findings. All of the studies were conducted in the United States and none focused on children or patients with TBI. There were no differences in estimates between studies that utilized out-of-hospital GCS scores or ED GCS scores (Table 4). Results were also unchanged when analyses were restricted to low risk of bias studies, or when studies with potential overlap20,59 were excluded. No study utilized data from the NTDB. Results were unchanged using the profile likelihood method.
Figure 10. Pooled AUROC of intubation for the total Glasgow Coma Scale versus the Simplified Motor Scale

<table>
<thead>
<tr>
<th>Author, year</th>
<th>N</th>
<th>AUROC Difference (95% CI)</th>
<th>tGCS (95% CI)</th>
<th>SMS (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult or mixed or unclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill et al., 2005</td>
<td>8,412</td>
<td>0.04 (0.02 to 0.06)</td>
<td>0.86 (NR)</td>
<td>0.83 (NR)</td>
</tr>
<tr>
<td>Gill et al., 2006</td>
<td>7,209</td>
<td>0.04 (0.03 to 0.05)</td>
<td>0.83 (0.81 to 0.84)</td>
<td>0.79 (0.77 to 0.80)</td>
</tr>
<tr>
<td>Haukoos et al., 2007</td>
<td>21,170</td>
<td>0.05 (0.04 to 0.06)</td>
<td>0.86 (0.85 to 0.87)</td>
<td>0.81 (0.80 to 0.82)</td>
</tr>
<tr>
<td>Thompson, et al., 2011</td>
<td>19,408</td>
<td>0.05 (-0.01 to 0.11)</td>
<td>0.70 (0.63 to 0.77)</td>
<td>0.65 (0.62 to 0.67)</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>9,816</td>
<td>0.03 (0.02 to 0.04)</td>
<td>0.86 (0.85 to 0.87)</td>
<td>0.83 (0.82 to 0.83)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 55.2%, p = 0.063)</td>
<td></td>
<td>0.04 (0.03 to 0.05)</td>
<td>0.84 (0.82 to 0.86)</td>
<td>0.78 (0.75 to 0.82)</td>
</tr>
<tr>
<td>Profile likelihood method</td>
<td></td>
<td>0.04 (0.03 to 0.05)</td>
<td>0.83 (0.77 to 0.88)</td>
<td>0.78 (0.71 to 0.85)</td>
</tr>
</tbody>
</table>

AUROC=area under the receiver operating characteristic curve; CI=confidence interval; n=number; NR=not reported; SMS=Simplified Motor Scale; tGCS=total Glasgow Coma Scale
One study found the out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity than the SMS (cutoff of ≤1) for identifying patients who underwent emergency intubation (76%, 95% CI 74 to 77 vs. 73%, 95% CI 71 to 74) and slightly lower specificity (89%, 95% CI 89 to 89 vs. 91%, 95% CI 90 to 91; Table 5).54

**Motor Glasgow Coma Scale Versus Simplified Motor Scale**

There was no clear difference between the mGCS versus the SMS in ability to discriminate between patients who undergo emergency intubation from those who do not undergo emergency intubation (4 studies, mean difference in AUROC 0.000, 95% CI -0.007 to 0.007, $I^2=0\%$).20,36,48,59 There was no statistical heterogeneity and findings were unchanged in sensitivity and stratified analyses.

**Trauma Center Need**

Two studies (n=811,143 and n=393,877) evaluated the predictive utility of the tGCS versus the mGCS for identifying patients in need of trauma center care (defined as ISS score of >5, ICU admission >24 hours, need for urgent surgery, or death in the ED).53,61 Differences in the AUROC (0.617 vs. 0.609 [CIs not reported] and 0.641, 95% CI 0.639 to 0.642 vs. 0.603, 95% CI 0.602 to 0.604), sensitivity (30% vs. 27% [CIs not reported] and 28%, 95% CI 28 to 28 vs. 25%, 95% CI 25 to 25), and specificity (93% vs. 95% [CIs not reported] and 94%, 95% CI 94 to 94 vs. 95%, 95% CI 95 to 95) were small. The adjusted risk estimates were also similar (odds ratio [OR] 3.03, 95% CI 2.95 to 3.13 for tGCS score of ≤13 vs. >13 and OR 3.37, 95% CI 3.27 to 3.48 for mGCS score of ≤5 vs. >5).

**Other Outcomes**

Two studies found the tGCS was better able to discriminate those with major injury (defined as an ISS score of >15) from those without major injury (AUROC 0.720, 95% CI 0.715 to 0.724 vs 0.681, 95% CI 0.677 to 0.68655 and 0.648, 95% CI 0.646 to 0.650 vs. 0.606, 95% CI 0.605 to 0.60861). One study found the out-of-hospital tGCS (cutoff of ≤13) associated with slightly higher sensitivity than the mGCS (cutoff of ≤5) for identifying people with major injury (defined as an ISS score of >15) from those without major injury (sensitivity 31%, 95% CI 31 to 32 vs. 28%, 95% CI 28 to 28 and slightly lower specificity 91%, 95% CI 91 to 91 vs. 93%, 95% CI 93 to 93).61 Two studies found that the tGCS was better able to discriminate those admitted to the ICU from those not admitted to the ICU (AUROC 0.772, 95% CI 0.754 to 0.790 vs. 0.721, 95% CI 0.705 to 0.738, p<0.00151 and 0.625, 95% CI 0.623 to 0.626 vs. 0.583, 95% CI 0.581 to 0.584, difference 0.042, 95% CI 0.041 to 0.04361). One study (n=2,231) of children with TBI found that the tGCS was better able to discriminate those with a length of stay 5 days or longer from those with a stay of more than 5 days (0.683, 95% CI 0.660 to 0.706 vs. 0.644, 95% CI 0.622 to 0.666, p<0.001), those discharged to rehabilitation from those not discharged to rehabilitation (0.804, 95% CI 0.782 to 0.826 vs. 0.766, 95% CI 0.740 to 0.792, p<0.001), and those dependent on a caregiver from those not dependent on a caregiver following discharge (0.757, 95% CI 0.732 to 0.783 vs. 0.747, 95% CI 0.722 to 0.772, p=0.06).51
Key Question 1a. How does predictive utility 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, 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 ED or urban vs. rural location)?

Key Points

- **Age:** Differences in the AUROC between the tGCS versus the mGCS were similar in studies that enrolled children and those that enrolled mixed populations of adults and children (SOE: Low).

- **Type of trauma:** Differences in the AUROC between the tGCS versus the mGCS were similar in studies that evaluated patients with TBI and those that enrolled mixed trauma patients (SOE: Low).

- **Out-of-hospital vs. ED assessment:** One study of adults found no differences between out-of-hospital and ED GCS scores on discrimination for in-hospital mortality or neurosurgical intervention but another study of adults or children found out-of-hospital GCS scores associated with higher discrimination for in-hospital mortality than ED scores (AUROC 0.754 vs. 0.635, p-value not reported). Differences in the AUROC between the tGCS versus the mGCS were similar in studies that evaluated out-of-hospital GCS scores and those that used ED scores (SOE: Insufficient).

- **No study evaluated how intoxication status, blood pressure, intubation status, receipt of field intubation, or level/training of field assessors impacts comparative predictive utility of the tGCS versus the mGCS or SMS.

Detailed Synthesis

**Age**

No study that evaluated mixed populations of adults and children performed analyses stratified according to age group. Among the head-to-head studies on predictive utility, two studies on in-hospital mortality, one study on neurosurgical interventions, and two studies on severe brain injuries focused on children. For all of these outcomes, differences between the tGCS versus the mGCS in the AUROC slightly favored the tGCS and estimates were similar in studies that evaluated children and studies that evaluated adults or mixed populations (Table 4). For in-hospital mortality, the mean difference in AUROC was 0.006 (95% CI 0.002 to 0.011, I²=0%) in 2 studies of children and 0.019 (95% CI 0.012 to 0.025, I²=75%) in 10 studies of adults or mixed populations. For neurosurgical intervention, results were similar in one study of children (mean difference in AUROC 0.034, 95% CI 0.009 to 0.059) and six studies of adults or mixed populations of adults and children (mean difference in AUROC 0.031, 95% CI 0.018 to 0.043, I²=72%). For severe brain injury, the mean difference in AUROC was higher in one study of children (0.121, 95% CI 0.068 to 0.174) than in four studies of adults or mixed populations of adults and children (0.046, 95% CI 0.038 to 0.054, I²=0%), but there was no statistically significant interaction with age (p=0.07).
The study in children used a broader definition for severe brain injury (based on head CT imaging findings or need for intervention) than the studies conducted in adults or mixed populations, which focused on CT imaging findings. One study of children up to 18 years of age found no clear differences in AUROC estimates for mGCS between the subgroup of children 0 to 3 years of age and the entire cohort for survival to hospital discharge (0.936, 95% CI 0.911 to 0.962 vs. 0.941, 95% CI 0.926 to 0.957), craniotomy (0.659, 95% CI 0.597 to 0.721 vs. 0.638, 95% CI 0.601 to 0.675), or ICU admission (0.723, 95% CI 0.696 to 0.750 vs. 0.721, 95% CI 0.705 to 0.738), or length of stay greater than 4 days (0.589, 95% CI 0.555 to 0.623 vs. 0.644, 95% CI 0.622 to 0.666). The mGCS was associated with slightly better discrimination for being dependent on a caregiver in those 0 to 3 years (0.787, 95% CI 0.752 to 0.821 vs. 0.747, 95% CI 0.722 to 0.772).

Two studies of children with TBI found a tGCS score of 12 or 13 or less associated with sensitivity of 80 percent and 100 percent for in-hospital mortality; specificity was 74 percent and 86 percent (Appendix I). In three studies of adults or mixed populations of adults and children, sensitivity of the tGCS (cutoff of ≤13) ranged from 71 percent to 80 percent and specificity from 68 percent to 88 percent (Table 5).

One study evaluated the diagnostic accuracy of tGCS in older (≥70 years) versus younger (<70 years) adults (Appendix I). Based on a tGCS cutoff of 13 or less, it found that sensitivity of the tGCS was lower in adults 70 years or older and worse by more than 20 percent versus those younger for in-hospital mortality (51% vs. 86%), severe TBI (28% vs. 53%), neurosurgical intervention (43% vs. 66%), and emergency intubation (58% vs. 78%). Specificity was greater than 90 percent on these outcomes in adults 70 years or older and 5 percent to 10 percent higher than adults younger than 70 years. At a cutoff of 14 or less, sensitivity improved in patients 70 years or older by about 10 percent on all outcomes and specificity decreased by 5 percent to 10 percent, but the differences compared with patients younger than 70 years of age remained similar. In older adults, decreases in the tGCS to 14 from 15 (adjusted OR [AOR] 1.40, 95% CI 1.07 to 1.83) and from 13 to 14 (AOR 2.34, 95% CI 1.57 to 3.52) were associated with greater risk of in-hospital mortality than in adults younger than 70 years of age (AOR 1.22, 95% CI 0.88 to 1.71 and 1.45, 95% CI 0.91 to 2.30).

Type of Trauma

No study that evaluated mixed populations of trauma patients performed analyses stratified according to the type of trauma. Among the head-to-head studies on predictive utility, differences in the AUROC between the tGCS and the mGCS were similar among studies that focused on patients with TBI versus those that evaluated patients with mixed trauma (Table 4). For in-hospital mortality, the difference in the AUROC was 0.009 (95% CI -0.002 to 0.020) in 3 studies of TBI patients and 0.017 (95% CI 0.009 to 0.024) in 10 studies of mixed trauma patients. For neurosurgical intervention, the difference in the AUROC was 0.017 (95% CI -0.022 to 0.056) in two studies of TBI patients and 0.031 (95% CI 0.018 to 0.044) in six studies of mixed trauma patients. For emergency intubation, the difference in the AUROC was 0.011 (95% CI -0.010 to 0.032) in one study of TBI patients and 0.034 (95% CI 0.020 to 0.048) in six studies of mixed trauma patients.

No head-to-head study evaluated the predictive utility of the tGCS versus the SMS specifically in patients with TBI, or evaluated effects of type of TBI or presence of other injuries on predictive utility. One study of patients with multiple injuries (based on head and skeletal injury AIS scores ≥3) found the tGCS (cutoff of ≤12) associated with sensitivity of 87 percent
(95% CI 78 to 94) and specificity of 71 percent (95% CI 68 to 74, Appendix I). One study of TBI patients found the tGCS associated with an AUROC for in-hospital mortality of 0.85 (95% CI 0.80 to 0.90) among all patients and 0.82 (95% CI 0.77 to 0.86) among the subgroup of patients with a GCS score less than 15 (Appendix I).

Field Versus Emergency Department Assessment

Among the head-to-head studies on predictive utility, differences between the tGCS versus the mGCS or the tGCS versus the SMS were similar among studies that used out-of-hospital GCC scores versus ED GCS score (Table 4). For in-hospital mortality, the difference in AUROC for studies that used out-of-hospital GCS scores was 0.016 (95% CI 0.007 to 0.024, I²=91%) and for studies that used ED GCS scores the difference was 0.020 (95% CI 0.006 to 0.034, I²=23%). Findings were similar for other outcomes and for comparisons of tGCS versus SMS. Of the studies that were pooled, two that used out-of-hospital GCS scores and two that used ED scores were conducted in the same trauma center/system. Differences in the AUROC between the tGCS versus the mGCS or SMS were very similar when results based on out-of-hospital versus ED GCS scores were compared from each system. For example, for in-hospital mortality, a study from the Loma Linda trauma center found out-of-hospital tGCS associated with an AUROC for in-hospital mortality of 0.89 for the tGCS versus 0.88 for the mGCS and 0.86 for the SMS. Using ED scores, the AUROCs were 0.906, 0.894, and 0.878, respectively. A study from the Denver trauma system found AUROCs for in-hospital mortality of 0.90, 0.88, and 0.87 for the tGCS, mGCS, and SMS (missing GCS data excluded from analysis). AUROCs were similar in a study from the Denver trauma system that used ED GCS scores (0.92, 0.90, and 0.89, respectively). The degree to which the patient populations in studies conducted in the same trauma system overlapped in the studies was unclear. Sample sizes were not the same in the out-of-hospital and ED studies from the same trauma center/system, indicating that some patients with out-of-hospital GCS scores did not have ED scores, or vice versa. In addition, in the Denver studies, data collection dates were not identical for the out-of-hospital (1999 to 2008) and ED (1995 to 2004) studies.

Other Factors

No study evaluated how intoxication status, blood pressure, intubation status, receipt of field intubation, or level/training of field assessors impacts comparative predictive utility of the tGCS versus the mGCS or SMS.
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 in whom initial emergency medical services transport is to a higher or lower than appropriate level of care, or proportion transferred to a higher or lower level of care)?

Key Points
- No study compared effects on the tGCS versus the mGCS or SMS on rates of over- or under-triage.

Detailed Synthesis
No study compared effects of the tGCS versus the mGCS or SMS on rates of over- or under-triage. The head-to-head studies in Key Question 1 of tGCS versus the mGCS or SMS were not designed to assess effects on over- or under-triage, because all patients underwent a single triage decision based on the tGCS, and do not reflect differences in actual triage decisions. In additions, such studies do not account for other factors that impact triage decisions.

Nonetheless, measures of diagnostic accuracy may provide some indirect information about the potential degree of over-triage (1-specificity, or the false-positive rate) and under-triage (1-sensitivity, or the false-negative rate). However, data on diagnostic accuracy were very limited. For in-hospital mortality, three studies found that differences in sensitivity between the tGCS and mGCS ranged from 0 percent to 1 percent and differences in specificity ranged from 0 percent to 5 percent (Table 5). For other outcomes, data on sensitivity and specificity differences were limited to a single study.

One study on predictive utility included in Key Question 1 attempted to estimate the net effect on over- and under-triage. It found that replacing the tGCS (cutoff of ≤13) with the mGCS (cutoff of ≤5) in the 2011 Centers for Disease Control and Prevention (CDC) guideline for field triage would result in a net decrease in over- or under-triage of 0.4 percent (prevent over-triage in 1.7% and under-triage in 0.2%, while causing over-triage in 0.4% and under-triage in 1.1%), based on accuracy for identification of patients with trauma center need (ISS of >15, ICU admission ≥24 hours, need for urgent surgery, or death in the ED). In addition, it found that benefits of the mGCS on rates of over- or under-triage may be higher, as 0.5 percent of patients who would have been under-triaged by mGCS would have received the appropriate triage decision based on other elements of the NTTP. Another study included in Key Question 1 found that based on diagnostic accuracy, rates of misclassification for in-hospital survival were very similar for the tGCS versus the mGCS (4.9% vs. 5.1%).
Key Question 2a. How do effects on over- and under-triage 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, 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 ED or urban vs. rural location)?

**Key Points**
- No evidence.

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., in-hospital mortality, morbidity, quality of life)?

**Key Points**
- No study compared effects on the tGCS versus the mGCS or SMS on clinical outcomes.

Key Question 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, 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 ED or urban vs. rural location)?

**Key Points**
- No evidence.

Key Question 4. In patients with known or suspected trauma, what is the comparative reliability (e.g., interrater and intrarater 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?

**Key Points**
- The interrater reliability of tGCS and mGCS appears to be high, but evidence was insufficient to determine if there were differences between scales (SOE: Insufficient).
- Three studies found the tGCS associated with a lower proportion of correct scores than the mGCS (differences in proportion of correct scores ranged from 6% to 27%), though the difference was statistically significant in only one study (SOE: Low).
Three studies found that training or use of a scoring aid increased the proportion of correct scores on both the tGCS and mGCS (increase in proportion of correct scores ranged from 32% to 70%) (SOE: Low).

Detailed Synthesis

Interrater Reliability

Interrater or inter-observer reliability refers to the extent to which different people using a scale arrive at the same rating for the same patient. Two studies evaluated the interrater reliability of the GCS in TBI patients presenting to the ED (Table 6, Appendix J).
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design</th>
<th>Risk of Bias</th>
<th>Country</th>
<th>Objective</th>
<th>N Type of Raters</th>
<th>Results Overall</th>
<th>Results by Patient Characteristics</th>
<th>Results by Provider Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holmes et al., 2005</td>
<td>Cross-sectional</td>
<td>High</td>
<td>USA</td>
<td>Compare the pediatric GCS in children ≤2 years to the standard GCS in children &gt;2 years.</td>
<td>N=102 patients Emergency Physicians</td>
<td>Weighted kappa (95% CI) across raters tGCS: 0.77 (0.38 to 1.00) for ≤2 year olds and 0.91 (0.75 to 1.00) for &gt;2 year olds mGCS: 0.91 (0.75 to 1.00) for ages combined</td>
<td>Weighted kappa (95% CI) across raters tGCS: 0.77 (0.38 to 1.00) for ≤2 year olds 0.91 (0.75 to 1.00) for &gt;2 year olds mGCS: not reported by age groups</td>
<td>None reported</td>
</tr>
<tr>
<td>Bledsoe et al., 2015</td>
<td>Cross-sectional</td>
<td>Low</td>
<td>USA</td>
<td>Evaluate tGCS and its components in standardized video vignettes.</td>
<td>N= 217 providers; 10 patient scenarios Advanced EMTs CCPs EMTs Nurses Paramedics Physicians Residents</td>
<td>Correct scores (95% CI), tGCS vs. mGCS Across all vignettes and participants: 33.1% (30.2 to 36.0) vs. 59.8% (58.1 to 61.5)</td>
<td>Correct scores tGCS Accuracy was lowest for scenarios with tGCS scores of 9 to 13 (&lt;20%; data taken from figure) mGCS Accuracy not reported by score or severity</td>
<td>Correct scores (95% CI) tGCS vs. mGCS by provider type Highest/best residents: 51% (44.5 to 57.5) vs. 78% (71.5 to 84.5) Lowest/worst nurses tGCS: 29% (10.3 to 47.7) EMTs mGCS: 51% (43.7 to 58.3)</td>
</tr>
<tr>
<td>Feldman et al., 2015</td>
<td>RCT</td>
<td>Low</td>
<td>USA</td>
<td>Assess ability of EMS personal to correctly score the tGCS and its components and to determine if scoring improves with the use of a scoring aid.</td>
<td>N=178 Providers EMTs Paramedics</td>
<td>Correct scores, tGCS vs. mGCS All scenarios: 41.0% vs. 50.6%</td>
<td>Correct, tGCS vs. mGCS Mild TBI scenarios: 54.2% vs. 74.6% Moderate TBI scenarios: 28.8% vs. 35.6% Severe TBI scenarios: 40.0% vs. 41.7%</td>
<td>None reported</td>
</tr>
<tr>
<td>Ease of Use/Correct Scoring</td>
<td>Author, Year of Study</td>
<td>Risk of Bias</td>
<td>Country</td>
<td>Objective</td>
<td>N</td>
<td>Type of Raters</td>
<td>Results Overall</td>
<td>Results by Patient Characteristics</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------</td>
<td>--------------</td>
<td>---------</td>
<td>-----------</td>
<td>---</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Heim et al., 2009</td>
<td>Cross-sectional High</td>
<td>Switzerland</td>
<td>Assess knowledge of GCS and scoring of a clinical scenario.</td>
<td>103 providers; 1 patient scenario Air rescue physicians</td>
<td>Incorrect (correct) scores tGCS: 36.9% (63.1%) mGCS: 27.2% (72.8%)</td>
<td>None reported</td>
<td>Incorrect (correct) scores by experience Registrars (trainees): 47.5% (52.5%) p=0.095 Fellow: 33.3% (66.7%) p=0.671 Consultant: 0% (100%) p&lt;0.05 Private practice: 36.8% (63.2%) p=1.00 Specialty was not associated with statistically significant differences in errors (anesthesia, internal medicine, general practice, other)</td>
<td></td>
</tr>
<tr>
<td>Field vs. ED Agreement</td>
<td>Kerby et al., 2007</td>
<td>High USA</td>
<td>Linkage of EMS and trauma registry data to determine if differences may cause inappropriate enrollment in out-of-hospital trials.</td>
<td>3,052 patients EMTs all levels ED personnel not specified</td>
<td>Weighed kappa (95% CI) tGCS vs. mGCS: 0.53 (0.48 to 0.58) vs. 0.48 (0.43 to 0.53)</td>
<td>None reported</td>
<td>Weighted kappa (95% CI) by transport time tGCS vs. mGCS &lt;20 minutes: 0.56 (0.50 to 0.61) vs. 0.52 (0.46 to 0.57) ≥20 minutes: 0.42 (0.32 to 0.52) vs. 0.35 (0.25 to 0.46)</td>
<td></td>
</tr>
</tbody>
</table>

CCP=critical care paramedic; CI=confidence interval; ED=emergency department; EMS=emergency medical services; EMT=emergency medical technician; GCS=Glasgow Coma Scale; KQ=Key Question; mGCS=motor Glasgow Coma Scale; n=number; RCT=randomized controlled trial; TBI=traumatic brain injury; tGCS=total Glasgow Coma Scale; vs=versus
One high risk of bias study (n=102 patients; number of raters unclear) conducted in the United States compared the interrater reliability of a pediatric version of the tGCS in preverbal children 2 years old and younger and the standard version in children older than two. 60 Two faculty emergency physicians assessed the same patients upon presentation to the ED. The agreement on tGCS was higher for older children using the standard tGCS (weighted kappa 0.91, 95% CI 0.75 to 1.00) than for younger children using the pediatric tGCS (weighted kappa 0.77, 95% CI 0.38 to 1.00), but estimates were imprecise, particularly for younger children, and CIs overlapped. The interrater reliability of the mGCS for all children in the sample (not reported by age group) was high (weighted kappa 0.91, 95% CI 0.75 to 1.00).

A moderate risk of bias study conducted in Japan assessed interrater reliability of the tGCS in 66 patients with suspected TBI. 69 The tGCS was assessed by two members of the medical team (number of assessors=33) upon arrival at the ED. The weighted kappa for the TBI patients was 0.74 (95% CI 0.71 to 0.76). The weighted kappa for all trauma patients was not reported.

Methodological shortcomings in the studies include lack of information on how patients were selected or use of a convenience sample, unclear blinding of assessors to other assessors’ ratings, and unclear timing of GCS assessment.

Ease of Use

Five studies (reported in 4 articles) evaluated the ability of medical personal to correctly score the tGCS and the mGCS, 70-72,74 based on simulated patients presented in video70,74 or written71,72 scenarios in which correct scores were determined by experts (Appendix J). In three studies, the assessors were emergency medical technicians (EMTs) and paramedics or paramedic students, 71,74 in one study the assessors were air rescue physicians, 72 and the fifth study evaluated assessors with various types and levels (e.g., EMT, nurse, physician, and resident) of training. 70 Four studies were conducted in the United States 70,71,74 and one in Switzerland. 72 Sample sizes ranged from 46 to 217 providers/raters. In four studies, the scenarios represented a spectrum of injury severity from mild to severe; the fifth study 72 used a single scenario in which the correct tGCS score was 6. Three studies 70-72 evaluated the proportion of correct scores with the tGCS and the mGCS and the other two 74 only evaluated the tGCS.

The studies differed in how they were designed. One study was a randomized controlled trial (RCT) (rated low risk of bias) that compared tGCS and mGCS scores with versus without the use of a scoring aid. 71 One article reported two studies (both rated moderate risk of bias): the first was a before-after study on the effects of video training on tGCS scoring, and the second was also a before-after study on the effects of video training on tGCS scoring, but in which participants were also randomized to use of a GCS scoring aid. 74 The other two studies (one rated low risk of bias 70 and the other high risk of bias 72) used a cross-sectional design. The RCT and one cross-sectional study were rated low risk of bias. 70,71 Methodological shortcomings in the high and moderate risk of bias studies included use of the same patient scenarios on repeat testing and unclear methods for determining correct answers; 74 the high risk of bias study only evaluated one scenario and allowed assessors 10 minutes to rate a written scenario. 72

Three studies consistently found that the proportion of correct scores was lower with the tGCS than with the mGCS. 70-72 One cross-sectional study found that the overall proportion of correct scores by assessors from different types and levels of training was 33.1 percent (95% CI 30.2 to 36.0) with the tGCS and 59.8 percent (95% CI 58.1 to 61.5) with the mGCS. 70 In two other studies, the proportion of correct scores was lower with the tGCS than with the mGCS, but differences were not statistically significant. A cross-sectional study found that the proportion of
correct scores by air rescue physicians was 63.1 percent (53.8% to 72.4%) for the tGCS and 72.8 percent (95% CI 64.2% to 81.4%) for the mGCS. An RCT found that the proportion of correct scores by EMTs or paramedics was lower using the tGCS than using the mGCS in assessors randomized to a scoring aid (56.7%, 95% CI 46.5 to 66.9 vs. 70.0%, 95% CI 60.5 to 79.5) as well as those randomized to no aid (25.0%, 95% CI 16.0 to 34.0 vs. 30.7%, 95% CI 21.1 to 40.3), though differences were not statistically significant. In this study, the overall rate of correct tGCS scores was 41 percent, and 69 percent of scores were within 1 point of the correct score. No other study reported the degree to which incorrect scores differed from correct scores (e.g., the proportion of scores that were incorrect by 1 point vs. those that differed by ≥2 points), or the proportion of scores that crossed GCS triage thresholds (e.g., patient scenario in which the correct tGCS score is 14 or 15 is scored as ≤13).

Three studies (reported in 2 articles) evaluated the effect of training and scoring aids on performance of the tGCS and the mGCS. In one study, 75 attendees at an Emergency Medical Services meeting evaluated four video scenarios depicting a spectrum of injury severity before and after watching a 13-minute training video. Across the patient scenarios, the proportion of correct scores on the tGCS increased from 15 percent before the video to 64 percent after the video (p<0.001); a similar pattern was observed for each individual patient scenario. In a second study reported in the same article, 46 students in a paramedic class watched the same video, and in addition were randomized to use or not use a tGCS reference card with the scoring for each GCS component. The proportion of correct tGCS scores improved both among those randomized to the reference card aid (50% pre-video and 100% post-video, p=0.001) as well as those randomized to no reference card aid (7.7% pre-video and 76.9% post-video, p<0.0001). In the third study, EMTs and paramedics (n=178) scored one out of nine possible written scenarios (depicting TBI and a spectrum of injury severity) after randomization to a scoring aid or no scoring aid. The proportion of correct scores was higher with the aid versus without the aid for both the tGCS (56.7% vs. 25.0%, difference 31.9%, 95% CI 18.3 to 45.6) and the mGCS (70.0% vs. 30.7%, difference 39.7%, 95% CI 26.2 to 53.1).

No study compared the ease of use of the tGCS or the mGCS in terms of the amount of time needed to complete assessments of trauma patients or satisfaction of assessors.

Key Question 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, 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 ED or urban vs. rural location)?

Key Points

- Evidence was insufficient to assess effects of patient or assessor characteristics on comparative interrater reliability of the tGCS versus the mGCS (SOE: Insufficient).
- No study evaluated how comparative interrater reliability or ease of use of the tGCS versus the mGCS vary according to assessment setting (SOE: Insufficient).
• The proportion of correct GCS scores was generally lowest for assessment of patient scenarios with moderate injury severity in three studies, including one study that evaluated the tGCS and the mGCS (SOE: Low).
• Evidence was insufficient to determine effects of level of training or professional background on the proportion of correct scores on the tGCS versus the mGCS (SOE: Insufficient).
• One study found agreement between out-of-hospital and ED scores was similar for the tGCS and the mGCS (SOE: Low).

Detailed Synthesis

Ease of Use

Patient Characteristics

Evidence regarding how patient or assessor characteristics impact ease of use of GCS scales was limited. Three studies that assessed the proportion of correct scores on the tGCS or mGCS varied according to TBI severity in the patient scenario assessed.\textsuperscript{70,71,74} In one study (75 assessors) the proportion of correct scores were 64.0 percent to 76.0 percent before viewing a training video and 89.3 percent to 98.6 percent using two patient scenarios of tGCS scores of 15.\textsuperscript{74} For the scenarios with a tGCS of 5, the proportion correct increased from 45.3 percent pre-video to 94.7 percent post-video, and in the scenario with a tGCS of 8, the proportion correct increased from 36.0 percent to 74.7 percent. In another study (178 assessors), the percent correct for both the tGCS and mGCS was highest without a scoring aid in the three mild TBI scenarios (proportion correct 44.8% for the tGCS and 58.6% for the mGCS) and lower in the three severe TBI scenarios (20% and 13.3%, respectively) and the three moderate TBI scenarios (10.3% and 20.7%, respectively).\textsuperscript{71} With a scoring aid, the proportion of correct scores at all severity levels improved but still remained highest for mild TBI (63.3% for tGCS and 90.0% for mGCS) and lower in the moderate (46.7% and 50.0%, respectively) and severe scenarios (60.0% and 70.0%, respectively). The third study (217 assessors) found that the proportion of correct scores was lowest for scenarios in which the tGCS scores were 9 to 13 (<20%) and highest for scenarios in which the tGCS score was 3 (>90% correct).\textsuperscript{70}

Rater Characteristics

Evidence on how assessor characteristics such as type or level of training impacts the proportion of correct GCS scores was limited. One study (46 assessors) found that prior to watching a training video, neither level of training nor participation in a training course was associated with correct scoring.\textsuperscript{74} After watching the video, participants who had taken a trauma care course in the last 5 years were more likely to score the scenarios correctly (p=0.001). Another study evaluated the proportion of correct scores according to training background (advanced EMTs, EMTs, critical care paramedics, paramedics, nurses, resident physicians, and staff physicians).\textsuperscript{70} The proportion of correct scores was highest for residents with both the tGCS (51%, 95% CI 44.5 to 57.5) and the mGCS (78%, 95% CI 71.5 to 84.5) than the other assessor categories. The proportion of correct scores on the tGCS was lowest for nurses (29%, 95% CI 10.3 to 47.7) and the proportion of correct mGCS scores was lowest for EMTs (51%, 95% CI 43.7 to 58.3). A third study found that the highest rate of errors in scoring on the tGCS was in registrars (physicians who had not yet obtained the rank of consultant) at 47.5 percent, compared
with 33.3 percent in fellows and 36.8 percent of physicians in private practice. Consultants had no errors, but the sample size was small (n=8). There were no clear differences in the rate of errors between different specialties (e.g., anesthesia, internal medicine, general practice, or others).

**Field Versus Emergency Department Agreement**

No study evaluated how interrater reliability or ease of use vary according to assessment setting. Four studies evaluated agreement between out-of-hospital and ED GCS scores. One high risk of bias study (n=3,052) found that agreement between out-of-hospital and ED scores for adults with blunt or penetrating trauma were similar for the tGCS (weighted kappa 0.53, 95% CI 0.48 to 0.58) and the mGCS (weighted kappa 0.48, 95% CI 0.43 to 0.53). Agreement between out-of-hospital and ED scores was somewhat higher among patients with shorter versus those with longer transport time (weighted kappa for tGCS 0.56 vs. 0.42, respectively; weighted kappa for mGCS 0.52 vs. 0.35, respectively).

The three other studies focused on the tGCS without a comparison with the mGCS and generally found high levels of agreement between out-of-hospital and ED scores. One high risk of bias study (n=7,823) found no statistically significant differences between out-of-hospital and ED tGCS scores for adult trauma patients, with 82 percent of tGCS scores falling in the same category (mild 14-15, moderate 9-13, or severe 3-8), 3 percent higher GCS (less severe in ED) compared with out-of-hospital, and 15 percent lower GCS (more severe in ED). A moderate risk of bias study (n=1,181) found good agreement between out-of-hospital and ED tGCS scores for trauma patients 15 years old or older, based on an intraclass correlation coefficient of 0.74 (95% CI, 0.37 to 1.12). In this study, 96.3 percent of out-of-hospital-ED pairs were within a predetermined range of acceptability of 3 points. Another moderate risk of bias study (n=185) found good agreement between out-of-hospital and ED tGCS scores in children, with a weighted kappa of 0.74 (95% CI 0.63 to 0.85). Although ED scores tended to be higher than out-of-hospital scores, differences were small (0.44 points on average, with no difference in median scores).
Discussion

Key Findings and Strength of Evidence

The Key Findings of this review, with overall strength of evidence ratings, are summarized in Appendix G. Details about ratings for individual strength of evidence domains are shown in Appendix F.

Based on head-to-head studies, we found that the total Glasgow Coma Scale (tGCS) is associated with slightly better predictive utility than the motor component of the Glasgow Coma Scale (mGCS), based on the area under the receiver operating characteristic curve (AUROC), a measure of discrimination. The tGCS is better able than the mGCS to discriminate patients with trauma who undergo neurosurgical intervention, have severe brain injury (traumatic brain injury [TBI]), or undergo emergency intubation from patients who do not experience these outcomes. Evidence on discrimination for identifying patients with severe injury (based on the Injury Severity Score [ISS] or criteria for trauma center need) was limited, but reported similar findings. Although the differences in discrimination are statistically significant, their clinical significance is uncertain. The difference in the AUROC on each of these outcomes ranged from 0.03 to 0.05, or “small” based on our prespecified thresholds for interpreting differences in the AUROC. The tGCS was also better than the mGCS at discriminating trauma patients who died during hospitalization from those who survived hospitalization, but the difference in the AUROC was even smaller (0.01) than for nonmortality outcomes. Findings for the tGCS versus the Simplified Motor Scale (SMS) were similar to findings for the tGCS versus the mGCS for nonmortality outcomes, though the SMS performed slightly worse than the mGCS for in-hospital mortality. This means that for every 100 trauma patients, the tGCS is able to correctly discriminate 1 to 3 more cases of in-hospital mortality from cases without in-hospital mortality than the mGCS or the SMS. Across scales, discrimination was generally higher for in-hospital mortality (0.84 to 0.89) than for nonmortality outcomes (0.71 to 0.85). Although studies varied in how they defined neurosurgical interventions, severe brain injury, and emergency intubation, findings were generally similar across definitions for these outcomes. One study in children found a greater difference in the AUROC (0.121, 95% confidence interval [CI] 0.068 to 0.174) compared with four studies conducted in mixed populations of adults and children (0.046, 95% CI 0.038 to 0.054), using a broader definition for severe brain injury that included interventions in addition to computed tomography (CT) imaging findings. However, the test for an interaction effect was not statistically significant (p=0.07).

Findings for discrimination were generally robust in sensitivity and subgroup analyses based on factors such as the age group analyzed (children vs. mixed populations of children and adults), study year (data collected in 2006 or later or included data collected before 2006), assessment setting (out-of-hospital vs. emergency department [ED]), or risk of bias ratings. However, sensitivity and subgroup analyses were limited by small numbers of studies, particularly for nonmortality outcomes. In addition, no study that evaluated mixed populations of adults and children reported results stratified by age group and no study that evaluated mixed populations of trauma patients reported results stratified by type of trauma. Therefore, findings on effects of age on predictive utility are based on cross-study comparisons. For age and type of trauma, few studies specifically evaluated children or patients with TBI, though those available reported findings similar to studies that evaluated mixed populations of adults and children or mixed trauma populations. None of the head-to-head studies specifically evaluated older patients or reported findings in this subgroup, though one study found lower accuracy of the tGCS using...
the standard cutoff score of 13 or less in adults older than 70 years of age versus those younger.\textsuperscript{65} Another study found that among patients younger than 18 years of age, differences between the tGCS versus the mGCS on discrimination for in-hospital mortality and other outcomes were similar for children 0 to 3 years of age, in whom the verbal component is difficult to assess, and the whole cohort.\textsuperscript{51} There were also no clear differences when studies were stratified according to out-of-hospital versus ED assessment of the Glasgow Coma Scale (GCS). In addition, studies that evaluated out-of-hospital versus ED scores in the same trauma center reported similar discriminative performance. There was insufficient evidence to determine how intubation status, intoxication status, receipt of field interventions, timing of GCS assessment, or level of training of people administering the GCS impacted predictive utility. No study evaluated how different GCS assessments performed in intoxicated patients or after intubation, or how performance varied according to receipt of out-of-hospital interventions. In the case of intoxication or intubation, the tGCS is often limited to the motor component due to the inability to accurately assess the verbal and eye domains. Studies on the effects of alcohol intoxication have shown somewhat mixed results, with some finding little effect on tGCS scores and others reporting lower tGCS scores in certain subgroups.\textsuperscript{77-79} No study evaluated effects of the type or level of training of GCS assessors on predictive utility.

Several studies on discrimination for in-hospital mortality utilized data from the National Trauma Data Bank (NTDB). Over 700 centers across the United States contribute to the NTDB.\textsuperscript{80} This could result in double counting of patients analyzed in studies based on single centers and overweighting of such patients, if that center contributes data to the NTDB and depending on whether the NTDB and single-center studies utilized data from the same time frame. There was insufficient information on the NTDB Web site to determine the extent to which trauma centers reported in single-center studies contributed to NTDB. However, excluding NTDB studies had little impact on our findings regarding in-hospital mortality, and estimates from the NTDB studies were very similar to the estimates from the studies conducted at single trauma centers. In cases where there were multiple studies from the same center (e.g., Loma Linda\textsuperscript{20,36} and Denver\textsuperscript{48,51,59}) and potential overlap in patient populations, restricting the analysis to the most recent study based on out-of-hospital GCS data also had little impact on findings.

Few studies reported the comparative diagnostic accuracy (sensitivity, specificity) of the tGCS versus the mGCS, but findings were generally consistent with analyses based on the AUROC. Based on standard cutoffs for the tGCS (≤13), mGCS (≤5), and SMS (≤1), differences in sensitivity and specificity were small. The consistency of findings between measures of discrimination and diagnostic accuracy may be expected, given that discrimination is calculated from sensitivity and specificity over the range of test values. There was insufficient evidence to compare the performance of triage instruments based on other measures of predictive utility, such as calibration or adjusted risk estimates.

No study evaluated how using the tGCS versus the mGCS or SMS impacts the likelihood of over- or under-triage. Head-to-head studies of the tGCS versus the mGCS or SMS were not designed to assess effects on over- or under-triage because the mGCS and SMS were taken from the tGCS, with each patient only undergoing a single triage decision. Studies that evaluate the tGCS, mGCS, or SMS alone are not helpful for assessing effects on over- or under-triage because they cannot isolate the effects of the GCS assessment from the many other factors that impact triage decisions.\textsuperscript{4,81} Measures of diagnostic accuracy may provide some indirect indication of the potential degree of over- and under-triage, with 1-sensitivity indicating the
proportion of patients who experience the outcome who would be missed (under-triage) and 1-specificity indicating the proportion of patients without the diagnosis (over-triage). However, this is an oversimplification that assumes that GCS assessments are the primary or sole driver of triage decisions, even though such decisions are known to be multifactorial and depend on other patient factors (e.g., presence of hypotension, type of injury) and other variables (e.g., proximity to a trauma center). Nonetheless, as noted above, limited evidence suggests no marked differences in predictive utility between the tGCS, mGCS, and SMS, including two studies that attempted to assess overall impact on over- or under-triage based on diagnostic accuracy estimates.18,53

No study evaluated how using the tGCS versus the mGCS or SMS impacts the likelihood of clinical outcomes such as in-hospital mortality, morbidity, or quality of life. As for over- or under-triage, understanding comparative effects on clinical outcomes requires head-to-head studies in which trauma patients who are assessed using different GCS scales are followed over time, in order to assess effects of the GCS scales versus other factors that impact clinical outcomes.

Evidence on interrater reliability and ease of use was limited. For assessment of patients with trauma, there was insufficient evidence to determine comparative interrater reliability of the tGCS, mGCS, and SMS, as there was only one head-to-head study with methodological limitations and imprecise estimates.60 Other studies found the mGCS associated with higher interrater reliability than the tGCS, but were excluded either because they did not report results separately for trauma patients20,36 or because they focused on interrater reliability of the GCS among hospitalized patients. Studies that addressed ease of use were limited to those that evaluated whether the measures were scored correctly when applied to video or written patient scenarios, compared with a reference standard (usually expert assessment). Three studies found that the percentage of correct scores was higher for the mGCS than the tGCS,70-72 though the difference was statistically significant in only one study.70 Limited evidence suggests that errors are more frequent when assessing patient scenarios, indicating moderate injury severity (tGCS scores of 9-13).70,71,74 For both scales, use of a scoring aid or training appears to improve the proportion of correct scores. No study evaluated other measures of ease of use, such as time to complete the assessment or assessor satisfaction.

One study found that agreement between field and ED scores was similar for the tGCS and mGCS.73 Although differences between field and ED scores were noted for both scales, the study also found that blood pressure readings changed. Therefore, some differences between field and ED scores may accurately reflect changing status of the patient due to receipt of out-of-hospital interventions and evolving clinical status, rather than true lack of agreement.

Findings in Relationship to What Is Already Known

A prior Centers for Disease Control and Prevention (CDC) guideline found limited evidence on the predictive utility of the tGCS versus the mGCS.4 Our review included a number of studies that were published after the CDC guideline and provides more robust findings, particularly since some studies evaluated very large samples, enabling precise estimates. Our findings on predictive utility are consistent with a prior systematic review comparing the tGCS versus the SMS that also found that discrimination was similar for these scales.15 Like the CDC guideline, we found insufficient evidence to determine effects on over- or under-triage, or on clinical outcomes. Our finding that the interrater reliability of the tGCS and the mGCS are similar was
based on very limited evidence in trauma patients; studies that included nontrauma patients have found the tGCS to be associated with lower interrater reliability.23,82

**Applicability**

Our findings on predictive utility of different GCS scales appear to have broad applicability to field triage in the United States, as they are primarily based on large studies conducted in U.S. trauma settings in mixed populations of adults and children with various types of trauma. We also restricted study inclusion to studies published in 1995 or later, with most studies utilizing data collected through the last 5 to 10 years, suggesting high applicability to use in the context of current trauma systems.

Nonetheless, we identified a number of factors that can impact applicability. Despite the broad applicability of the evidence, its applicability to specific patient populations (e.g., specific type of trauma, age, presence and severity of intoxication, presence of medical comorbidities, and presence of other injuries) is less certain. For example, a modified version of the tGCS is utilized in young children83 and the GCS was originally developed for assessment of TBI,11,12 not trauma injuries or conditions associated with impaired consciousness in general. Limited evidence from across-study comparisons suggests similar results in children versus mixed populations of adults plus children and in patients with TBI versus mixed trauma populations. Within the subgroup of patients with TBI, 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 injury84 (diffuse injury resulting from contrecoup forces), and TBI often occurs in conjunction with other injuries. However, no study evaluated how the type of TBI injury or co-occurring injuries impacts performance of GCS scales. No study evaluated how predictive utility varied according to the level or training of field personnel (e.g., emergency medical responder, emergency medical technician [EMT], advanced EMT, paramedic, physician, or nurse24). In fact, no study that used out-of-hospital scores reported the training of the people administering the GCS. Another factor that could impact applicability is that the performance of the tGCS and mGCS may be different when administered soon after injury (in the field) as opposed to later (after field stabilization and destination decisions have been made or patients have arrived in the ED). A number of studies on predictive utility were conducted in ED settings, which are more controlled and easier to study than field settings,25 but may be of limited applicability to field settings. However, we found that predictive utility was similar in studies that utilized out-of-hospital versus ED GCS scores. We also found no clear differences in estimates of predictive utility when we restricted analyses to studies conducted in U.S. settings or to more recent (2006 or later) studies, which may be more applicable to current U.S. practice.

The differences between the tGCS and the mGCS or SMS in mean AUROC ranged from 0.01 to 0.05. This indicates that the ability of the scales to distinguish patients who experience an outcome from those who do not based on a higher score is 1 percent to 5 percent higher with the tGCS than with the more abbreviated scales. These differences were statistically significant, in part due to the large sample sizes evaluated in the studies. Although we classified such differences as “small,” based on a priori thresholds, such thresholds are by nature somewhat arbitrary. The importance of “small” differences in discrimination depend in part on the seriousness of the outcome evaluated, the degree to which triage and other treatment decisions are based on the field triage scale, and the degree to which such actions impact clinical outcomes.
Studies on ease of use focused on the proportion of correct scores using video or written scenarios, as opposed to assessment of actual patients/situations. While some studies limited the time allowed for the assessment and utilized scenarios indicating a spectrum of injury intensity, the applicability of such studies to actual field conditions is uncertain. In addition, there was insufficient evidence to determine the effects of incorrect scoring on triage decisions (e.g., whether the incorrect score would result in a change from being above the threshold for transport to a high-level trauma center to below, or vice versa).

**Implications for Clinical and Policy Decisionmaking**

Our review has implications for clinical and policy decisionmaking. Because we found no evidence on effects on clinical outcomes or risk of over- or under-triage, decisions regarding the selection of field assessment scales for trauma must rely on comparative predictive utility. Therefore, decisions regarding implementation of simplified GCS scales for field triage must consider the potential trade-offs between predictive utility and ease of use. Although the tGCS appears to have slightly greater discrimination than the mGCS or SMS for in-hospital mortality, severe brain injury, and markers of severe injury such as receipt of neurosurgical interventions or emergency intubation, differences were relatively small and of uncertain clinical significance. It is also possible that some of the differences in predictive utility could be reduced because field triage personnel also use other factors to inform triage decisions. Limited evidence suggests that the mGCS may be easier to score correctly than the tGCS, which may offset its slightly lower predictive utility.17,71,72

Similar results for the mGCS and the SMS might be expected because the SMS utilizes the same information as the mGCS, with the only difference being that for the SMS, patients with low scores on the mGCS (0 to 4) are collapsed into a single category. Therefore, any differences in predictive utility between the SMS and mGCS are likely to be primarily related to ease of use and reliability. Therefore, studies in which the mGCS and SMS scores are derived from the same assessment are inadequate for evaluating comparative performance; rather, studies that independently apply the SMS and mGCS are needed.

Evidence on how factors related to patients, assessors, and settings impacts predictive utility is limited. However, even if such differences exist, there may be advantages to having a single scale that can be applied across trauma scenarios, instead of requiring field assessors to select from among different scales for particular situations, even if the predictive utility of the single scale is slightly lower in certain situations.

Although evidence on comparative interrater reliability was very limited, scales with poor interrater reliability would also be expected to be associated with low predictive utility, which depends in part on the reliability of assessments. As differences in predictive utility were small, differences in interrater reliability are unlikely to be a major factor driving clinical and policy decisions regarding selection of field assessment scales. For all field assessment scales, training and use of scoring aids is likely to improve reliability and accuracy of scoring.

**Limitations of the Review Process**

Our review process had some limitations. Because of anticipated heterogeneity due to differences in patient populations, outcomes, assessment settings, and other factors, we used the random effects DerSimonian-Laird model to pool data. Statistical heterogeneity was moderate or high in some analyses. The DerSimonian-Laird estimator can result in confidence intervals that are too narrow when statistical heterogeneity is present.46 Therefore, we also performed analyses
using an alternative random effects model, the profile likelihood method, to evaluate whether findings were sensitive to the statistical model used. Results were similar using the profile likelihood method. Even when statistical heterogeneity was high, estimates for differences in AUROC across studies were generally quite similar, with statistical heterogeneity largely related to the presence of large sample sizes and very precise estimates. For example, for the analysis of tGCS versus mGCS for in-hospital mortality, the I² value was 85 percent, but differences in the AUROC ranged from 0.01 to 0.07 and all favored the tGCS (pooled estimate 0.016, 95% CI 0.009 to 0.022).

Another limitation is that we had to impute CIs for some studies included in the pooled analyses. However, findings were similar when we used alternative imputation methods. Also, no head-to-head study performed the tGCS and mGCS or SMS separately. Rather, the mGCS and SMS were retrospectively determined for each patient from the tGCS. Therefore, the tGCS and mGCS or SMS were not performed independently, and it is uncertain how findings on the other GCS components may have impacted scoring on the motor component. Because data to estimate the correlation between tGCS and mGCS scores were limited, we assumed moderate correlation. Findings were similar in sensitivity analyses that utilized alternative correlation assumptions.

Most studies on predictive accuracy focused on differences in discrimination based on the AUROC, with only a few studies reporting sensitivity and specificity of the tGCS, mGCS, or SMS using standard thresholds. Analyses based on the AUROC tend to favor measures based on scales with more input points, which “smooth” the curve and could explain some of the superior discrimination observed with the tGCS (3 to 15 scale) over more abbreviated scales. Studies that reported sensitivity and specificity based on standard cutoffs reported small differences between the tGCS and the mGCS.

No study compared effects of the tGCS versus the mGCS or the SMS on risk of over- or under-triage, as measured by differences in the rate of inappropriate transport decisions as defined using standardized criteria for need for high level trauma center care. In addition, no study evaluated rates of transfer to a higher or lower level of care, which may be a marker for over- or under-triage. Should such studies have been available, their interpretation would have been a challenge because factors other than the severity of trauma, such as the clinical course, geographic proximity, and availability of resources, may impact transfer decisions. In addition, minimally injured over-triaged patients may be discharged directly home rather than transferred to lower level care.

Most studies on predictive utility had methodological limitations, including failure to report attrition, missing data, and unclear methods for measuring outcomes. However, restricting analyses to low risk of bias studies had little impact on findings. In addition, studies that reported low missing GCS data reported results similar to those with high or unclear missing data. Many of the studies on reliability and ease of use of the tGCS and mGCS scales had major flaws, as reflected by insufficient and low strength of evidence ratings.

We restricted analyses to English-language studies, which could result in language bias. However, we identified no foreign-language study that appeared to meet inclusion criteria, and our focus was on studies conducted in U.S. trauma settings, which are unlikely to be published in languages other than English. We were limited in our ability to assess publication bias, given the relatively small number of studies. Although we searched for relevant studies on ClinicalTrials.gov, this database focuses on clinical trials. We did not identify any ongoing
studies on predictive utility of the GCS, which are unlikely to be registered in this database, or relevant clinical trials.

A number of studies were based on the large NTDB, which incorporates data from over 700 trauma centers across the United States, with very large sample samples. We could not reliably determine the degree to which studies that analyzed data from single centers analyzed populations with overlap with the NTDB studies. We performed sensitivity analyses in which NTDB studies were excluded, which had little impact on findings. Similarly, in situations where there was more than one study from a single trauma center with potential overlap, we found that results were similar when we focused on the most recent study from each center that used out-of-hospital GCS scores.

Gaps in the Evidence Base

The most important gap in the evidence base was the lack of evidence on effects of using different GCS scales on risk of over- or under-triage or on clinical outcomes. Although there were a fair number of head-to-head studies for predictive utility, including studies with large samples, the studies mostly evaluated mixed populations of adults and children with various types of trauma. Evidence to determine how predictive utility differs in subgroups defined by patient characteristics such as type of trauma (including presence and type of TBI), degree of intoxication, intubation status, and receipt of cointerventions was limited. There was also insufficient evidence to determine how the type or level of training of field personnel impacts predictive utility. A number of analyses and subgroup analyses were based on small numbers of studies, and should be interpreted with caution. We also identified little evidence on measures of predictive utility other than discrimination (e.g., diagnostic accuracy, calibration, adjusted risk estimates, and risk reclassification rates).

The literature on interrater reliability and ease of use was very limited. There was only one head-to-head study of interrater reliability in trauma patients, and it had methodological limitations and imprecise estimates. Studies on ease of use focused on scoring of written or video patient scenarios and did not address factors such as the time needed to complete the assessment or assessor satisfaction.

Future Research Needs

We identified several important future research needs. Head-to-head observational or randomized studies that assess one set of patients with the tGCS and another set with the SMS or mGCS are needed to understand effects on clinical outcomes as well as risk of over- or under-triage. Alternatively, studies that utilize the tGCS and simplified scales in the same patients could assess potential effects on over- or under-triage when the field assessment scales are incorporated into field assessment guidelines. However, such studies would represent less direct evidence, since they would not be based on actual differences in triage decisions. For over- and under-triage, studies should utilize standardized, validated measures to determine the appropriateness of transport and triage decisions. For predictive utility, prospective studies that independently assess patients using the tGCS and the mGCS or SMS would be useful for confirming the findings of the currently available retrospective studies, in which the mGCS or SMS were not independently assessed. Studies are needed to better understand the predictive utility in important subpopulations, including children, older patients, patients with specific types of trauma, and patients who have received field interventions prior to assessment. For patients who are intoxicated or intubated, studies that measure how frequently the tGCS reverts to the
mGCS due to the inability to assess the other GCS components would be helpful. Studies that evaluate how the predictive utility of the tGCS compares with the mGCS or SMS according to the level or type of training of assessing personnel in the field are also needed. Studies that assess measures of predictive utility other than discrimination (e.g., calibration, adjusted risk estimates, diagnostic accuracy, risk reclassification) would be useful for providing more complete information on predictive utility. Finally, head-to-head studies on interrater reliability and ease of use (including time to use, assessor comfort or satisfaction, and impacts of educational efforts) when the scales are administered to trauma patients in field settings are needed to better understand how these factors may impact decisions regarding selection of field assessment scales.

Conclusions

The tGCS is associated with slightly greater discrimination than the mGCS or SMS for in-hospital mortality, receipt of neurosurgical interventions, severe brain injury, overall injury severity, and emergency intubation, with differences in the AUROC ranging from 0.01 to 0.05. The clinical significance of small differences in discrimination is likely to be small, and could be offset by factors such as convenience and ease of use. Research is needed to understand how use of the tGCS versus the mGCS or SMS impacts clinical outcomes and risk of over- or under-triage.
References


44. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74.


# Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHRQ</td>
<td>Agency for Healthcare Research and Quality</td>
</tr>
<tr>
<td>AIS</td>
<td>Abbreviated Injury Scale</td>
</tr>
<tr>
<td>AOR</td>
<td>Adjusted Odds Ratio</td>
</tr>
<tr>
<td>AUROC</td>
<td>Area Under the Receiver Operating Characteristic Curve</td>
</tr>
<tr>
<td>CCP</td>
<td>Critical Care Paramedic</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CER</td>
<td>Comparative Effectiveness Review</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CINAHL</td>
<td>Cumulative Index to Nursing and Allied Health Literature</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>ED</td>
<td>Emergency Department</td>
</tr>
<tr>
<td>EMS</td>
<td>Emergency Medical Services</td>
</tr>
<tr>
<td>EMT</td>
<td>Emergency Medical Technician</td>
</tr>
<tr>
<td>GCS</td>
<td>Glasgow Coma Scale</td>
</tr>
<tr>
<td>HAPI</td>
<td>Health &amp; Psychosocial Instruments</td>
</tr>
<tr>
<td>ICU</td>
<td>Intensive Care Unit</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile Range</td>
</tr>
<tr>
<td>ISS</td>
<td>Injury Severity Score</td>
</tr>
<tr>
<td>KQ</td>
<td>Key Question</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Medical Literature Analysis and Retrieval System Online</td>
</tr>
<tr>
<td>mGCS</td>
<td>Motor Glasgow Coma Scale</td>
</tr>
<tr>
<td>n</td>
<td>Number</td>
</tr>
<tr>
<td>NHTSA</td>
<td>National Highway Traffic Safety Administration</td>
</tr>
<tr>
<td>NLR</td>
<td>Negative Likelihood Ratio</td>
</tr>
<tr>
<td>NPV</td>
<td>Negative Predictive Value</td>
</tr>
<tr>
<td>NR</td>
<td>Not Reported</td>
</tr>
<tr>
<td>NTDB</td>
<td>National Trauma Data Bank</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>PENTA</td>
<td>Pediatric Trauma Registry</td>
</tr>
<tr>
<td>PICOTS</td>
<td>Populations, Interventions, Comparators, Outcomes, Timing, Types of Studies, and Setting</td>
</tr>
<tr>
<td>PLR</td>
<td>Positive Likelihood Ratio</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive Predictive Values</td>
</tr>
<tr>
<td>QUADAS</td>
<td>Quality Assessment of Diagnostic Accuracy Studies</td>
</tr>
<tr>
<td>QUIPS</td>
<td>Quality in Prognostic Studies</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized Controlled Trial</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SMS</td>
<td>Simplified Motor Score</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>SOE</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>TBI</td>
<td>Traumatic Brain Injury</td>
</tr>
<tr>
<td>TEP</td>
<td>Technical Expert Panel</td>
</tr>
<tr>
<td>tGCS</td>
<td>Total Glasgow Coma Scale</td>
</tr>
<tr>
<td>TX</td>
<td>Texas</td>
</tr>
<tr>
<td>U.S. and USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>vs.</td>
<td>Versus</td>
</tr>
</tbody>
</table>
Appendix A. Search Strategies

Database: Ovid MEDLINE(R) and Ovid OLDMEDLINE(R)
Search Strategy:

--------------------------------------------------------------------------------
1     exp Glasgow Coma Scale/ (7598)
2     exp Trauma Severity Indices/ (26320)
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] (17103)
4     2 and 3 (8526)
5     1 or 4 (8526)
6     exp Craniocerebral Trauma/ (133918)
7     (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] (136121)
8     6 or 7 (196584)
9     exp Emergencies/ (35777)
10    exp Emergency Medical Services/ (105134)
11    (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. (40109)
12    exp Emergency Treatment/ (100260)
13    exp emergency medicine/ (10629)
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. (40109)
15    exp accidents/ (152529)
16    (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] (302727)
17    15 and 16 (13350)
18    9 or 10 or 12 or 13 or 14 or 17 (249494)
19    5 and 8 and 18 (990)
20    limit 19 to english language (889)
21    limit 19 to abstracts (928)
22    20 or 21 (973)

Database: Ovid MEDLINE(R)
Search Strategy:

--------------------------------------------------------------------------------
3  ((glasgow adj3 coma*) or tgc or mge or ges).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)

Database: CINAHL
Search Strategy:

1  (MH "Head Injuries+") (29202)
2  (tbi or ((head or brain* or cereb* or crani* or skull*) n3 (injur* or traum* or wound* or damag*))).mp. (31690)
3  1 or 2 (38057)
4  (MH "Trauma Severity Indices+") (10932)
((glasgow n3 coma*) or tgcs or mgcs or gcs) (6372)
4 or 5 (11,836)
(MH "Emergency Medical Services+") (69526)
(MH "Emergency Medical Technicians") or (MH "Emergency Medical Technician Attitudes") (8776)
(MH "Physicians, Emergency") or (MH "Emergency Nurse Practitioners") (2383)
(MH "Emergency Nursing") (11557)
pre-hospital* or prehospital or paramedic* or emt or ems or emergency medical technician* or ambulance* (27129)
(emergency or emergencies or accident*) n5 (triage or priorit* or classif* or identif*) (2730)
7 or 8 or 9 or 10 or 11 or 12 (90644)
3 and 6 and 13 (774)

Database: CINAHL
Search Strategy:
--------------------------------------------------------------------------------
(MH "trauma+") or (MH "wounds and injuries+") or (MH "emergency patients+") or (MH "accidents+") or (MH "violence+") (286638)
(tbi or ((head or brain* or cereb* or crani* or skull*) n3 (injur* or traum* or wound* or damag*)) (31922)
1 or 2 (293880)
(MH "Trauma Severity Indices+") (10983)
((glasgow n3 coma*) or tgcs or mgcs or gcs) (6432)
4 or 5 (11915)
(MH "Emergency Medical Services+") (69734)
(MH "Emergency Medical Technicians") or (MH "Emergency Medical Technician Attitudes") (8800)
(MH "Physicians, Emergency") or (MH "Emergency Nurse Practitioners") (2395)
(MH "Emergency Nursing") (11562)
pre-hospital* or prehospital or paramedic* or emt or ems or emergency medical technician* or ambulance* (27291)
(emergency or emergencies or accident*) n5 (triage or priorit* or classif* or identif*) (2762)
7 or 8 or 9 or 10 or 11 or 12 (91020)
3 and 6 and 13 (2364)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials
Search Strategy:
--------------------------------------------------------------------------------
injur* or traum* or wound* or damag*.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (47806)
((glasgow adj3 coma*) or tgcs or mgcs or gcs).mp. (1148)
((traum* or injur*) adj3 sever* adj5 (rated or rating* or scale*)].mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (123)
2 or 3 (1175)
5 (pre-hospital* or prehospital or paramedic* or emt or ems or emergency medical technician* or ambulance* or emergency or emergencies or accident* or triage or priorit* or classif* or identif*).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (125583)
6 1 and 4 and 5 (217)

Database: EBM Reviews - Cochrane Database of Systematic Reviews
Search Strategy:

1 (injur* or traum* or wound* or damag*).mp. [mp=title, abstract, full text, keywords, caption text] (4650)
2 ((glasgow adj3 coma*) or tgcs or mgcs or gcs).mp. (88)
3 ((traum* or injur*) adj3 sever* adj5 (rated or rating* or scale*)).mp. [mp=title, abstract, full text, keywords, caption text] (24)
4 2 or 3 (97)
5 (pre-hospital* or prehospital or paramedic* or emt or ems or emergency medical technician* or ambulance* or emergency or emergencies or accident* or triage or priorit* or classif* or identif*).mp. [mp=title, abstract, full text, keywords, caption text] (8818)
6 1 and 4 and 5 (80)

Database: PsycINFO
Search Strategy:

1 exp Traumatic Brain Injury/ (13891)
2 exp Head Injuries/ (5271)
3 exp trauma/ (59345)
4 exp accidents/ (11000)
5 exp violence/ (62018)
6 (tbi or ((head or brain* or cereb* or crani* or skull*) adj3 (injur* or traum* or wound* or damag*)>).mp. (47797)
7 1 or 2 or 3 or 4 or 5 or 6 (155040)
8 ((glasgow adj3 coma*) or tgcs or mgcs or gcs).mp. (4566)
9 ((traum* or injur*) adj5 (critical* or sever* or threat*) adj7 (rat* or scale* or index* or classif* or identif*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (1171)
10 8 or 9 (5448)
11 exp Emergency Services/ (6316)
12 exp accidents/ (11000)
13 (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. (2424)
14 (emergency or emergencies or triage or ((priorit* or early or earlie* or rapid* or quick* or swift*) adj5 (treat* or therap* or interven* or interven* or transport* or procedur*)}).mp. (60383)
15 11 or 12 or 13 or 14 (72337)
16 7 and 10 and 15 (546)

A-4
Database: Health and Psychosocial Instruments
Search Strategy:
---------------------------------------------------------------
1  (tbi or ((head or brain* or cereb* or crani* or skull*) adj3 (injur* or traum* or wound* or damag*))).mp. (1769)
2  ((glasgow adj3 coma*) or tgcs or mgcs or gcs).mp. (329)
3  ((traum* or injur*) adj3 sever* adj5 (rated or rating* or scale*)).mp. [mp=title, acronym, descriptors, measure descriptors, sample descriptors, abstract, source] (23)
4  2 or 3 (351)
5  (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. (96)
6  (emergency or emergencies or accident* or triage or priorit*).mp. (1413)
7  5 or 6 (1486)
8  1 and 4 and 7 (8)
9  4 and 7 (23)

Database: Health and Psychosocial Instruments
Search Strategy:
---------------------------------------------------------------
1  (injur* or traum* or wound* or damag*).mp. [mp=title, acronym, descriptors, measure descriptors, sample descriptors, abstract, source] (7032)
2  ((glasgow adj3 coma*) or tgcs or mgcs or gcs).mp. (330)
3  ((traum* or injur*) adj3 sever* adj5 (rated or rating* or scale*)).mp. [mp=title, acronym, descriptors, measure descriptors, sample descriptors, abstract, source] (23)
4  2 or 3 (351)
5  (pre-hospital* or prehospital or paramedic* or emt or ems or emergency medical technician* or ambulance* or emergency or emergencies or accident* or triage or priorit* or classif* or identif*).mp. [mp=title, acronym, descriptors, measure descriptors, sample descriptors, abstract, source] (6531)
6  1 and 4 and 5 (29)
### Appendix B. Inclusion and Exclusion Criteria

#### Table B-1. Inclusion and exclusion criteria

<table>
<thead>
<tr>
<th></th>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>Patients with known or suspected trauma.</td>
<td>Nonhuman population, patients without known or suspected trauma, patients transferred from another hospital.</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td><strong>GCS motor score (mGCS):</strong> Focus on studies of the mGCS using a cutoff score of ≤5 to indicate patients who require high level trauma care, but will include studies that use alternative cutoffs or modifications of mGCS.</td>
<td>Studies that evaluate the utility of mGCS or tGCS in combination with other predictors, including guidelines and triage criteria.</td>
</tr>
<tr>
<td></td>
<td><strong>GCS total score (tGCS):</strong> Focus on studies that use a cutoff tGCS score of ≤13 to indicate patients who require high level trauma care, but will include studies that use alternative cutoffs or modifications of tGCS.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Potential modifiers:</strong> 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).</td>
<td></td>
</tr>
<tr>
<td><strong>Comparisons</strong></td>
<td><strong>Main KQs:</strong> Head-to-head comparisons of mGCS vs. tGCS</td>
<td>Other measures, comparison of transferred and direct transport patients, or no comparison for main KQs</td>
</tr>
<tr>
<td></td>
<td><strong>Sub KQs:</strong> No comparison required</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td><strong>KQ1:</strong> Predictive utility for mortality, morbidity, ISS ≥16, or 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 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.</td>
<td>Costs, prevalence rates.</td>
</tr>
<tr>
<td></td>
<td><strong>KQ2:</strong> Over- or under-triage, proportion of patients who are transferred to a higher or lower level of care.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>KQ3:</strong> 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 followup.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>KQ4:</strong> Reliability (e.g., inter-rater and intra-rater kappa); ease of use (e.g., time to complete, measures of missing data, user reported satisfaction).</td>
<td></td>
</tr>
<tr>
<td><strong>Timing</strong></td>
<td>Administered soon after injury (in the field) or immediately upon arrival in the emergency department.</td>
<td>After admission to the ICU, after &gt;24 hours in the hospital.</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
<td>Out-of-hospital setting (in the field) or immediately upon arrival at the hospital emergency department.</td>
<td>Studies conducted in the ICU or in the developing world.</td>
</tr>
<tr>
<td><strong>Study Design</strong></td>
<td>Randomized trials, cohort, and case-control studies.</td>
<td>Case reports, case series, cross-sectional studies, and modeling studies.</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td>English-language abstracts (includes English-language abstracts of non English-language papers) and papers.</td>
<td>Non English-language papers.</td>
</tr>
</tbody>
</table>
Appendix C. List of Included Studies


Appendix D. List of Excluded Studies

Table D-1. Exclusion code key

<table>
<thead>
<tr>
<th>Codes</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Background or discussion paper only, no data for evidence</td>
</tr>
<tr>
<td>3</td>
<td>Wrong population (nonhuman population, patients without known or suspected trauma, transferred patients)</td>
</tr>
<tr>
<td>4</td>
<td>Wrong intervention (measure other than the GCS, composite variables [GCS + something], guidelines/triage criteria)</td>
</tr>
<tr>
<td>5</td>
<td>Wrong outcome (costs, prevalence rates, doesn’t report outcome of interest)</td>
</tr>
<tr>
<td>6</td>
<td>Wrong study design (case reports, case series, cross-sectional studies [KQ 1-3, only], and modeling studies)</td>
</tr>
<tr>
<td>7</td>
<td>Wrong publication type (opinion, editorial, letter, guideline document not used for background)</td>
</tr>
<tr>
<td>8</td>
<td>Wrong comparison (other measures, non head-to-head comparison studies for main KQs, direct vs. transferred patients)</td>
</tr>
<tr>
<td>9</td>
<td>Wrong setting (in hospital or ICU not ED, not immediately upon arrival in ED [&gt;4 hours], studies conducted in the developing world, unable to determine where/when GCS administered)</td>
</tr>
<tr>
<td>10</td>
<td>Not English language but may be relevant</td>
</tr>
<tr>
<td>11</td>
<td>Review not meeting our requirements (i.e. wrong study designs included, no risk of bias assessment, only one library searched, nonsystematic review, more updated review available)</td>
</tr>
<tr>
<td>12</td>
<td>Studies outside of search dates (published before January 1995)</td>
</tr>
<tr>
<td>13</td>
<td>Indirect studies (tGCS, mGCS, or SMS only) for KQ 1 that do not address one of the subgroups of interest (e.g., pediatrics, elderly, intoxicated individuals, intubated patients, TBI vs. other trauma patients, etc.)</td>
</tr>
</tbody>
</table>

ED=emergency department; GCS=Glasgow Coma Scale; ICU=intensive care unit; KQ=key question; mGCS=motor only component of GCS; SMS=Simplified Motor Score; TBI=traumatic brain injury; tGCS=total score of GCS; vs.=versus


Exclusion: 11


Exclusion: 4


Exclusion: 4


Exclusion: 6


Exclusion: 4


Exclusion: 4

Exclusion: 4


Exclusion: 4


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 11

Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 4


Exclusion: 3


Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 3

Baker M. Reviewing the application of the Glasgow Coma Scale: does it have interrater reliability? BJNN. 2008;4(7).

Exclusion: 11


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4

Exclusion: 4


Exclusion: 7


Exclusion: 13


Exclusion: 5


Exclusion: 7


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 3


Exclusion: 4


Exclusion: 4

Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 3


Exclusion: 5


Exclusion: 3


Exclusion: 5


Exclusion: 5


Exclusion: 3


Exclusion: 5


Exclusion: 3

Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 12


Exclusion: 11


Exclusion: 6


Exclusion: 9


Exclusion: 3


Exclusion: 4


Exclusion: 7


Exclusion: 4


Exclusion: 5


Exclusion: 7


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 6


Exclusion: 4


Exclusion: 4

Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4

Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 7

Cone DC, Domeier R. Examining the applicability of guidelines promulgated in 2003 by the American College of Surgeons' Committee on Trauma. J Trauma. 2005 Nov;59(5):1273; author reply -4.

Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Criss EA. Research review. JEMS. 2007;32(4). Exclusion: 5

Criss EA. Research review. JEMS. 2008;33(2) PMID: 18395599. Exclusion: 5

Criss EA. Research review. JEMS. 2010;35(2) PMID: 20569863. Exclusion: 7


Exclusion: 9


Exclusion: 6


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 4


Exclusion: 3


Exclusion: 4


Exclusion: 5


Exclusion: 5


Exclusion: 13


Exclusion: 13


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 7


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 7


Exclusion: 9


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4

D-14

Exclusion: 3


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 6


Exclusion: 4


Exclusion: 7


Exclusion: 4


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 3


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 9


Exclusion: 6


Exclusion: 4


Exclusion: 4

Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 13


Exclusion: 4


Exclusion: 5

Exclusion: 3


Exclusion: 5


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 6


Exclusion: 9


Exclusion: 5


Exclusion: 13

Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 13


Exclusion: 13


Exclusion: 4


Exclusion: 13


Exclusion: 4


**Exclusion: 4**


**Exclusion: 4**


**Exclusion: 4**


**Exclusion: 4**


**Exclusion: 4**


**Exclusion: 4**


**Exclusion: 4**


**Exclusion: 3**


**Exclusion: 3**


**Exclusion: 3**


**Exclusion: 3**


**Exclusion: 13**


**Exclusion: 4**

Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 10


Exclusion: 9


Exclusion: 9


Exclusion: 13


Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4

Exclusion: 4


Exclusion: 3


Exclusion: 13


Exclusion: 13


Exclusion: 4/4


Exclusion: 4


Exclusion: 4


Exclusion: 5/5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 13

Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 7


Exclusion: 13


Exclusion: 4


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 5


Exclusion: 13


Exclusion: 9


Exclusion: 4


Exclusion: 13


Exclusion: 9


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 13


Exclusion: 5


Exclusion: 5


Exclusion: 4


Exclusion: 13


Exclusion: 5


Exclusion: 5


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 4
PMID: 9846260.
Exclusion: 4

PMID: 19131821.
Exclusion: 5

PMID: 20838131.
Exclusion: 4

PMID: 16096552.
Exclusion: 4

PMID: 20370756.
Exclusion: 5

PMID: 21839444.
Exclusion: 3

Nore AK, Ommundsen OE, Steine S. [How to distinguish between illness, injury or intoxication in the emergency unit?]. Tidsskr Nor Laegeforen. 2001 Mar 30;121(9):1055-8.
PMID: 11354881.
Exclusion: 10

PMID: 12352488.
Exclusion: 13

PMID: 11245749.
Exclusion: 9

PMID: 18581133.
Exclusion: 13

PMID: 19359919.
Exclusion: 4

PMID: 19301371.
Exclusion: 4

PMID: 18022928.
Exclusion: 9

PMID: 21040114.
Exclusion: 4

Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 4


Exclusion: 4

Exclusion: 11


Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 9


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 11


Exclusion: 4


Exclusion: 4

Exclusion: 4


Exclusion: 3


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4

Exclusion: 11


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 8


Exclusion: 4


Exclusion: 8


Exclusion: 4


Exclusion: 4


Exclusion: 12

Exclusion: 7


Exclusion: 7


Exclusion: 4


Exclusion: 4


Exclusion: 5


Exclusion: 5


Exclusion: 4


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 7


Exclusion: 4


Exclusion: 4


Exclusion: 9


Exclusion: 13

Exclusion: 13


Exclusion: 4


Exclusion: 10


Exclusion: 4


Exclusion: 4


Exclusion: 11


Exclusion: 4


Exclusion: 4


Exclusion: 7


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 9

Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 6


Exclusion: 9


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 5


Exclusion: 9


Exclusion: 4


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 5


Exclusion: 6


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4

Exclusion: 13


Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 10


Exclusion: 13 for KQ1, 5 for KQ2


Exclusion: 9


Exclusion: 9


Exclusion: 13


Exclusion: 13


Exclusion: 4


Exclusion: 4


Exclusion: 4

Exclusion: 5


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 3


Exclusion: 5


Exclusion: 3


Exclusion: 5


Exclusion: 4

**Exclusion:** 4


**Exclusion:** 4


**Exclusion:** 6


**Exclusion:** 4


**Exclusion:** 5


**Exclusion:** 4


**Exclusion:** 5


**Exclusion:** 4


**Exclusion:** 8


**Exclusion:** 8


**Exclusion:** 3


**Exclusion:** 13


**Exclusion:** 4


**Exclusion:** 4


**Exclusion:** 4


**Exclusion:** 4


**Exclusion:** 4
Exclusion: 5

Exclusion: 4


Exclusion: 7

Exclusion: 4

Exclusion: 4

Exclusion: 4

Exclusion: 4

Exclusion: 4

Exclusion: 4

Exclusion: 4

Exclusion: 4

Exclusion: 4

Exclusion: 3


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 9


Exclusion: 4


Exclusion: 4


Exclusion: 4


Exclusion: 4
Appendix E. Risk of Bias Criteria

Risk Prediction Studies

Criteria:

- The study sample adequately represents the population of interest
- The study data available (i.e., participants not lost to followup) adequately represent the study sample
- The prognostic factor is measured in a similar way for all participants
- The outcome of interest is measured in a similar way for all participants
- Important potential confounding factors are appropriately accounted for
- The observed effect of the prognostic factor on the outcome is very likely to be distorted by another factor related to prognostic factor and outcome

Definitions of risk of bias based on above criteria:

Low: The least risk of bias, and 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.

Moderate: 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.

High: 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.
Reliability and Ease of Use Studies

Criteria:

Patient Selection
For assessments of interrater reliability and field versus emergency department (ED) agreement
- Are raters rating the same patient?
- Are raters rating all patients in a set time frame or a random selection of patients?
- Are no or only small numbers of patients dropped as two ratings were not possible?
For assessments of ease of use
- Is a rationale given for the sample size/number of dual ratings?
- Is scoring/assessment done on more than one patient and a range of patient situations?

Index Test(s)
For assessments of interrater reliability and field versus ED agreement
- Are raters blinded to the other rater?
Not applicable for assessments of ease of use

Reference Standard
For assessments of interrater reliability and field versus ED agreement
- Is the approach to scoring agreement explained and appropriate (e.g. Kappa, % agreement, exact match or in same category)?
For assessments of ease of use
- Is how the correct answer was determined clearly described?
- Was the correct answer verified with more than one expert?

Flow and Timing
For assessments of interrater reliability
- Are ratings of the patient being made within a reasonable amount of time or time within which it could be expected that score did not change (at the same time or within minutes)?
For assessments of field versus ED agreement
- Is ED rating immediately upon arrival?
For assessments of ease of use
- Is it clear if the field rating is before or after resuscitation?
- Was the scoring done in a way to simulate field or ED (e.g., limited amount of time, video preferable to written etc.)?

Definition of risk of bias based on above criteria:

Low: Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles missing data in a reasonable manner; includes a large number (>100), broad-spectrum of patients with and without disease; study attempts to enroll a random or consecutive sample of patients who meet inclusion criteria screening cutoffs pre-stated.
Moderate: Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (50 to 100 subjects) and a “medium” spectrum of patients (i.e. applicable to most screening settings).

High: Has important limitation such as: uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample size of very narrow selected spectrum of patients.

References
Appendix F. Strength of Evidence Domains and Definitions

Strength of Evidence Criteria

The set of five required domains comprises the main constructs that Evidence-based Practice Centers (EPCs) should use for all major outcomes and comparisons of interest. As briefly defined below in Table F-1, these domains represent related but separate concepts, and each is scored independently. The concepts are explained in more detail in below.

**Table F-1. Required domains and their definitions**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Definition and Elements</th>
<th>Score and Application</th>
</tr>
</thead>
</table>
| Study Limitations   | Study limitations is the degree to which the included studies for a given outcome have a high likelihood of adequate protection against bias (i.e., good internal validity), assessed through two main elements:  
• Study design: Whether RCTs or other designs such as nonexperimental or observational studies.  
• Study conduct. Aggregation of ratings of risk of bias of the individual studies under consideration. | Score as one of three levels, separately by type of study design:  
• Low level of study limitations  
• Medium level of study limitations  
• High level of study limitations |
| Directness          | Directness relates to (a) whether evidence links interventions directly to a health outcome of specific importance for the review, and (b) for comparative studies, whether the comparisons are based on head-to-head studies. The EPC should specify the comparison and outcome for which the SOE grade applies.  
Evidence may be indirect in several situations such as:  
• The outcome being graded is considered intermediate (such as laboratory tests) in a review that is focused on clinical health outcomes (such as morbidity, mortality).  
• Data do not come from head-to-head comparisons but rather from two or more bodies of evidence to compare interventions A and B—e.g., studies of A vs. placebo and B vs. placebo, or studies of A vs. C and B vs. C but not direct comparisons of A vs. B.  
• Data are available only for proxy respondents (e.g., obtained from family members or nurses) instead of directly from patients for situations in which patients are capable of self-reporting and self-report is more reliable.  
Indirectness always implies that more than one body of evidence is required to link interventions to the most important health outcome. | Score as one of two levels:  
• Direct  
• Indirect  
If the domain score is indirect, EPCs should specify what type of indirectness accounts for the rating. |
| Consistency         | Consistency is the degree to which included studies find either the same direction or similar magnitude of effect. EPCs can assess this through two main elements:  
• Direction of effect: Effect sizes have the same sign (that is, are on the same side of no effect or a MID)  
• Magnitude of effect: The range of effect sizes is similar. EPCs may consider the overlap of CIs when making this evaluation.  
The importance of direction vs. magnitude of effect will depend on the key question and EPC judgments. | Score as one of three levels:  
• Consistent  
• Inconsistent  
• Unknown (e.g., single study)  
Single-study evidence bases (including mega-trials) cannot be judged with respect to consistency. In that instance, use “Consistency unknown (single study)” |
<table>
<thead>
<tr>
<th>Domain</th>
<th>Definition and Elements</th>
<th>Score and Application</th>
</tr>
</thead>
</table>
| Precision       | Precision is the degree of certainty surrounding an effect estimate with respect to a given outcome, based on the sufficiency of sample size and number of events.  
  • A body of evidence will generally be imprecise if the OIS is not met. OIS refers to the minimum number of patients (and events when assessing dichotomous outcomes) needed for an evidence base to be considered adequately powered.  
  • If EPCs performed a meta-analysis, then EPCs may also consider whether the CI crossed a threshold for an MID.  
  • If a meta-analysis is infeasible or inappropriate, EPCs may consider the narrowness of the range of CIs or the significance level of p-values in the individual studies in the evidence base. | Score as one of two levels:  
  • Precise  
  • Imprecise  
  A precise estimate is one that would allow users to reach a clinically useful conclusion (e.g., treatment A is more effective than treatment B). |
| Reporting Bias  | Reporting bias results from selectively publishing or reporting research findings based on the favorability of direction or magnitude of effect. It includes:  
  • Study publication bias, i.e., nonreporting of the full study.  
  • Selective outcome reporting bias, i.e., nonreporting (or incomplete reporting) of planned outcomes or reporting of unplanned outcomes.  
  • Selective analysis reporting bias, i.e., reporting of one or more favorable analyses for a given outcome while not reporting other, less favorable analyses.  
  Assessment of reporting bias for individual studies depends on many factors—e.g., availability of study protocols, unpublished study documents, and patient-level data. Detecting such bias is likely with access to all relevant documentation and data pertaining to a journal publication, but such access is rarely available. Because methods to detect reporting bias in observational studies are less certain, this guidance does not require EPCs to assess it for such studies. | Score as one of two levels:  
  • Suspected  
  • Undetected  
  Reporting bias is suspected when:  
  • Testing for funnel plot asymmetry demonstrates a substantial likelihood of bias,  
  And/or  
  • A qualitative assessment suggests the likelihood of missing studies, analyses, or outcomes data that may alter the conclusions from the reported evidence.  
  Undetected reporting bias includes all alternative scenarios. |

CI = confidence internal; EPC = Evidence-based Practice Center; MID = minimally important difference; OIS = optimal information size; SOE = strength of evidence

**Study Limitations Domain**

**Definition**  
Scoring the study limitations domain is the essential starting place for grading strength of the body of evidence. It refers to the judgment that the findings from included studies of a treatment (or treatment comparison) for a given outcome are adequately protected against bias (i.e., have good internal validity), based on the design and conduct of those studies. That is, EPCs assess the ability of the evidence to yield an accurate estimate of the true effect without bias (nonrandom error).

**Directness Domain**

**Definition**  
Directness of evidence expresses how closely available evidence measures an outcome of interest. Assessing directness has two parts: directness of outcomes and directness of
comparisons. Applicability of evidence (external validity) is considered explicitly but separately from strength of evidence.

**Consistency Domain**

**Definition**

Consistency refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitudes of effect) across individual studies within an evidence base. EPCs may choose which of these two notions of consistency (direction or magnitude) they are scoring; they should be explicit about this choice.

**Precision Domain**

**Definition**

Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. It is based on the potential for random error evaluated through the sufficiency of sample size and, in the case of dichotomous outcomes, the number of events. A precise body of evidence should enable decisionmakers to draw conclusions about whether one treatment is inferior, equivalent, or superior to another.

**Reporting Bias**

**Definition**

Reporting bias occurs when authors, journals, or both decide to publish or report research findings based on their direction or magnitude of effect. Table F-2 defines the three main types of reporting bias that either authors or journals can introduce: publication bias and outcome and analysis reporting bias.

**Four Strength of Evidence Levels**

The four levels of grades are intended to communicate to decisionmakers EPCs’ confidence in a body of evidence for a single outcome of a single treatment comparison. Although assigning a grade requires judgment, having a common understanding of the interpretation will be useful for helping EPCs as they conduct their own global assessment and for improving consistency across reviewers and EPCs.

Table F-2 summarizes the four levels of grades that EPCs use for the overall assessment of the body of evidence. Grades are denoted high, moderate, low, and insufficient. They are not designated by Roman numerals or other symbols. EPCs should apply discrete grades and should not use designations such as “low to moderate” strength of evidence.
Each level has two components. The first, principal definition concerns the level of confidence that EPCs place in the estimate of effect (direction or magnitude of effect) for the benefit or harm; this equates to their judgment as to how much the evidence reflects a true effect. The second, subsidiary definition involves an assessment of the level of deficiencies in the body of evidence and belief in the stability of the findings, based on domain scores and a more holistic, summary appreciation of the possibly complex interaction among the individual domains.

Assigning a grade of high, moderate, or low implies that an evidence base is available from which to estimate an effect for either the benefit or the harm. The designations of high, moderate, and low should convey how confident EPCs would be about decisions based on evidence of differing grades, which can be based on either quantitative or qualitative assessment.

For comparative effectiveness questions, the comparison is typically a choice of either direction (A>B, A=B, A<B) or magnitude (difference between A and B). In some instances assigning different grades regarding the direction and the magnitude of an effect may be appropriate. An example of this situation is when studies consistently find that an intervention improves an outcome (e.g., apnea-hypopnea index is reduced by a statistically significant amount or beyond a minimally important difference), but the degree of heterogeneity about the estimate is high (e.g., range -10 to -46 events/minute; $I^2 = 86\%$).

The importance of the distinctions among high, moderate, and low levels (and the distinction with insufficient strength of evidence) can vary by the type of outcome, comparison, and decisionmaker. EPCs understand that some stakeholders may want to take action only when evidence is of high or moderate strength, whereas others may want to understand clearly the implications of low versus insufficient evidence. Even when strength of evidence is low or insufficient, consumers, clinicians, and policymakers may find themselves in the position of having to make choices and decisions, and they may consider factors other than the evidence from a specific systematic review, such as patient values and preferences, costs, or resources.

References
## Appendix G. Strength of Evidence Table

### Table G-1. Strength of evidence

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Study Design Number of Studies (N)</th>
<th>Study Limitations</th>
<th>Directness</th>
<th>Consistency</th>
<th>Precision</th>
<th>Reporting Bias</th>
<th>Main Findings</th>
<th>Strength of Evidence Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KQ 1. Predictive Utility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-hospital mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tGCS vs. mGCS: Discrimination</td>
<td>12 (756,145)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent*</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC: 0.015 (0.009 to 0.022)</td>
<td>Moderate</td>
</tr>
<tr>
<td>tGCS (≤13) vs. mGCS (≤5): Diagnostic accuracy</td>
<td>5 (1,183,836)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Imprecise</td>
<td>Not detected</td>
<td>Differences in sensitivity 0% to 3%, difference in specificity 0% to 5%</td>
<td>Low</td>
</tr>
<tr>
<td>tGCS vs. SMS: Discrimination</td>
<td>5 (110,435)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.030 (0.024 to 0.036)</td>
<td>Moderate</td>
</tr>
<tr>
<td>tGCS (≤13) vs. SMS (≤1): Diagnostic accuracy</td>
<td>1 (52,412)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Precise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 75% (73 to 76) vs. 72% (70 to 74); specificity (95% CI) 88% (87 to 88) vs. 89% (89 to 87)</td>
<td>Low</td>
</tr>
<tr>
<td>mGCS vs. SMS: Discrimination</td>
<td>4 (56,223)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.014 (0.006 to 0.021)</td>
<td>Moderate</td>
</tr>
<tr>
<td>Neurosurgical intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tGCS vs. mGCS: Discrimination</td>
<td>7 (429,124)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.031 (0.018 to 0.044)</td>
<td>Moderate</td>
</tr>
<tr>
<td>tGCS (≤13) vs. mGCS (≤5): Diagnostic accuracy</td>
<td>2 (362,432)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Inconsistent for sensitivity, consistent for specificity</td>
<td>Imprecise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 51% (50 to 52) vs. 46% (45 to 48) and 63% (38 to 84) vs. 68% (43 to 87); specificity (95% CI) 86% (86 to 86) vs. 88% (88 to 88) and 82% (80 to 84) vs. 83% (81 to 85)</td>
<td>Insufficient for sensitivity, low for specificity</td>
</tr>
<tr>
<td>tGCS vs. SMS: Discrimination</td>
<td>5 (108,635)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.032 (0.025 to 0.039)</td>
<td>Moderate</td>
</tr>
<tr>
<td>tGCS (≤13) vs. SMS (≤1): Diagnostic accuracy</td>
<td>1 (52,412)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Precise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 60% (56 to 63) vs. 53% (49 to 56); specificity (95% CI) 85% (84 to 85) vs. 86% (86 to 87)</td>
<td>Low</td>
</tr>
<tr>
<td>mGCS vs. SMS: Discrimination</td>
<td>4 (56,223)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.002 (-0.005 to 0.010)</td>
<td>Moderate</td>
</tr>
<tr>
<td>Key Question</td>
<td>Outcome</td>
<td>Number of Studies (N)</td>
<td>Study Design</td>
<td>Directness</td>
<td>Consistency</td>
<td>Precision</td>
<td>Reporting Bias</td>
<td>Main Findings</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Severe brain injury</td>
<td>tGCS vs. mGCS: Discrimination</td>
<td>5 (134,186)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent*</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.050 (0.034 to 0.065)</td>
</tr>
<tr>
<td></td>
<td>IGCS (≤13) vs. mGCS (≤5): Diagnostic accuracy</td>
<td>1 (1,410)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Precise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 62% (55 to 68) vs. 61% (54 to 67); specificity (95% CI) 85% (83 to 88) vs. 89% (88 to 91)</td>
</tr>
<tr>
<td></td>
<td>IGCS vs. SMS: Discrimination</td>
<td>5 (100,223)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent*</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.048 (0.038 to 0.059)</td>
</tr>
<tr>
<td></td>
<td>IGCS (≤13) vs. SMS (≤1): Diagnostic accuracy</td>
<td>1 (52,412)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Precise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 45% (44 to 46) vs. 41% (40 to 42); specificity (95% CI) 89% (89 to 90) vs. 90% (90 to 91)</td>
</tr>
<tr>
<td></td>
<td>mGCS vs. SMS: Discrimination</td>
<td>4 (56,223)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.000 (-0.008 to 0.007)</td>
</tr>
<tr>
<td>Emergency intubation</td>
<td>IGCS vs. mGCS: Discrimination</td>
<td>6 (436,391)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent*</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.034 (0.020 to 0.048)</td>
</tr>
<tr>
<td></td>
<td>IGCS vs. mGCS Diagnostic accuracy</td>
<td>1 (370,352)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Precise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 84% (83 to 84) vs. 81% (81 to 82); specificity (95% CI) 90% (90 to 90) vs. 92.0% (92 to 92)</td>
</tr>
<tr>
<td></td>
<td>IGCS vs. SMS: Discrimination</td>
<td>5 (108,635)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent*</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.040 (0.030 to 0.050)</td>
</tr>
<tr>
<td></td>
<td>IGCS (≤13) vs. SMS (≤1): Diagnostic accuracy</td>
<td>1 (52,412)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Precise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 76% (74 to 77) vs. 73% (71 to 74); specificity (95% CI) 89% (89 to 89) vs. 91% (90 to 91)</td>
</tr>
<tr>
<td></td>
<td>mGCS vs. SMS: Discrimination</td>
<td>4 (56,223)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Difference in AUROC 0.000 (-0.007 to 0.007)</td>
</tr>
<tr>
<td>Key Question Outcome</td>
<td>Study Design Number of Studies (N)</td>
<td>Study Limitations</td>
<td>Directness</td>
<td>Consistency</td>
<td>Precision</td>
<td>Reporting Bias</td>
<td>Main Findings</td>
<td>Strength of Evidence Grade</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------------------</td>
<td>-------------------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------------</td>
<td>---------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td><strong>Trauma center need</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGCS vs. mGCS: Discrimination, diagnostic accuracy</td>
<td>2 (1,181,535)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>AUROC 0.62 vs. 0.61 and 0.64 vs. 0.60, sensitivity 30% vs. 27% and 28% vs. 25%, specificity 93% vs. 95% and 94% vs. 95%</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Severe injury (ISS &gt;15)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGCS vs. mGCS: Discrimination, diagnostic accuracy</td>
<td>2 (474,427)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>AUROC (95% CI) 0.720 (0.715 to 0.724) vs. 0.681 (0.677 to 0.686) and 0.648 (0.646 to 0.650) vs. 0.606 (0.605 to 0.608)</td>
<td>Low</td>
</tr>
<tr>
<td>IGCS vs. mGCS Diagnostic accuracy</td>
<td>1 (370,392)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Precise</td>
<td>Not detected</td>
<td>Sensitivity (95% CI) 31% (31 to 32) vs. 28% (28 to 28) and specificity (95% CI) 91% (91 to 91) vs. 93% (93 to 93)</td>
<td>Moderate</td>
</tr>
<tr>
<td><strong>KQ 1a. Effects of patient and assessment setting on comparative predictive utility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age: Discrimination</td>
<td>14 (810,600)</td>
<td>Moderate</td>
<td>Indirect</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Differences in the AUROC were similar in studies that enrolled children and those that enrolled mixed populations of adults and children</td>
<td>Low</td>
</tr>
<tr>
<td>Type of trauma: Discrimination</td>
<td>14 (810,600)</td>
<td>Moderate</td>
<td>Indirect</td>
<td>Consistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Differences in the AUROC were similar in studies that evaluated patients with TBI and those that enrolled mixed trauma patients</td>
<td>Low</td>
</tr>
<tr>
<td>Assessment setting: Discrimination</td>
<td>12 (797,826)</td>
<td>Moderate</td>
<td>Indirect</td>
<td>Inconsistent</td>
<td>Precise</td>
<td>Not detected</td>
<td>Differences in the AUROC were inconsistent in two studies that compared field and ED assessments. Differences in discrimination were similar in studies that used field versus ED scores.</td>
<td>Insufficient</td>
</tr>
<tr>
<td><strong>KQ 2. Under- and over-triage</strong></td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>Insufficient</td>
</tr>
<tr>
<td><strong>KQ 3. Clinical outcomes</strong></td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Key Question Outcome</td>
<td>Study Design</td>
<td>Number of Studies (N)</td>
<td>Study Limitations</td>
<td>Directness</td>
<td>Consistency</td>
<td>Precision</td>
<td>Reporting Bias</td>
<td>Main Findings</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>KQ 4. Reliability and ease of use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrater reliability</td>
<td></td>
<td>2 (135)</td>
<td>High</td>
<td>Indirect</td>
<td>Consistent</td>
<td>Imprecise</td>
<td>Not detected</td>
<td>The interrater reliability of the tGCS and mGCS appears to be high, but evidence was insufficient to determine if there were differences between scales</td>
</tr>
<tr>
<td>Ease of use: Proportion of correct scores using written or video patient scenarios</td>
<td></td>
<td>3 (498)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Imprecise</td>
<td>Not detected</td>
<td>Three studies found the tGCS associated with a lower proportion of correct scores than the mGCS (differences in proportion of correct scores ranged from 6% to 27%), though the difference was statistically significant in only one study</td>
</tr>
<tr>
<td>Ease of use: Effects of training on proportion of correct scores</td>
<td></td>
<td>3 (299)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Imprecise</td>
<td>Not detected</td>
<td>Three studies found that training or use of a scoring aid increased the proportion of correct scores on both the tGCS and mGCS (increase in proportion of correct scores ranged from 32% to 70%)</td>
</tr>
<tr>
<td>Key Question Outcome</td>
<td>Study Design</td>
<td>Directness</td>
<td>Consistency</td>
<td>Precision</td>
<td>Reporting Bias</td>
<td>Main Findings</td>
<td>Strength of Evidence Grade</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>Interrater reliability or ease of use</td>
<td>1 (3,052)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Unable to determine</td>
<td>Imprecise</td>
<td>Not detected</td>
<td>Evidence was insufficient to assess effects of patient, assessor, or setting on comparative interrater reliability of the tGCS versus the mGCS</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Injury severity: Effects on proportion of correct scores</td>
<td>3 (470)</td>
<td>Moderate</td>
<td>Direct</td>
<td>Consistent</td>
<td>Imprecise</td>
<td>Not detected</td>
<td>The proportion of correct GCS scores was generally lowest for assessment of patient scenarios with moderate injury severity in three studies, including one study that evaluated the tGCS and the mGCS</td>
<td>Low</td>
</tr>
</tbody>
</table>

AUROC=area under the receiver operating characteristic curve; ED= emergency department; ISS=injury severity score; KQ= Key Question; mGCS= motor Glasgow coma scale; n= number; SMS=simplified motor scale; TBI=traumatic brain injury; tGCS= total Glasgow Coma Scale; vs.= versus

*I-square >50% but range in differences in AUROC across studies <0.05
†I-square 0% in 4 studies of mixed populations of adults and children, pooled estimate similar to estimate in mixed populations, estimate higher in study of children but no statistically significant subgroup effect
# Appendix H. Head-to-Head Studies for Predictive Utility

Table H-1. Characteristics of head-to-head studies

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design</th>
<th>Eligibility Criteria</th>
<th>Population Characteristics</th>
<th>Setting and Dates Assessments Performed</th>
<th>N</th>
<th>Outcomes (Proportion with Outcome)</th>
</tr>
</thead>
</table>
| Acker, et al., 2014 | Retrospective cohort | Children ≤18 years old who were admitted to the hospital with a diagnosis of TBI and had complete tGCS and mGCS data available. | Age (mean, years): 6.9 (SD 5.8)  
Male: 65%  
Race: NR  
TBI: 100%  
ISS (median): 17 (IQR: 10-26)  
tGCS on presentation (median): 15 (IQR: 8-15)  
mGCS on presentation (median): 6 (IQR: 4-6)  
Cause of injury  
-Fall: 21%  
-MVC: 22%  
-NAT: 18%  
-Other: 39% | USA, Colorado Urban  
2 Level 1 pediatric trauma centers  
2002 to 2011 | 2,231 | Need for craniotomy (10.4%)  
Need for ICP monitoring (16.9%)  
Admission to the ICU (56.5%)  
Hospital stay of ≥5 days (30.4%)  
Discharge to rehabilitation (13.2%)  
Dependence on caretakers at followup (76.9%)  
Mortality (8.4%) |
| Al-Salamah, et al., 2004 | Retrospective analysis of prospective cohort | Patients who had an injury caused by any mechanism, ISS >12, transported by land ambulance, entered into the Ontario Trauma Registry Comprehensive Data set | Age (mean, years): 44 (SD 21)  
Male: 70%  
Race: NR  
Primary site of injury on arrival to ED  
-Head and neck: 32%  
-Chest and abdomen: 11%  
-Lower extremity: 3%  
-Upper extremity: 3%  
-Spine: 2%  
-Multiple sites: 36%  
-Unknown: 13%  
Endotracheal intubation before arrival to ED: 0.3%  
Required intubation in ED: 16% | Canada, Ontario Trauma registry  
72% urban, 28% suburban or rural  
1994 to 2002 | 795 | Mortality (18%)  
ICU admission (8%)  
Composite outcome of ICU admission or requiring intubation in the ED (NR) |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Glasgow Coma Scale Used</th>
<th>Personnel Performing Assessments and Where Assessed</th>
<th>Potential Confounders</th>
<th>Results: Univariate</th>
<th>Method for Constructing Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acker, et al., 2014</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>On presentation, but otherwise not described</td>
<td>Mentions univariate analysis was adjusted using the Bonferroni method for multiple comparisons, but adjustments not described and only goodness of fit data reported</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Al-Salamah, et al., 2004</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Trauma team, not otherwise described</td>
<td>Only diagnostic accuracy and discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------</td>
<td>--------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acker, et al., 2014</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Al-Salamah, et al., 2004 | NR | **Test characteristics (95% CI)* of mortality, tGCS (score ≤13) vs. mGCS (score ≤5)**  
Sensitivity: 80.28% (72.78 to 86.48) vs. 80.28% (72.78 to 86.48)  
Specificity: 67.99% (64.26 to 71.56) vs. 73.05% (69.47 to 76.42%)  
PLR: 2.51 (2.18 to 2.88) vs. 2.98 (2.56 to 3.46)  
NLR: 0.29 (0.21 to 0.41) vs. 0.27 (0.19 to 0.38)  
PPV: 35.29% (30.08 to 40.78) vs. 39.31% (33.65 to 45.19)  
NPV: 94.07% (91.54 to 96.02) vs. 94.46% (92.09 to 96.28) |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Discrimination or Calibration</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acker, et al., 2014</td>
<td>AUROC (95% CI), p-value, tGCS vs. mGCS</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>All ages (0-18 years)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Survived to hospital discharge: 0.949 (0.938 to 0.961) vs. 0.941 (0.926 to 0.957), p=0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Craniotomy: 0.642 (0.603 to 0.681) vs. 0.638 (0.601 to 0.675), p=0.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICU admission: 0.772 (0.754 to 0.790) vs. 0.721 (0.705 to 0.738), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LOS &gt;4 days: 0.683 (0.660 to 0.706) vs. 0.644 (0.622 to 0.666), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discharge to rehabilitation: 0.804 (0.782 to 0.826) vs. 0.766 (0.740 to 0.792), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dependent on caregiver: 0.757 (0.732 to 0.783) vs. 0.747 (0.722 to 0.772), p=0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICP monitoring: 0.808 (0.784 to 0.832) vs. 0.774 (0.748 to 0.800), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Youngest age group (0-3 years)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Survived to hospital discharge: 0.949 (0.934 to 0.964) vs. 0.936 (0.911 to 0.962), p=0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Craniotomy: 0.680 (0.617 to 0.743) vs. 0.659 (0.597 to 0.721), p=0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICU admission: 0.786 (0.758 to 0.814) vs. 0.723 (0.696 to 0.750), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LOS &gt;4 days: 0.630 (0.594 to 0.666) vs. 0.589 (0.555 to 0.623), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discharge to rehabilitation: 0.772 (0.732 to 0.811) vs. 0.713 (0.667 to 0.760), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dependent on caregiver: 0.808 (0.774 to 0.842) vs. 0.787 (0.752 to 0.821), p=0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICP monitoring: 0.728 (0.686 to 0.769) vs. 0.685 (0.643 to 0.726), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Al-Salamah, et al., 2004</td>
<td>Mortality, tGCS vs. mGCS</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>AUROC: 0.82 vs. 0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hosmer-Lemeshow Goodness of Fit p-value: &lt;0.01 vs. &lt;0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICU admission, tGCS vs. mGCS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p-value: 0.02 vs. 0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICU admission or required intubation in the ED, tGCS vs. mGCS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p-value: &lt;0.001 vs. &lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
<td>---------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Beskind, et al., 2014</td>
<td>Retrospective cohort</td>
<td>Trauma patients presenting to the ED via EMS at a level 1 trauma center</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Beskind, et al., 2014</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Out-of-hospital, otherwise not described</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Beskind, et al., 2014</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------------------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
</tbody>
</table>
| Beskind, et al., 2014 | **AUROC (95% CI). tGCS vs. mGCS**  
Survival to discharge: 0.899 (0.874 to 0.923) vs. 0.888 (0.864 to 0.913), mean difference=0.010 (0.002 to 0.018)  
Intubation in out-of-hospital setting or ED: 0.966 (0.955 to 0.976) vs. 0.948 (0.933 to 0.963), mean difference=0.018 (0.011 to 0.024)  
Neurosurgical intervention: 0.690 (0.661 to 0.718) vs. 0.671 (0.643 to 0.699), mean difference=0.019 (0.008 to 0.029)  
*Patients with ISS ≥16 (n=1,151)*  
Survival to discharge: 0.844 (0.815 to 0.874) vs. 0.837 (0.808 to 0.866), mean difference=0.008 (-0.001 to 0.018)  
Intubation in out-of-hospital setting or ED: 0.914 (0.895 to 0.932) vs. 0.905 (0.884 to 0.926), mean difference=0.009 (0.0001 to 0.017)  
Neurosurgical intervention: 0.571 (0.533 to 0.609) vs. 0.570 (0.531 to 0.608), mean difference=0.002 (-0.013 to 0.016)  
*Patients with head AIS ≥3 (n=1,165; TBI)*  
Survival to discharge: 0.869 (0.838 to 0.899) vs. 0.855 (0.824 to 0.886), mean difference=0.014 (0.005 to 0.023)  
Intubation in out-of-hospital setting or ED: 0.918 (0.899 to 0.937) vs. 0.907 (0.884 to 0.929), mean difference=0.012 (0.002 to 0.021)  
Neurosurgical intervention: 0.596 (0.558 to 0.635) vs. 0.602 (0.565 to 0.640), mean difference=-0.006 (-0.021 to 0.009) | Low           |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design</th>
<th>Eligibility Criteria</th>
<th>Population Characteristics</th>
<th>Setting and Dates Assessments Performed</th>
<th>N</th>
<th>Outcomes (Proportion with Outcome)</th>
</tr>
</thead>
</table>
Male: 66.1%  
Race: NR  
ISS (median): 9 (IQR: 4-13)  
Survival: 95.7%  
Trauma center need: 38.7%  
GCS score ≤13: 16.8%  
mGCS score ≤5: 14.2%  
SBP<90 mm Hg: 5.2%  
Respiratory rate <10 or >29: 6.3%  
Any step 1 criteria of the NTTP: 23%  
Penetrating injury: 11.6%  
Flail chest: 0.4%  
Open skull fracture: <0.1%  
≥2 long bone fractures: 1.3%  
Pelvic fracture: 6.3%  
Crush injury: 0.5%  
Amputation: 0.2%  
Paralysis: 0.4%  
Any step 2 criteria of the NTTP: 19.9%  
Any step 1 or 2 criteria of the NTTP: 46.5% | USA  
Trauma registry  
2007 to 2008 | 811,143 | Trauma center need (38.7%): ISS >15; ICU admission of ≥24 hours; need for urgent surgery (ED disposition to the OR); or death in the ED |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Glasgow Coma Scale Used</th>
<th>Personnel Performing Assessments and Where Assessed</th>
<th>Potential Confounders</th>
<th>Results: Univariate</th>
<th>Method for Constructing Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown, et al., 2014</td>
<td>tGCS vs. mGCS (from tGCs)</td>
<td>Out-of-hospital, otherwise not described</td>
<td>Adjusted for other triage criteria in the first 2 steps of the NTTP (SBP, respiratory rate, and anatomy of injury)</td>
<td>NR</td>
<td>Forward stepwise logistic regression</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Brown, et al., 2014 | **OR (95% CI)** tGCS (score ≤13) vs. mGCS (score ≤5)**  
|                   | All patients, with missing data imputed: 3.03 (2.94 to 3.13, p<0.01) vs. 3.37 (3.27 to 3.48, p<0.01)  
|                   | Only completed cases (59% of subjects had tGCS vs. 58% had mGCS present): 4.84 (4.40 to 4.57) vs. 4.87 (4.70 to 4.97) | **Need for trauma center, tGCS (score ≤13) vs. mGCS (score ≤5)**  
|                   |                                                                                      | GCS scores alone  
|                   |                                                                                      | Sensitivity: 30.3% vs. 26.7%  
|                   |                                                                                      | Specificity: 93.1% vs. 95.1%  
|                   |                                                                                      | Accuracy: 66.3% vs. 66.1%  
|                   |                                                                                      | **r^2**: 0.882 vs. 0.964  
|                   |                                                                                      | **GCS scores incorporated into the NTTP Step 1 and 2 criteria**  
|                   |                                                                                      | Sensitivity: 62.1% vs. 60.4%  
|                   |                                                                                      | Specificity: 65.7% vs. 67.1%  
<p>|                   |                                                                                      | Accuracy: 64.2% vs. 64.2% |</p>
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Discrimination or Calibration</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown, et al., 2014</td>
<td><strong>AUROC, tGCS vs. mGCS</strong>&lt;br&gt;GCS scores alone: 0.617 vs. 0.609, p&lt;0.01&lt;br&gt;GCS scores incorporated into NTTP Step 1 and 2 criteria: 0.639 vs. 0.637, p=0.10</td>
<td>Moderate</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>Retrospective cohort</td>
<td>Patients ≥16 years transported from the scene to a hospital by EMS, entered into the Ohio Trauma Registry, with complete EMS GCS scores</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>tGCS vs. SMS (from tGCS)</td>
<td>Out-of-hospital, obtained by EMS providers</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>NR</td>
<td><em><em>Test characteristics (95% CI)</em> tGCS ≤13 vs. SMS ≤1 vs. SMS 0</em>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Mortality</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 75.03% (73.45 to 76.56) vs. 72.20% (70.57 to 73.79) vs. 66.91% (65.20 to 68.58)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 87.63% (87.34 to 87.92) vs. 89.42% (89.14 to 89.69) vs. 93.80% (93.58 to 94.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 6.07 (5.88 to 6.26) vs. 6.82 (6.60 to 7.06) vs. 10.79 (10.34 to 11.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.28 (0.27 to 0.30) vs. 0.31 (0.29 to 0.33) vs. 0.35 (0.34 to 0.37)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 27.20% (26.25 to 28.17) vs. 29.59% (28.55 to 30.64) vs. 39.92% (38.57 to 41.28)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 98.28% (98.15 to 98.40) vs. 98.12% (97.99 to 98.25) vs. 97.87% (97.74 to 98.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TBI</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 45.40% (44.30 to 46.50) vs. 40.81% (39.72 to 41.89) vs. 30.12% (29.12 to 31.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 89.30% (89.01 to 89.59) vs. 90.50% (90.22 to 90.77) vs. 94.10% (93.88 to 94.32)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 4.24 (4.09 to 4.40) vs. 4.30 (4.13 to 4.47) vs. 5.11 (4.86 to 5.37)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.61 (0.60 to 0.62) vs. 0.65 (0.64 to 0.67) vs. 0.74 (0.73 to 0.75)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 43.20% (42.13 to 44.27) vs. 43.50% (42.38 to 44.64) vs. 47.79% (46.60 to 49.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 90.12% (89.84 to 90.40) vs. 89.51% (89.22 to 89.79) vs. 88.25% (87.96 to 88.54)</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>AUROC (95% CI), tGCS vs. SMS</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Non-parametric analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality: 0.85 (0.84 to 0.86) vs. 0.82 (0.81 to 0.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TBI: 0.72 (0.71 to 0.72) vs. 0.66 (0.65 to 0.66)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neurosurgical intervention: 0.75 (0.73 to 0.77) vs. 0.70 (0.68 to 0.72)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Any emergency intubation: 0.86 (0.85 to 0.87) vs. 0.83 (0.82 to 0.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ED intubation: 0.86 (0.86 to 0.87) vs. 0.83 (0.82 to 0.84)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parametric analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality: 0.87 (0.86 to 0.88) vs. 0.86 (0.85 to 0.88)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TBI: 0.80 (0.80 to 0.81) vs. 0.78 (0.76 to 0.80)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neurosurgical intervention: 0.82 (0.81 to 0.84) vs. 0.81 (0.78 to 0.84)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Any emergency intubation: 0.90 (0.90 to 0.91) vs. 0.91 (0.90 to 0.91)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ED intubation: 0.91 (0.90 to 0.91) vs. 0.91 (0.90 to 0.92)</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------------</td>
<td>--------------------------------</td>
</tr>
</tbody>
</table>
| Caterino and Raubenolt, 2012 Continued | **Neurosurgical intervention** | Sensitivity: 60.05% (56.53 to 63.50) vs. 52.93% (49.37 to 56.46) vs. 42.24% (38.76 to 45.78)  
Specificity: 84.70% (84.39 to 85.01) vs. 86.40% (86.10 to 86.69) vs. 90.70% (90.45 to 90.95)  
PLR: 3.92 (3.69 to 4.17) vs. 3.89 (3.63 to 4.17) vs. 4.54 (4.17 to 4.95)  
NLR: 0.47 (0.43 to 0.51) vs. 0.54 (0.51 to 0.59) vs. 0.64 (0.60 to 0.68)  
PPV: 5.64% (5.15 to 6.15) vs. 5.59% (5.08 to 6.14) vs. 6.47% (5.81 to 7.18)  
NPV: 99.29% (99.20 to 99.36) vs. 99.18% (99.09 to 99.26) vs. 99.04% (98.95 to 99.13) | |
| | **Any emergency intubation** | Sensitivity: 75.50% (74.13 to 76.83) vs. 72.71% (71.30 to 74.09) vs. 63.49% (61.98 to 64.99)  
Specificity: 88.90% (88.62 to 89.18) vs. 90.60% (90.34 to 90.86) vs. 94.70% (94.50 to 94.90)  
PLR: 6.80 (6.59 to 7.01) vs. 7.74 (7.48 to 8.00) vs. 11.98 (11.46 to 12.52)  
NLR: 0.28 (0.26 to 0.29) vs. 0.30 (0.29 to 0.32) vs. 0.39 (0.37 to 0.40)  
PPV: 35.87% (34.84 to 36.91) vs. 38.88% (37.77 to 40.00) vs. 49.63% (48.25 to 51.01)  
NPV: 97.78% (97.64 to 97.92) vs. 97.58% (97.44 to 97.72) vs. 96.93% (96.77 to 97.08) | |
| | **ED intubation** | Sensitivity: 76.89% (75.43 to 78.31) vs. 74.09% (72.57 to 75.57) vs. 64.61% (62.96 to 66.23)  
Specificity: 88.20% (87.91 to 88.48) vs. 89.83% (89.56 to 90.09) vs. 94.00% (93.79 to 94.21)  
PLR: 6.52 (6.32 to 6.72) vs. 7.28 (7.05 to 7.53) vs. 10.77 (10.32 to 11.24)  
NLR: 0.26 (0.25 to 0.28) vs. 0.29 (0.27 to 0.31) vs. 0.38 (0.36 to 0.39)  
PPV: 30.82% (29.83 to 31.82) vs. 33.22% (32.15 to 34.30) vs. 42.41% (41.05 to 43.78)  
NPV: 98.24% (98.11 to 98.36) vs. 98.07% (97.94 to 98.19) vs. 97.49% (97.35 to 97.63) |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design</th>
<th>Eligibility Criteria</th>
<th>Population Characteristics</th>
<th>Setting and Dates Assessments Performed</th>
<th>N</th>
<th>Outcomes (Proportion with Outcome)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cicero and Cross, 2013</td>
<td>Retrospective cohort</td>
<td>Patients in the NTDB data set from 2007-2009, ages &lt;19 years. Exclusion: interfacility transfers, ED LOS &gt;7 days or greater than the total recorded hospital LOS.</td>
<td>Age (mean, years): 12.6 (SD 5.5) Male: 67% Nonwhite race: 38% ED LOS (mean, minutes): 227 (SD 229) Hospital LOS (mean, days): 3.8 (SD 6.8) ISS (mean): 9.9 (SD 10.3)</td>
<td>USA Trauma registry 2007 to 2009</td>
<td>104,035</td>
<td>Mortality (3.8%) Death on arrival (NR): having a recorded ED disposition of death regardless of duration of resuscitation efforts Major injury (15%): having a recorded ISS &gt;15 ED LOS (NA): duration from arrival until disposition or death Hospital LOS (NA): duration of admission to any hospital inpatient service</td>
</tr>
<tr>
<td>Corrigan, et al., 2014</td>
<td>Retrospective cohort</td>
<td>Patients in the NTDB data set with a diagnosis of TBI, ages ≥18 years, were not transferred in from another hospital, did not die in the ED, with no missing data.</td>
<td>NR</td>
<td>USA Trauma registry 2007 to 2010</td>
<td>77,470</td>
<td>Days in the ICU (NA) Discharged alive (NR) LOS days (NA) Discharged home, if alive (NR)</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
<td>Potential Confounders</td>
<td>Results: Univariate</td>
<td>Method for Constructing Multivariate Model</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>---------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Cicero and Cross, 2013</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Out-of-hospital, otherwise not described</td>
<td>Only discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Corrigan, et al., 2014</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Out-of-hospital, otherwise not described</td>
<td>Only discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cicero and Cross, 2013</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrigan, <em>et al.</em>, 2014</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cicero and Cross, 2013</td>
<td>tGCS vs. mGCS</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUROC (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Overall mortality: 0.946 (0.941 to 0.951) vs. 0.940 (0.935 to 0.945)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Death on arrival: 0.958 (0.953 to 0.963) vs. 0.953 (0.948 to 0.959)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Major injury: 0.720 (0.715 to 0.724) vs. 0.681 (0.677 to 0.686)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Likelihood of surviving at arrival to ED (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tGCS=3: 0.71 (0.70 to 0.72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tGCS=15: 1 (1.0 to 1.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>LOS tGCS=3 vs. tGCS=14 or 15</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ED LOS (hours): 2 vs. 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hospital LOS (days): 8 vs. approximately 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrigan, et al., 2014</td>
<td>tGCS vs. mGCS</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICU days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AIC: 371699 vs. 373272</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R^2$: 0.1318 vs. 0.1140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discharged alive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AIC: 31456 vs. 32351</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SC: 31520.430 vs. 32416.138 c-index: 0.886 vs. 0.878</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LOS days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AIC: 461601 vs. 462758</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R^2$: 0.0956 vs. 0.0820</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discharged home (if alive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AIC: 71373 vs. 72631</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SC: 71437.519 vs. 72695.471 c-index: 0.763 vs. 0.750</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
<td>Population Characteristics</td>
<td>Setting and Dates Assessments Performed</td>
<td>N</td>
<td>Outcomes (Proportion with Outcome)</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
<td>----------------------------------------</td>
<td>---</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Davis, et al., 2006</td>
<td>Retrospective registry cohort</td>
<td>Adult patients with moderate-to-severe TBI (head/neck AIS ≥3) and available GCS scores. Exclusion: head/neck AIS was defined by a neck injury.</td>
<td>NR</td>
<td>USA, California (San Diego) Urban, other data NR</td>
<td>12,882</td>
<td>Mortality (NR) Neurosurgical intervention (NR): composite endpoint, which included mortality, craniotomy, invasive intracranial pressure monitoring, or ICU admission &gt;48 hours</td>
</tr>
<tr>
<td>Eken, et al., 2009</td>
<td>Prospective cohort</td>
<td>Patients &gt;17 years old with an altered level of consciousness, after any trauma to the head, neurological complaints of lateralizing motor, and/or sensory deficits, dysarthria, dysphasia, or facial asymmetry were eligible. Exclusion: patients who were intubated or administered sedative or paralytic agents before presentation to ED.</td>
<td>Age (median, years): 59 (range: 18-97) Male: 64% Race: NR</td>
<td>Turkey Tertiary care ED of hospital Level IV trauma center 2006</td>
<td>185</td>
<td>3-month mortality (25%) Hospital mortality (14%) 3-month morbidity using an MRS (39%)</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
<td>Potential Confounders</td>
<td>Results: Univariate</td>
<td>Method for Constructing Multivariate Model</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------</td>
<td>-----------------------------------------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Davis, et al., 2006</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>In-field and upon admission to ED, otherwise not described</td>
<td>NR</td>
<td>NR</td>
<td>Linear regression model adjusted for field GCS, otherwise not described.</td>
<td></td>
</tr>
<tr>
<td>Eken, et al., 2009</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>On presentation to ED, otherwise not described</td>
<td>Only discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------</td>
<td>--------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davis, <em>et al.</em>, 2006</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eken, <em>et al.</em>, 2009</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davis, et al., 2006</td>
<td><strong>Reported AUROC (optimized threshold value)</strong>&lt;br&gt;<strong>Mortality</strong>&lt;br&gt;Preadmission tGCS (field or arrival): 0.84 (5.016)&lt;br&gt;Preadmission mGCS: 0.83 (3.010)&lt;br&gt;Field tGCS: 0.84 (5.016)&lt;br&gt;Arrival tGCS: 0.84 (6.024)&lt;br&gt;<strong>Neurosurgical intervention</strong>&lt;br&gt;Preadmission tGCS (field or arrival): 0.80 (11.016)&lt;br&gt;Preadmission mGCS 0.78 (5.010)&lt;br&gt;Field tGCS: 0.80 (12.024)&lt;br&gt;Arrival tGCS: 0.83 (12.024)</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eken, et al., 2009</td>
<td><strong>Reported AUROC (95% CI) tGCS vs. mGCS</strong>&lt;br&gt;3-month mortality: 0.726 (0.656 to 0.789) vs. 0.679 (0.606 to 0.745)&lt;br&gt;Hospital mortality: 0.735 (0.655 to 0.797) vs. 0.662 (0.589 to 0.730)&lt;br&gt;Modified Rankin Scale 3-6, all patients: 0.720 (0.650 to 0.784) vs. 0.651 (0.578 to 0.720)&lt;br&gt;MRS 3-6, patients with trauma: 0.776 (0.657 to 0.869) vs. 0.706 (0.582 to 0.811)&lt;br&gt;MRS 3-6, patients without trauma: 0.655 (0.562 to 0.740) vs. 0.597 (0.503 to 0.686)</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
<td>Population Characteristics</td>
<td>Setting and Dates Assessments Performed</td>
<td>N</td>
<td>Outcomes (Proportion with Outcome)</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Gill, et al., 2005 | Retrospective cohort | Patients of all ages presenting to level 1 trauma center who met standard trauma alert criteria | Age (median, years): 24 (IQR: 15-38)  
Male: 71.5%  
Race: NR  
Trauma mechanism  
-MVC: 60.8%  
-Homicide and injury purposely inflicted by other people: 20.7%  
-Motor vehicle, nontraffic accidents: 3.8%  
-Other accidents: 3.0%  
-Suicide and self-inflicted injury: 1.9%  
-Other road vehicle accidents: 1.4% | USA, California (Loma Linda)  
Urban, University  
Level 1 trauma center and children's hospital  
1990 to 2002 | 8,432 | ED intubation (26.4%)  
Neurosurgical intervention (9.3%)  
Clinically significant brain injury (17.1%)  
Mortality (11.4%) |
| Gill, et al., 2006 | Retrospective cohort | Patients of all ages presenting to level 1 trauma center who met standard trauma alert criteria | Age (median, years): 24 (IQR: 16-38)  
Male: 70%  
Race: NR | USA, California (Loma Linda)  
Urban, University  
Level 1 trauma center and children's hospital  
1990 to 2002 | 7,233 | ED intubation (26%)  
Neurosurgical intervention (9%)  
Clinically significant brain injury (17%)  
Mortality (10%) |
| Haukoos, et al., 2007 | Retrospective cohort | All adult and pediatric patients who presented to the ED and were included in the trauma registry | Age (median, years): 32 (IQR: 21-45)  
Male: 71%  
Race: NR  
ISS score (median): 9 (IQR: 2-14)  
Trauma mechanism  
-MVC: 49%  
-Homicide and injury purposely inflicted by other people: 21%  
-Accidental falls: 17%  
-Other accidents: 5%  
-Suicide and self-inflicted injury: 2%  
-Other road vehicle crashes: 2%  
-Motor vehicle nontraffic crash: 1% | USA, Colorado  
Urban, Denver Health Medical Center  
Level 1 trauma center  
1995 to 2004 | 21,170 | Intubation, out-of-hospital or ED (18%)  
Brain injury (14%)  
Neurosurgical intervention (7%)  
Mortality (5%) |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Glasgow Coma Scale Used</th>
<th>Personnel Performing Assessments and Where Assessed</th>
<th>Potential Confounders</th>
<th>Results: Univariate</th>
<th>Method for Constructing Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gill, et al., 2005</td>
<td>tGCS vs. mGCS (from tGCS) vs. SMS (from tGCS)</td>
<td>Administered in ED by ED physicians</td>
<td>Only discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Gill, et al., 2006</td>
<td>tGCS vs. mGCS (from tGCS) and SMS (from tGCS)</td>
<td>Administered out-of-hospital, otherwise not described</td>
<td>Only discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Haukoos, et al., 2007</td>
<td>tGCS vs. mGCS (from tGCS) vs. SMS (from tGCS)</td>
<td>Administered in ED by ED physicians</td>
<td>Only discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill, et al., 2005</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill, et al., 2006</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haukoos, et al., 2007</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill, et al., 2005</td>
<td>Reported AUROC for tGCS vs. mGCS vs. SMS (CI’s not reported)</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ED intubation: 0.865 vs. 0.826 vs. 0.826</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neurosurgical intervention: 0.874 vs. 0.848 vs. 0.851</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brain injury: 0.826 vs. 0.789 vs. 0.791</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality: 0.906 vs. 0.894 vs. 0.878</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gill, et al., 2006</td>
<td>Reported AUROC (95% CI) for tGCS vs. mGCS vs. SMS</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ED intubation: 0.83 (0.81 to 0.84) vs. 0.79 (0.78 to 0.80) vs. 0.79 (0.77 to 0.80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neurosurgical intervention: 0.86 (0.85 to 0.88) vs. 0.84 (0.82 to 0.85) vs. 0.83 (0.81 to 0.84)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clinically significant brain injury (TBI): 0.83 (0.82 to 0.84) vs. 0.79 (0.78 to 0.81) vs. 0.79 (0.77 to 0.80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hospital mortality: 0.89 (0.88 to 0.90) vs. 0.88 (0.87 to 0.89) vs. 0.86 (0.86 to 0.89)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haukoos, et al., 2007</td>
<td>Reported AUROC (95% CI) for tGCS vs. mGCS vs. SMS</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intubation: 0.86 (0.85 to 0.87) vs. 0.81 (0.80 to 0.82) vs. 0.81 (0.80 to 0.82)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brain injury: 0.76 (0.75 to 0.77) vs. 0.71 (0.70 to 0.72) vs. 0.71 (0.70 to 0.72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neurosurgical intervention: 0.83 (0.82 to 0.84) vs. 0.80 (0.79 to 0.81) vs. 0.80 (0.79 to 0.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality: 0.92 (0.91 to 0.93) vs. 0.90 (0.89 to 0.91) vs. 0.89 (0.88 to 0.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
<td>Population Characteristics</td>
<td>Setting and Dates Assessments Performed</td>
<td>N</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Healey, et al., 2003</td>
<td>Retrospective cohort</td>
<td>Patients in the NTDB data set with complete GCS data.</td>
<td>Age: NR</td>
<td>USA Trauma registry 1994 to 2001</td>
<td>202,255</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Male: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Race: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GCS score=15: 80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holmes, et al., 2006</td>
<td>Prospective cohort</td>
<td>Pediatric patients &lt;18 years with blunt head trauma presenting to the ED. Exclusion: children with trivial head trauma defined by falls from ground level or trauma resulting from walking or running into stationary objects if the only abnormal finding was a scalp laceration or abrasion, and children transferred who had undergone CT scanning before transfer.</td>
<td>Ages ≤2 years: 16% Ages &gt;2 years: 84% Male: NR Race: NR</td>
<td>USA, California (Davis) Level 1 trauma center 1998 to 2001</td>
<td>2,043</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
<td>Potential Confounders</td>
<td>Results: Univariate</td>
<td>Method for Constructing Multivariate Model</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------</td>
<td>-----------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Healey, et al., 2003</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Out-of-hospital, otherwise not described</td>
<td>Only diagnostic accuracy and discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Holmes, et al., 2006</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Presentation to ED, otherwise not described</td>
<td>Only discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healey, et al., 2003</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holmes, et al., 2005</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healey, et al., 2003</td>
<td><strong>tGCS vs. mGCS</strong>&lt;br&gt;AUROC (95% CI): 0.891 (0.888 to 0.894) vs. 0.873 (0.870 to 0.875), p=0.000&lt;br&gt;Misclassification: 4.9% vs. 5.1%</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holmes, et al., 2005</td>
<td><strong>Reported AUROC (95% CI) tGCS vs. mGCS</strong>&lt;br&gt;<strong>TBI on cranial CT scan</strong>&lt;br&gt;Ages ≤2 years: 0.72 (0.56 to 0.87) vs. 0.60 (0.48 to 0.72)&lt;br&gt;Ages &gt;2 years: 0.82 (0.76 to 0.87) vs. 0.71 (0.65 to 0.77)&lt;br&gt;<strong>TBI in need of acute intervention</strong>&lt;br&gt;Ages ≤2 years: 0.97 (0.94 to 1.0) vs. 0.76 (0.59 to 0.93)&lt;br&gt;Ages &gt;2 years 0.87 (0.83 to 0.92) vs. 0.76 (0.71 to 0.81)</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
<td>Population Characteristics</td>
<td>Setting and Dates Assessments Performed</td>
<td>N</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Kupas et al., 2016</td>
<td>Retrospective cohort</td>
<td>Adults, age ≥18 years, with complete data available in the Pennsylvania Trauma System Foundation's registry.</td>
<td>Age (median, years): 50 Male: 62.2% White: 78.6% Black: 13.9% Asian: 0.9% SBP ≥90 mm Hg: 95.6% GCS score of 15: 74.8% Injury type - Blunt: 87.9% - Penetrating: 8.4% - Burn: 3.7%</td>
<td>USA, Pennsylvania Level I, II, III, or IV trauma centers 1999 to 2013</td>
<td>370,392</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
<td>Potential Confounders</td>
<td>Results: Univariate</td>
<td>Method for Constructing Multivariate Model</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Kupas <em>et al</em>., 2016</td>
<td>tGCS vs. mGCS (from IGCS)</td>
<td>Out-of-hospital, otherwise not described</td>
<td>Only diagnostic accuracy and discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kupas et al., 2016</td>
<td>NR</td>
<td>Test characteristics (95% CI) tGCS (score ≤13) vs. mGCS (score ≤5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mortality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 69.8% (69.2 to 70.4) vs. 67.3% (66.7 to 67.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 88.1% (88.0 to 88.2) vs. 90.1% (90.0 to 90.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 12.88 (12.52 to 13.25) vs. 13.551 (13.183 to 13.920)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.76 (0.75 to 0.76) vs. 0.726 (0.722 to 0.730)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ISS score &gt;15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 31.3% (31.0 to 31.6) vs. 28.0% (27.7 to 28.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 91.3% (91.2 to 91.4) vs. 92.8% (92.7 to 92.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 2.54 (2.52 to 2.56) vs. 2.56 (2.54 to 2.59)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.53 (0.53 to 0.54) vs. 0.51 (0.51 to 0.52)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ICU admission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 27.3% (27.1 to 27.5) vs. 23.9% (23.7 to 24.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 91.6% (91.4 to 91.8) vs. 92.7% (92.6 to 92.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 1.52 (1.52 to 1.53) vs. 1.51 (1.50 to 1.51)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.37 (0.37 to 0.38) vs. 0.37 (0.37 to 0.38)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intubation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 83.7% (83.3 to 84.2) vs. 81.3% (80.9 to 81.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 90.0% (89.9 to 90.1) vs. 92.0% (91.9 to 92.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 28.62 (27.72 to 29.53) vs. 28.70 (27.85 to 29.55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.62 (0.62 to 0.62) vs. 0.57 (0.57 to 0.58)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trauma center need</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 28.2% (27.9 to 28.4) vs. 25.1% (24.9 to 25.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 93.7% (93.6 to 93.8) vs. 95.0% (94.9 to 95.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 2.19 (2.17 to 2.20) vs. 2.21 (2.19 to 2.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.38 (0.37 to 0.38) vs. 0.35 (0.34 to 0.35)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 33.5% (33.0 to 34.0) vs. 30.5% (30.0 to 31.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 86.5% (86.4 to 86.6) vs. 88.4% (88.3 to 88.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 2.81 (2.75 to 2.87) vs. 2.89 (2.83 to 2.96)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.87 (0.87 to 0.88) vs. 0.86 (0.86 to 0.87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Craniotomy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 51.4% (50.2 to 52.5) vs. 46.5% (45.4 to 47.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 85.9% (85.8 to 86.0) vs. 87.8% (87.7 to 87.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 6.03 (5.76 to 6.30) vs. 5.88 (5.61 to 6.12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.94 (0.94 to 0.94) vs. 0.94 (0.94 to 0.94)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------------------------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Kupas et al., 2016 | Reported AUROC (95% CI) tGCS (score ≤13) vs. mGCS (score ≤5)  
ISS score of >15: 0.648 (0.646 to 0.650) vs. 0.606 (0.605 to 0.608), difference=0.042 (0.041 to 0.043)  
ISS score of >24: 0.719 (0.716 to 0.721) vs. 0.680 (0.677 to 0.682), difference=0.039 (0.038 to 0.041)  
Mortality: 0.831 (0.828 to 0.834) vs. 0.803 (0.800 to 0.806), difference=0.028 (0.026 to 0.030)  
ICU admission: 0.625 (0.623 to 0.626) vs. 0.583 (0.581 to 0.584), difference=0.042 (0.041 to 0.043)  
Intubation: 0.904 (0.902 to 0.907) vs. 0.884 (0.882 to 0.887), difference=0.020 (0.019 to 0.021)  
Trauma center need: 0.641 (0.639 to 0.642) vs. 0.603 (0.602 to 0.604), difference=0.038 (0.037 to 0.039)  
Surgery: 0.612 (0.608 to 0.615) vs. 0.597 (0.595 to 0.600), difference=0.014 (0.013 to 0.16)  
Craniotomy: 0.724 (0.718 to 0.730) vs. 0.676 (0.670 to 0.682), difference=0.048 (0.044 to 0.052) | Moderate |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design</th>
<th>Eligibility Criteria</th>
<th>Population Characteristics</th>
<th>Setting and Dates Assessments Performed</th>
<th>N</th>
<th>Outcomes (Proportion with Outcome)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross, <em>et al.</em>, 1998</td>
<td>Retrospective cohort</td>
<td>All patients ≥13 years transported directly to the trauma center. Exclusion: patients seen initially at another hospital and transferred to the trauma center.</td>
<td>Age (mean, years): 37.1 (range: 13-95) Male: 69% Race: NR Airway intubation in the field: 3.5% Blunt mechanism of injury: 85% ISS (mean): 14.4 ISS (median): 13 No head injury: 43.8% AIS≤2 (concussion): 25% AIS=3: 16.3%</td>
<td>USA, New Jersey Level 1 trauma center 1994 to 1996</td>
<td>1,410</td>
<td>Severe head injury (14.8%): AIS ≥4 or AIS=5 Mortality (6%)</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
<td>Potential Confounders</td>
<td>Results: Univariate</td>
<td>Method for Constructing Multivariate Model</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------</td>
<td>----------------------</td>
<td>--------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Ross, et al., 1998</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Out-of-hospital, otherwise not described</td>
<td>Effect of shock on neurologic status (patients, n=3, with SBP&lt;90 mm Hg)</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross, et al., 1998</td>
<td>NR</td>
<td>Test characteristics (95% CI)* tGCS (score ≤13) vs. mGCS (score ≤5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIS score =5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 93.33% (83.80 to 98.15) vs. 90.16% (79.81 to 96.30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 84.51% (82.46 to 86.40) vs. 85.40% (83.40 to 87.24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 6.02 (5.23 to 6.94) vs. 6.17 (5.30 to 7.20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.08 (0.03 to 0.20) vs. 0.12 (0.05 to 0.25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 21.13% (16.38 to 26.55) vs. 21.83% (16.89 to 27.44)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 99.65% (99.11 to 99.90) vs. 99.48% (98.88 to 99.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AIS score ≥4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 61.72% (54.76 to 68.34) vs. 60.77% (53.79 to 67.43)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 85.47% (83.05 to 87.67) vs. 89.59% (87.73 to 91.26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 4.25 (3.52 to 5.13) vs. 5.84 (4.79 to 7.12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.45 (0.38 to 0.53) vs. 0.44 (0.37 to 0.52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 48.68% (42.52 to 54.87) vs. 50.40% (44.05 to 56.73)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 90.91% (88.81 to 92.73) vs. 92.92% (91.29 to 94.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 71.28% (61.02 to 80.14) vs. 72.34% (62.15 to 81.07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 84.95% (82.91 to 86.84%) vs. 86.02% (84.03 to 87.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 4.74 (3.95 to 5.68) vs. 5.17 (4.31 to 6.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.45 (0.25 to 0.47) vs. 0.32 (0.23 to 0.45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 25.28% (20.16 to 30.96) vs. 26.98% (21.61 to 32.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 97.64% (96.59 to 98.44) vs. 97.75% (96.73 to 98.53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Craniotomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 63.16% (38.36 to 83.71) vs. 68.42% (43.45 to 87.42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 81.81% (79.68 to 83.81) vs. 82.82% (80.73 to 84.77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 3.47 (2.42 to 4.98) vs. 3.98 (2.87 to 5.52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.45 (0.25 to 0.81) vs. 0.38 (0.20 to 0.74)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 4.53% (2.36 to 7.78) vs. 5.16% (2.78 to 8.66)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 99.39% (98.74 to 99.75) vs. 99.48% (98.88 to 99.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross, et al., 1998</td>
<td>NR</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
<td>Population Characteristics</td>
<td>Setting and Dates Assessments Performed</td>
<td>N</td>
<td>Outcomes (Proportion with Outcome)</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------</td>
<td>---</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Thompson, et al., 2011</td>
<td>Retrospective cohort</td>
<td>All adult and pediatric patients who presented to the ED and were included in the trauma registry</td>
<td>Age (median, years): 33 (IQR: 22-48) Male: 71% Race: NR ISS (median): 9 (IQR: 4-17) Out-of-hospital GCS score (median): 15 (IQR: 14-15) Mechanism of injury -Blunt: 81% -Penetrating, stab: 7% -Penetrating, gunshot: 6% -Other: 6%</td>
<td>USA, Colorado Urban, Denver Health Medical Center Level 1 trauma center 1999 to 2008</td>
<td>19,408</td>
<td>Emergency tracheal intubation (18%) Clinically meaningful brain injury (18%) Need for neurosurgical intervention (8%) Mortality (6%)</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
<td>Potential Confounders</td>
<td>Results: Univariate</td>
<td>Method for Constructing Multivariate Model</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Thompson, et al., 2011</td>
<td>tGCS vs. mGCS (from tGCS) vs. SMS (from tGCS)</td>
<td>Out-of-hospital, otherwise not described</td>
<td>Only diagnostic accuracy and discrimination reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thompson, et al., 2011</td>
<td>NR</td>
<td><strong>Test characteristics (95% CI) SMS=0 vs. SMS≤1</strong>  &lt;br&gt;Primary analysis, with missing GCS data multiply imputed  &lt;br&gt;Emergency tracheal intubation  &lt;br&gt;Sensitivity: 0.61 (0.56 to 0.66) vs. 0.67 (0.62 to 0.72)  &lt;br&gt;Specificity: 0.65 (0.58 to 0.73) vs. 0.62 (0.56 to 0.68)  &lt;br&gt;PLR: 1.75 (1.47 to 2.09) vs. 1.79 (1.58 to 2.03)  &lt;br&gt;NLR: 0.60 (0.55 to 0.66) vs. 0.52 (0.47 to 0.58)  &lt;br&gt;Brain injury  &lt;br&gt;Sensitivity: 0.55 (0.51 to 0.59) vs. 0.62 (0.59 to 0.65)  &lt;br&gt;Specificity: 0.64 (0.56 to 0.71) vs. 0.61 (0.54 to 0.67)  &lt;br&gt;PLR: 1.53 (1.30 to 1.81) vs. 1.59 (1.38 to 1.83)  &lt;br&gt;NLR: 0.70 (0.66 to 0.75) vs. 0.63 (0.58 to 0.67)  &lt;br&gt;Neurosurgical intervention  &lt;br&gt;Sensitivity: 0.66 (0.62 to 0.70) vs. 0.74 (0.70 to 0.77)  &lt;br&gt;Specificity: 0.63 (0.55 to 0.70) vs. 0.59 (0.53 to 0.65)  &lt;br&gt;PLR: 1.78 (1.52 to 2.08) vs. 1.82 (1.60 to 2.07)  &lt;br&gt;NLR: 0.54 (0.50 to 0.59) vs. 0.44 (0.40 to 0.49)  &lt;br&gt;Mortality  &lt;br&gt;Sensitivity: 0.83 (0.75 to 0.91) vs. 0.86 (0.78 to 0.94)  &lt;br&gt;Specificity: 0.63 (0.56 to 0.70) vs. 0.59 (0.54 to 0.65)  &lt;br&gt;PLR: 2.25 (1.89 to 2.68) vs. 2.13 (1.92 to 2.37)  &lt;br&gt;NLR: 0.27 (0.19 to 0.37) vs. 0.23 (0.15 to 0.34)  &lt;br&gt;Composite outcome (any one of the outcomes)  &lt;br&gt;Sensitivity: 0.53 (0.49 to 0.57) vs. 0.59 (0.55 to 0.63)  &lt;br&gt;Specificity: 0.66 (0.58 to 0.74) vs. 0.64 (0.57 to 0.70)  &lt;br&gt;PLR: 1.57 (1.30 to 1.89) vs. 1.63 (1.41 to 1.88)  &lt;br&gt;NLR: 0.71 (0.67 to 0.76) vs. 0.64 (0.60 to 0.69)  &lt;br&gt;Sensitivity analysis, with missing GCS data excluded  &lt;br&gt;Emergency tracheal intubation  &lt;br&gt;Sensitivity: 0.63 (0.62 to 0.65) vs. 0.69 (0.67 to 0.70)  &lt;br&gt;Specificity: 0.61 (0.60 to 0.62) vs. 0.60 (0.59 to 0.61)  &lt;br&gt;PLR: 1.62 (1.57 to 1.68) vs. 1.71 (1.66 to 1.76)  &lt;br&gt;NLR: 0.60 (0.58 to 0.63) vs. 0.52 (0.50 to 0.55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Thompson, *et al.*, 2011 | **Reported AUROC (95% CI)** tGCS vs. mGCS vs. SMS  
*Primary analysis, with missing GCS data multiply imputed*  
Emergency tracheal intubation: 0.70 (0.63 to 0.77) vs. 0.65 (0.60 to 0.70) vs. 0.65 (0.62 to 0.67)  
Brain injury: 0.66 (0.60 to 0.71) vs. 0.61 (0.57 to 0.65) vs. 0.61 (0.58 to 0.64)  
Neurosurgical intervention: 0.70 (0.64 to 0.77) vs. 0.66 (0.61 to 0.71) vs. 0.66 (0.64 to 0.69)  
Mortality: 0.82 (0.74 to 0.90) vs. 0.76 (0.70 to 0.83) vs. 0.74 (0.70 to 0.77)  
Composite (any one of the outcomes): 0.66 (0.60 to 0.72) vs. 0.61 (0.57 to 0.66) vs. 0.61 (0.58 to 0.64)  
*Sensitivity analysis, with missing GCS data excluded*  
Emergency tracheal intubation: 0.80 (0.79 to 0.81) vs. 0.77 (0.76 to 0.78) vs. 0.77 (0.76 to 0.78)  
Brain injury: 0.75 (0.74 to 0.76) vs. 0.70 (0.69 to 0.71) vs. 0.70 (0.69 to 0.71)  
Neurosurgical intervention: 0.79 (0.78 to 0.81) vs. 0.77 (0.75 to 0.78) vs. 0.77 (0.76 to 0.78)  
Mortality: 0.90 (0.89 to 0.91) vs. 0.88 (0.87 to 0.89) vs. 0.87 (0.86 to 0.88)  
Composite (any one of the outcomes): 0.77 (0.76 to 0.78) vs. 0.72 (0.72 to 0.73) vs. 0.72 (0.71 to 0.73) | Moderate |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Results: Multivariate</th>
<th>Measures of Diagnostic Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thompson, et al. 2011</td>
<td><strong>Brain injury</strong>&lt;br&gt;Sensitivity: 0.57 (0.56 to 0.59) vs. 0.63 (0.61 to 0.65)&lt;br&gt;Specificity: 0.60 (0.59 to 0.60) vs. 0.58 (0.58 to 0.59)&lt;br&gt;PLR: 1.42 (1.37 to 1.47) vs. 1.51 (1.46 to 1.55)&lt;br&gt;NLR: 0.72 (0.69 to 0.75) vs. 0.64 (0.61 to 0.67)&lt;br&gt;<strong>Neurosurgical intervention</strong>&lt;br&gt;Sensitivity: 0.68 (0.66 to 0.70) vs. 0.75 (0.73 to 0.77)&lt;br&gt;Specificity: 0.59 (0.58 to 0.60) vs. 0.57 (0.56 to 0.58)&lt;br&gt;PLR: 1.65 (1.59 to 1.72) vs. 1.74 (1.68 to 1.80)&lt;br&gt;NLR: 0.54 (0.50 to 0.59) vs. 0.44 (0.40 to 0.48)&lt;br&gt;<strong>Mortality</strong>&lt;br&gt;Sensitivity: 0.85 (0.83 to 0.87) vs. 0.88 (0.86 to 0.90)&lt;br&gt;Specificity: 0.59 (0.59 to 0.60) vs. 0.57 (0.56 to 0.58)&lt;br&gt;PLR: 2.08 (2.02 to 2.14) vs. 2.04 (1.99 to 2.10)&lt;br&gt;NLR: 0.26 (0.23 to 0.30) vs. 0.22 (0.18 to 0.25)&lt;br&gt;<strong>Composite outcome (any one of the outcomes)</strong>&lt;br&gt;Sensitivity: 0.55 (0.54 to 0.57) vs. 0.61 (0.59 to 0.62)&lt;br&gt;Specificity: 0.62 (0.61 to 0.63) vs. 0.61 (0.60 to 0.62)&lt;br&gt;PLR: 1.44 (1.39 to 1.48) vs. 1.54 (1.50 to 1.59)&lt;br&gt;NLR: 0.73 (0.70 to 0.75) vs. 0.65 (0.63 to 0.67)</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Van de Voorde, et al., 2008</td>
<td>Prospective cohort</td>
<td>TBI patients (defined as LOS in hospital &gt;48 hours or death, and any brain AIS'90 score) ages 0-18 years admitted in 2005 to 1 of 18 participating hospitals. Excluded if had a high AIS'90 score in any other body region that was thought to contribute significantly to outcome, if they were ictal or postictal on first GCS assessment of if data collection was insufficient.</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>Van de Voorde, et al., 2008</td>
<td>tGCS vs. mGCS (from tGCS)</td>
<td>Best GCS on scene, or upon ED admission if no pre-hospital intervention</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Van de Voorde, et al., 2008</td>
<td>NR</td>
<td>Test characteristics (95% CI)* of mortality, sensitivity and specificity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tGCS score &lt;15: 100% (69.15 to 100) vs. 56.10% (44.70 to 67.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tGCS score &lt;14: 100% (69.15 to 100) vs. 70.73% (59.65 to 80.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tGCS score &lt;13: 100% (69.15 to 100) vs. 74.39% (63.56 to 83.40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mGCS score &lt;6: 100% (69.15 to 100) vs. 74.36% (63.21 to 83.58)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mGCS score &lt;5: 100% (69.15 to 100.0) vs. 85.90% (76.17 to 92.74)</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Van de Voorde, et al., 2008</td>
<td>NR</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Please see Appendix C. Included Studies for full study references.

AIC= Akaike information criterion; AIS= Abbreviated Injury Scale; AUROC= area under the receiver operating characteristic curve; BP= blood pressure; CI= confidence interval; CT= computed tomography; ED= emergency department; EMS= emergency medical services; GCS= Glasgow Coma Scale; ICP= intracranial pressure; ICU= intensive care unit; IQR= interquartile range; ISS= injury severity score; LOS= length of stay; mGCS= motor scale of GCS; MRS= Modified Rankin Scale; MVC= motor vehicle crash; N= number; NA= not available; NAT= nonaccidental trauma; NLR= negative likelihood ratio; NPV= negative predictive value; NR= not reported; NTDB= National Trauma Data Bank; NTTP= National Trauma Triage Protocol; OR= odds ratio; PENTA= pediatric trauma registry; PLR= positive likelihood ratio; PPV= positive predictive value; RR= relative risk; SBP= systolic blood pressure; SC= Schwartz criterion; SD= standard deviation; SMS= 3-point simplified motor score; TBI= traumatic brain injury; tGCS= total GCS; vs.= versus

*Calculated
# Appendix I. Indirect Studies for Predictive Utility

## Table I-1. Characteristics of indirect predictive utility studies

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design</th>
<th>Eligibility Criteria</th>
<th>Population Characteristics</th>
<th>Setting and Dates Assessments Performed</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caterino, <em>et al.</em>, 2011</td>
<td>Retrospective cohort</td>
<td>Patients ≥16 years transported from the scene to a hospital by EMS, entered into the Ohio Trauma Registry, with complete EMS GCS scores</td>
<td>Age ≥ 70 years: 30% (n=15,708) Male: 56% White: 80% Nonwhite: 13.5% Hispanic: 1.4% ISS score &lt;15 (mild): 65% ISS score &gt;15 (moderate, severe): 26.6% Survived to hospital discharge: 94.2% Intubated in out-of-hospital setting or ED: 7.6% Trauma type was TBI:15.2% EMS GCS score &lt;13: 12.8% Initial EMS BP &lt;90 mmHg: 5.2% Initial EMS BP ≥90 mmHg: 92.6% Neurosurgical intervention: 1.5%</td>
<td>USA, Ohio Urban and rural settings Ohio Trauma Registry 2002 to 2007</td>
<td>52,412</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Outcomes</td>
<td>Glasgow Coma Scale Used</td>
<td>Personnel Performing Assessments and Where Assessed</td>
<td>Potential Confounders</td>
<td>Results: Univariate</td>
</tr>
<tr>
<td>--------------</td>
<td>----------</td>
<td>-------------------------</td>
<td>--------------------------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Caterino, et al., 2011</td>
<td>Mortality (in hospital) Clinical brain injury (skull fractures with underlying brain injury, intracranial hemorrhage, cerebral contusion, or nonspecific intracranial injury) Neurosurgical intervention (operations on the brain, skull, or meninges, including diagnostic and therapeutic procedures such as shunts, craniotomies, and ventriculostomies) Emergency intubation</td>
<td>tGCS</td>
<td>Out-of-hospital, obtained by EMS providers</td>
<td>Under estimation of confidence levels</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adults ≥70 years of age vs. &lt;70 years of age</strong></td>
<td><strong>Test characteristics (95% CI) of tGCS score 13 elders vs. adults</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported OR (95% CI)</td>
<td><strong>Mortality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality, tGCS 14 vs. 15: 1.40 (1.07 to 1.83) vs. 1.22 (0.88 to 1.71)</td>
<td>Sensitivity: 50.7% (47.5 to 53.9) vs. 85.7% (84.1 to 87.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality, tGCS 13 vs. 14: 2.34 (1.57 to 3.52) vs. 1.45 (0.91 to 2.30)</td>
<td>Specificity: 93.8% (93.4 to 94.2) vs. 85.0% (84.6 to 85.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBI, tGCS 14 vs. 15: 2.50 (2.06 to 3.02) vs. 2.51 (2.24 to 2.81)</td>
<td>PLR: 8.20 (7.51 to 8.96) vs. 5.72 (5.55 to 5.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBI, tGCS 13 vs. 14: 1.00 (0.71 to 1.43) vs. 1.11 (0.93 to 1.33)</td>
<td>NLR: 0.52 (0.49 to 0.56) vs. 0.17 (0.15 to 0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurosurgical intervention, tGCS 14 vs. 15: 0.67 (0.38 to 1.20) vs. 2.02 (1.45 to 2.80)</td>
<td><strong>TBI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurosurgical intervention, tGCS 13 vs. 14: 2.41 (1.05 to 5.55) vs. 1.59 (1.03 to 2.42)</td>
<td>Sensitivity: 27.5% (25.7 to 29.3) vs. 53.0% (51.6 to 54.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intubation, tGCS 14 vs. 15: 2.22 (1.59 to 3.10) vs. 3.12 (2.60 to 3.74)</td>
<td>Specificity: 94.3% (93.9 to 94.7) vs. 87.1% (86.8 to 87.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intubation, tGCS 13 vs. 14: 1.16 (0.63 to 2.12) vs. 1.50 (1.18 to 1.92)</td>
<td>PLR: 4.85 (4.41 to 5.34) vs. 3.26 (3.16 to 3.35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reported OR (95% CI), tGCS score ≤14 in adults ≥70 years of age vs. ≤13 in adults &lt;70 years of age</strong></td>
<td>NLR: 0.77 (0.75 to 0.79) vs. 0.44 (0.42 to 0.45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality: 4.68 (2.90 to 7.54)</td>
<td><strong>Neurosurgical intervention</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBI: 1.84 (1.45 to 2.34)</td>
<td>Sensitivity: 42.7% (35.7 to 49.9) vs. 65.9% (61.9 to 69.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurosurgical intervention: 0.39 (0.20 to 0.78)</td>
<td>Specificity: 91.5% (91.0 to 91.9) vs. 81.8% (81.4 to 82.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intubation: 0.38 (0.26 to 0.56)</td>
<td>PLR: 5.02 (4.24 to 5.94) vs. 3.61 (3.40 to 3.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test characteristics (95% CI) of tGCS score &lt;14 elders vs. adults</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td>Sensitivity: 59.2% (56.1 to 62.3) vs. 88.2 (86.7 to 89.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specificity: 85.1% (84.6 to 85.7) vs. 76.9% (76.4 to 77.3)</td>
<td>PLR: 3.99 (3.74 to 4.26) vs. 3.81 (3.72 to 3.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLR: 0.48 (0.44 to 0.52) vs. 0.15 (0.14 to 0.17)</td>
<td><strong>TBI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity: 42.7% (40.7 to 44.7) vs. 65.0% (63.7 to 66.3)</td>
<td>Specificity: 86.8% (86.3 to 87.4) vs. 80.0% (79.6 to 80.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLR: 3.24 (3.04 to 3.46) vs. 3.26 (3.16 to 3.35)</td>
<td>NLR: 0.66 (0.64 to 0.68) vs. 0.44 (0.42 to 0.45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caterino, et al., 2011</td>
<td>NR</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
<td>Population Characteristics</td>
<td>Setting and Dates Assessments Performed</td>
<td>N</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>Caterino, et al., 2011</td>
<td>Continued</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Johnson and Krishnamurthy, 1996 | Retrospective cohort | Children seen by the neurosurgical service at the Children's Hospital in Washington D.C. during 1985 to 1988. | Age (mean, years): 7.33 (SD 5.08)  
Male: NR  
Race: NR  
GCS (mean): 13.37 (SD 3.32)  
GCS score ≤8: 6.7%  
ISS (mean): 10.22 (SD 9.93)  
TS (mean): 14.31 (SD 2.75)  
Revised TS (mean): 7.17 (SD 1.46)  
Arrived intubated: 0.9%  
Mortality: 1.9%  
PICU LOS (mean, days): 3.71 (SD 7.85)  
*Mechanism of injury*  
-MVA: 45.7%  
-Falls: 32.2%  
-Abuse: 2.5%  
-Assault: 8.3%  
-Struck by object: 7.6  
-Other: 3.7% | USA, Washington D.C.  
Urban, Children's hospital  
Level 1 trauma center  
1985 to 1988 | 841 |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Outcomes</th>
<th>Glasgow Coma Scale Used</th>
<th>Personnel Performing Assessments and Where Assessed</th>
<th>Potential Confounders</th>
<th>Results: Univariate</th>
<th>Method for Constructing Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caterino, <em>et al.</em>, 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continued</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson and Krishnamurthy, 1996</td>
<td>Mortality</td>
<td>TGC5</td>
<td>Senior neurosurgical resident performed assessment within 30 minutes of arrival to trauma center</td>
<td>Only diagnostic accuracy reported; no adjustment performed</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------</td>
<td>--------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caterino, et al., 2011 Continued</td>
<td><strong>Neurosurgical intervention</strong>&lt;br&gt;Sensitivity: 49.7% (42.6 to 56.9) vs. 74.9% (71.2 to 78.4)&lt;br&gt;Specificity: 82.8% (82.2 to 83.4) vs. 73.9% (73.5 to 74.4)&lt;br&gt;PLR: 2.89 (2.51 to 3.34) vs. 2.88 (2.74 to 3.02)&lt;br&gt;NLR: 0.61 (0.53 to 0.70) vs. 0.34 (0.30 to 0.39)&lt;br&gt;<strong>Intubation</strong>&lt;br&gt;Sensitivity: 66.3% (62.2 to 70.1) vs. 84.0% (82.7 to 85.2)&lt;br&gt;Specificity: 84.2% (83.6 to 84.8) vs. 79.0% (78.6 to 79.5)&lt;br&gt;PLR: 4.20 (3.9 to 4.5) vs. 4.01 (3.91 to 4.11)&lt;br&gt;NLR: 0.40 (0.36 to 0.45) vs. 0.20 (0.19 to 2.20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson and Krishnamurthy, 1996</td>
<td><em><em>Test characteristics (95% CI)</em> of tGCS score &lt;13</em>*&lt;br&gt;Sensitivity: 100% (79.41 to 100)&lt;br&gt;Specificity: 85.58% (82.99 to 87.90)&lt;br&gt;PLR: 6.93 (5.87 to 8.19)&lt;br&gt;NLR: 0&lt;br&gt;PPV: 11.85% (6.93 to 18.53)&lt;br&gt;NPV: 100% (99.48 to 100)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caterino, <em>et al.</em>, 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continued</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson and Krishnamurthy, 1996</td>
<td>NR</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
<td>Population Characteristics</td>
<td>Setting and Dates Assessments Performed</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>----</td>
<td></td>
</tr>
</tbody>
</table>
| Leijdesdorff, et al., 2014 | Retrospective cohort | Trauma patients with severe TBI (AIS codes for intracranial injury and skeletal injury with severity code ≥3). Exclusion: patients deceased at the scene of the accident and not admitted to the hospital. | Age (mean, years): 45.2 (SD 23.2)  
Male: 61.2%  
Race: NR  
ISS (mean): 18.3 (SD 9.2)  
ISS <16: 40.9%  
ISS ≥16: 58.8%  
In hospital mortality: 8.1%  
Trauma type  
- Brainstem: 1.2%  
- Cerebellum: 7%  
- Cerebrum: 86.2%  
- Contusion: 51.8%  
- Hemorrhage: 54.7%  
- Skull fracture: 43.8%  
- Skull base fracture: 25%  
- Skull vault fracture: 23.9%  
GCS score ≤13: 25% | Netherlands  
Urban  
Level I, II, and III trauma centers  
2003 to 2011 | 1,250 |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Outcomes</th>
<th>Glasgow Coma Scale Used</th>
<th>Personnel Performing Assessments and Where Assessed</th>
<th>Potential Confounders</th>
<th>Results: Univariate</th>
<th>Method for Constructing Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leijdesdorff, et al., 2014</td>
<td>Mortality</td>
<td>tGCS</td>
<td>NR</td>
<td>Data missing for n=3 patients with severe TBI</td>
<td><strong>Crude OR (95% CI) for risk of in-hospital mortality with GCS score &gt;12 as reference</strong>&lt;br&gt;GCS score 8–12: 5.57 (2.36 to 13.15)&lt;br&gt;GCS score &lt;8: 28.09 (13.95 to 56.58)&lt;br&gt;GCS score unknown: 5.20 (2.44 to 11.07)</td>
<td>Multivariate logistic regression analysis, age, GCS, and ISS independent prognostic factors for mortality</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Leijdesdorff, et al., 2014 | Adjusted OR (95% CI) for risk of in-hospital mortality with GCS score >12 as reference  
GCS score 8–12: 3.89 (1.61 to 9.40)  
GCS score <8: 19.24 (9.11 to 40.62)  
GCS score unknown: 4.46 (2.05 to 9.68) | Test characteristics (95% CI)*  
GCS score ≤12 vs. GCS score <8  
Sensitivity: 87.2% (77.7 to 93.7) vs. 71.8% (60.5 to 81.4)  
Specificity: 70.7% (67.5 to 73.7) vs. 85.9% (83.4 to 88.1)  
PLR: 2.97 (2.60 to 3.40) vs. 5.10 (4.11 to 6.32)  
NRL: 0.18 (0.10 to 0.32) vs. 0.33 (0.23 to 0.47)  
PPV: 21.0% (16.7 to 25.8) vs. 31.3% (24.6 to 38.6)  
NPV: 98.4% (97.1 to 99.2) vs. 97.1% (95.7 to 98.2) |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Discrimination or Calibration</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leijdesdorff, <em>et al.</em>, 2014</td>
<td>NR</td>
<td>Moderate</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| Majdan, et al., 2015 | Prospective cohort | All patients with GCS score ≤12 within 48 hours after the accident and/or AIS head score >2 were included in the study. | Age (median, years): 50 (IQR: 29-69)  
Male: 72%  
Race: NR  
ISS (median): 26 (IQR: 17–41)  
6-month mortality: 39%  
Intubated in out-of-hospital setting: 55%  
**Trauma type**  
- Traumatic subarachnoid hemorrhage: 59%  
- Epidural hematoma: 17%  
- Subdural hematoma hypotension: 54%  
- Hypotension: 11%  
- Hypoxia: 19%  
Mechanism of injury  
- Traffic accident: 39%  
- Same-level fall: 28%  
- High-level fall: 12%  
- Violence: 2%  
- Other cause: 16%  
- Unknown cause: 2%  
Head trauma: 100%  
GCS score ≤12: 100%  
Filed GCS (median): 6 (IQR: 3-11) | Austria Urban  
International Neurotrauma Research Organization (INRO)  
In the field and hospital  
2009 to 2012 | 445 |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Outcomes</th>
<th>Glasgow Coma Scale Used</th>
<th>Personnel Performing Assessments and Where Assessed</th>
<th>Potential Confounders</th>
<th>Results: Univariate</th>
<th>Method for Constructing Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majdan, et al., 2015</td>
<td>Prognostic performance of mGCS in the field and at admission</td>
<td>mGCS, tGCS</td>
<td>Physicians or paramedics in the field and by an anesthesiologist at admission</td>
<td>No specific measures for interrater disagreements</td>
<td>Reported OR (95% CI) of field vs. admission mGCS on 6-month mortality, with mGCS score=6 as reference &lt;br&gt; mGCS score=1: 23.2 (7.7 to 69.4) vs. 4 (1.7 to 9.1) &lt;br&gt; mGCS score=2: 41.1 (6.3 to 267.9) vs. 2.5 (0.4 to 17.6) &lt;br&gt; mGCS score=3: 5.9 (1.6 to 21.5) vs. 1.5 (0.4 to 6.1) &lt;br&gt; mGCS score=4: 4.4 (1.4 to 14.4) vs. 1.2 (0.4 to 3.9) &lt;br&gt; mGCS score=5: 6.2 (1.9 to 20.3) vs. 1.2 (0.4 to 3.8)</td>
<td>Logistic regression models fit using GCS motor score and pupillary reactivity in field and at admission as single predictors, 6-month mortality as response variable</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------</td>
<td>--------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Majdan, et al., 2015 | **Reported OR (95% CI of field vs. admission mGCS on 6-month mortality, with mGCS score ≤5 as reference**  
*Adjusted for age and pupillary reaction*  
mGCS score=1: 5.3 (2.5 to 11.4) vs. 2.9 (1.4 to 6.1)  
mGCS score=2: 15.6 (2.2 to 108) vs. 1.9 (0.2 to 21.6)  
mGCS score=3: 1.4 (0.47 to 4.4) vs. 1.5 (0.6 to 7.2)  
mGCS score=4: 1.8 (0.74 to 4.2) vs. 0.8 (0.2 to 2.8)  
*Adjusted for age, CT classification, hypoxia, hypotension, traumatic subarachnoid hemorrhage, epidural hematoma, and pupillary reaction*  
mGCS score=1: 3 9 (1.1 to 8.7) vs. 1.29 (0.45 to 3.7)  
mGCS score=2: 53.2 (2.7 to 1040) vs. unable to report due to singularity  
mGCS score=3: 0.63 (0.14 to 2.8) vs. 3.1 (0.49 to 19.1)  
mGCS score=4: 0.44 (0.11 to 1.7) vs. 0.8 (0.17 to 3.8) | NR |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Discrimination or Calibration</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majdan, <em>et al.</em>, 2015</td>
<td>Univariable AUROC of mGCS to predict 6-month mortality field vs. admission: 0.754 vs. 0.635</td>
<td>Moderate</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Eligibility Criteria</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td>Retrospective cohort</td>
<td>Patients 5 to 18 years old with blunt TBI, transported by the Milwaukee County EMS (MCEMS) System to the Children's Hospital of Wisconsin (CHW), with data on out-of-hospital and ED GCS scores. Exclusion: children without both scores documented, preexisting neurological illness, history of bleeding disorder, penetrating head trauma, and those transferred from other centers.</td>
</tr>
<tr>
<td>Reisner, et al., 2014</td>
<td>Retrospective cohort</td>
<td>Patients with available GCS score and at least one reliable high-mortality TBI value in the initial 15 minutes of transportation.</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Outcomes</td>
<td>Glasgow Coma Scale Used</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td>Mortality GOS score</td>
<td>tGCS</td>
</tr>
<tr>
<td>Reisner, et al., 2014</td>
<td>Mortality AIS</td>
<td>tGCS</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results: Multivariate</td>
<td>Measures of Diagnostic Accuracy</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td>NR</td>
<td>Test characteristics (95% CI),* out-of-hospital vs. ED tGCS score ≤13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mortality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 80.0% (28.4 to 99.5) vs. 80.0% (28.4 to 99.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 74.4% (67.4 to 80.6) vs. 76.0% (68.3 to 82.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 3.13 (1.89 to 5.18) vs. 3.34 (1.97 to 5.64)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.27 (0.05 to 1.55) vs. 0.26 (0.05 to 1.52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 8.0% (2.2 to 19.2) vs. 10.3% (2.9 to 24.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 99.3% (95.9 to 100) vs. 99.1% (95.1 to 100)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GOS score of severe disability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity: 39.7% (27.0 to 53.4) vs. 39.6% (25.8 to 54.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 80.7% (72.2 to 87.5) vs. 80.2% (70.8 to 87.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLR: 2.05 (1.26 to 3.36) vs. 2.00 (1.17 to 3.41)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLR: 0.75 (0.60 to 0.94) vs. 0.75 (0.59 to 0.97)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPV: 51.1% (35.8 to 66.3) vs. 50.0% (33.4 to 66.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NPV: 72.4% (63.8 to 80.0) vs. 72.6% (63.1 to 80.8)</td>
</tr>
<tr>
<td>Reisner, et al., 2014</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Discrimination or Calibration</td>
<td>Risk of Bias</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Nesia, et al., 2012</td>
<td>NR</td>
<td>Moderate</td>
</tr>
<tr>
<td>Reisner, et al., 2014</td>
<td>AUROC (95% CI), any tGCS vs. GCS &lt;15 vs. GCS &lt;8 Head AIS ≥5: 0.90 (0.86 to 0.93) vs. 0.80 (0.76 to 0.85) vs. 0.59 (0.52 to 0.66) All-cause mortality: 0.85 (0.80 to 0.90) vs. 0.82 (0.77 to 0.86) vs. 0.65 (0.59 to 0.70) Head AIS ≥5/procedure: 0.89 (0.86 to 0.92) vs. 0.78 (0.73 to 0.82) vs. 0.55 (0.49 to 0.62)</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Please see Appendix C. Included Studies for full study references.

AIS=Abbreviated Injury Scale; AUROC=area under the receiver operating characteristic curve; BP=blood pressure; CHW=Children’s Hospital of Wisconsin; CI=confidence interval; CT=computed tomography; ED=emergency department; EMS=emergency medical services; GCS=Glasgow Coma Scale; GOS=Glasgow Outcome Scale; ICP=intracranial pressure; INRO=International Neurotrauma Research Organization; IQR=interquartile range; ISS=injury severity score; LOS=length of stay; MCEMS=Milwaukee County Emergency Medical Services; mGCS=motor scale of GCS; MVA= motor vehicle accident; n=number; NLR=negative likelihood ratio; NPV=negative predictive value; NR=not reported; OR=odds ratio; PICU=pediatric intensive care unit; PLR=positive likelihood ratio; PPV=positive predictive value; PRBC=packed red blood cells; SD=standard deviation; TBI=traumatic brain injury; tGCS=total GCS; TS=Trauma Scores; vs.=versus

*Calculated
## Appendix J. Studies of Reliability and Ease of Use

### Table J-1. Characteristics of reliability and ease of use studies

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design Objective</th>
<th>GCS Used Measures Assessed</th>
<th>Type of Assessment or Intervention</th>
<th>Setting of Assessment</th>
<th>Personnel Performing Assessments</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbabi, et al., 2004</td>
<td>Cross-sectional study comparing out-of-hospital and ED data.</td>
<td>tGCS</td>
<td>Field vs. ED agreement</td>
<td>2 Level 1 Trauma Centers (1 urban) in the USA with actual patient data. Conducted from January 1994 to December 2001</td>
<td>Not reported</td>
<td>Eligible patients: 19,409 Analyzed patients: 7,823 (had field and ED GCS data)</td>
</tr>
<tr>
<td>Bledsoe, et al., 2015</td>
<td>Cross-sectional study to evaluate tGCS and its components in standardized video vignettes by EMS personnel in educational settings.</td>
<td>tGCS and mGCS (taken from tGCS)</td>
<td>Ease of use: correct scoring vs. expert scoring (2 board certified neurologists)</td>
<td>Setting NR, providers were from Nevada, Texas, Florida, and Minnesota, and provided simulation of 10 standardized video vignettes. Conducted from January to March in 2013</td>
<td>AEMT CCP EMT Nurse Paramedic Physician Resident</td>
<td>217 Providers AEMT: 25 CCP: 6 EMT: 19 Nurse: 82 Paramedic: 43 Physician: 10 Resident: 22 Not stated: 10</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Eligibility Criteria</td>
<td>Provider Characteristics</td>
<td>Patient Characteristics</td>
<td>Outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arbabi, et al., 2004</td>
<td>Adult, ≥18 years old, trauma patients. Exclusion: transfers from an outside hospital, burn patients</td>
<td>Not reported</td>
<td>Not reported separately for subset with GCS data</td>
<td>Agreement on GCS category (3-8, 9-13, 14-15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bledsoe, et al., 2015</td>
<td>Convenience sample of attendees at educational setting.</td>
<td>Age (mean, years): 36.2</td>
<td>Vignettes used, patient characteristics NR.</td>
<td>Accuracy: percent correct score</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male: 53.0%</td>
<td>Number of vignettes at each tGCS Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race: NR</td>
<td>3: 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Years of experience</td>
<td>5: 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;1 year: 22.7%</td>
<td>9: 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 to 10 years: 49.8%</td>
<td>11: 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;10 years: 34.6%</td>
<td>13: 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not stated: 5.1%</td>
<td>14: 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15: 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Arbabi, et al., 2004 | Agreement between field vs. ED assessment, no significant differences  
|                   | Same GCS category: 82% (6,382)                                             | High         |
|                   | Higher/improved category in ED: 3% (229)                                  |              |
|                   | Lower/worsened category in ED: 15% (1,212)                                |              |
| Bledsoe, et al., 2015 | Correct Scores (95% CI) tGCS vs. mGCS  
<p>|                   | Total across all vignettes and participants: 33.1% (30.2 to 36.0) vs. 59.8% (58.1 to 61.5) | Low          |
|                   | Highest percent correct by provider (in residents for both measures): 51% (44.5 to 57.5) vs. 78% (71.5 to 84.5) |              |
|                   | Lowest percent correct by provider: 29% (10.3 to 47.7) for nurses vs. 51% (43.7 to 58.3) for EMTs |              |
|                   | <strong>Other outcomes</strong>                                                        |              |
|                   | -9.2% assigned values that did not exist                                  |              |
|                   | -Accuracy was lowest for tGCS scores of 9 to 13 (&lt;20%; data taken from figure values NR) |              |</p>
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design Objective</th>
<th>GCS Used Measures Assessed</th>
<th>Type of Assessment or Intervention</th>
<th>Setting of Assessment</th>
<th>Personnel Performing Assessments</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinh, <em>et al.</em>, 2013</td>
<td>Cross-sectional study comparing EMS and ED vital signs.</td>
<td>tGCS</td>
<td>Field vs. ED agreement</td>
<td>Major trauma center in Sydney, Australia with actual field data. Conducted from January 2011 to October 2012</td>
<td>NR</td>
<td>Eligible patients: 1,265 Analyzed patients: 1,181</td>
</tr>
</tbody>
</table>

N = Number of assessments.
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Eligibility Criteria</th>
<th>Provider Characteristics</th>
<th>Patient Characteristics</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinh, et al., 2013</td>
<td>Patients ≥15 year old transported directly from the scene of injury via ambulance Exclusion: transfer from another hospital, transported by aeromedical retrieval, missing EMS data, inconsistent arrival time, or no vital signs at scene or ED.</td>
<td>None reported</td>
<td>Age (mean, years): 43 (SD 20) Male: 70% Race: NR Prehospital intubation: 1.2% Penetrating trauma: 7% GCS score ≤13: 10% ISS score &gt;15: 14.5% Mortality: 10.5%</td>
<td>Agreement using intra class coefficients and Bland-Altman plots</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dinh, et al., 2013</td>
<td>Intra class correlation coefficient: 0.74 (95% CI, 0.37 to 1.12)</td>
<td>Moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bland-Altman plot: 96.3% of out-of-hospital EO pairs within predetermined range of acceptability of 3 points</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design Objective</td>
<td>GCS Used Measures Assessed</td>
<td>Type of Assessment or Intervention</td>
<td>Setting of Assessment</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>-------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Feldman, et al., 2015</td>
<td>RCT to assess ability of EMS personal to correctly score the tGCS and its components and to determine if scoring improves with the use of a scoring aid.</td>
<td>tGCS and mGCS (taken from tGCS)</td>
<td>Ease of use: correct scoring with a scoring aid vs. scoring without a scoring aid</td>
<td>Setting NR, providers were from an urban, academic Level 1 trauma center in Ohio, provided with simulation of 9 standardized written scenarios. Conducted from April to June in 2013</td>
</tr>
<tr>
<td>Heim, et al., 2009</td>
<td>Cross-sectional survey to assess knowledge of GCS and scoring of a clinical scenario.</td>
<td>tGCS and mGCS (taken from tGCS)</td>
<td>Ease of use: knowledge and correct scoring of a clinical case</td>
<td>16 helicopter bases in Switzerland given simulated clinical scenario. Conducted in May 2004</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Eligibility Criteria</td>
<td>Provider Characteristics</td>
<td>Patient Characteristics</td>
<td>Outcomes</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Feldman, et al., 2015</td>
<td>EMTs or paramedics who had transported a patient to the ED and were &gt;18 years old.</td>
<td>Age (mean, years): 36 (SD 9)</td>
<td>Scenarios used, patient characteristics</td>
<td>Accuracy: complete agreement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male: 88.2%</td>
<td>NR. Mild, moderate, and severe TBI; no other information provided.</td>
<td>Score within 1 point</td>
</tr>
<tr>
<td></td>
<td></td>
<td>White: 83.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Black: 14.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>American Indian/Alaskan Native: 1.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asian/Pacific Islander: 0.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other race: 0.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Basic EMT: 46.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate EMT: 1.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Paramedic: 52.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Experience (mean, years): 12 (SD 4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heim, et al., 2009</td>
<td>Based at 1 of 16 participating helicopter bases, with prior training in air rescue,</td>
<td>Level</td>
<td>Patient in scenario has TBI, no other details reported</td>
<td>GCS scoring of case</td>
</tr>
<tr>
<td></td>
<td>which included registrars (in training), fellows, consultants (specialist in hospital),</td>
<td>Registrar: 38.8%</td>
<td></td>
<td>Knowledge of GCS components</td>
</tr>
<tr>
<td></td>
<td>or private practices.</td>
<td>Fellow: 35.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consultant: 7.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Private practice: 18.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specialty</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anesthesia: 61.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>General medicine: 18.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internal medicine: 16.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other: 3.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feldman, <em>et al.</em>, 2015</td>
<td><strong>Accuracy tGCS vs. mGCS</strong>&lt;br&gt; All scenarios: 41.0% vs. 50.6%&lt;br&gt; Mild TBI scenarios: 54.2% vs. 74.6%&lt;br&gt; Moderate TBI scenarios: 28.8% vs. 35.6%&lt;br&gt; Severe TBI scenarios: 40.0% vs. 41.7%&lt;br&gt; <strong>Accuracy tGCS with aid vs. tGCS without aid</strong>&lt;br&gt; All scenarios: 25.0% vs. 56.7%; difference: 31.9% (95% CI, 18.3 to 45.6)&lt;br&gt; Mild TBI scenarios: 44.8% vs. 63.3%; difference: 14.3% (95% CI, -6.1 to 34.6)&lt;br&gt; Moderate TBI scenarios: 10.3% vs. 46.7%; difference: 31.4% (95% CI, 10.5 to 52.3)&lt;br&gt; Severe TBI scenarios: 20.0% vs. 60.0%; difference: 40.0% (95% CI, 16.9 to 63.1)&lt;br&gt; <strong>Accuracy mGCS with aid vs. mGCS without aid</strong>&lt;br&gt; All scenarios: 30.7% vs. 70.0%; difference: 39.7% (95% CI, 26.2 to 53.1)&lt;br&gt; Mild TBI scenarios: 58.6% vs. 90.0%; difference: 29.3% (95% CI, 6.1 to 52.5)&lt;br&gt; Moderate TBI scenarios: 20.7% vs. 50.0%; difference: 29.3% (95% CI, 10.5 to 52.3)&lt;br&gt; Severe TBI scenarios: 13.3% vs. 70.0%; difference: 56.7% (95% CI, 36.2 to 77.1)</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heim, <em>et al.</em>, 2009</td>
<td><strong>Incorrect (correct) scores</strong>&lt;br&gt; tGCS: 36.9% (63.1%)&lt;br&gt; mGCS: 27.2% (72.8%)&lt;br&gt; Registrars (trainees): 47.5%&lt;br&gt; Fellow: 33.3%&lt;br&gt; Consultant: 0%&lt;br&gt; Private practice: 36.8%&lt;br&gt; Specialty was not associated with difference in errors (anesthesia, internal medicine, general practice, other)&lt;br&gt; All respondents knew the GCS&lt;br&gt; Incorrectly named components: 5.8%&lt;br&gt; Attributed wrong number of points: 3.9%&lt;br&gt; Knew minimum was 3: 100%&lt;br&gt; Knew maximum was 15: 99%</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design Objective</td>
<td>GCS Used Measures Assessed</td>
<td>Type of Assessment or Intervention</td>
<td>Setting of Assessment</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------</td>
<td>------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Holmes, et al., 2005</td>
<td>Cross-sectional study comparing the pediatric GCS in children ≤2 years and the standard GCS in children &gt;2 years.</td>
<td>tGCS and mGCS (taken from tGCS)</td>
<td>Interrater reliability</td>
<td>Level 1 trauma center in California using actual patient data. Conducted from 1998 to 2001</td>
</tr>
<tr>
<td>Kerby, et al., 2007</td>
<td>Cross-sectional study of linkage of EMS and trauma registry data.</td>
<td>tGCS and mGCS (taken from tGCS)</td>
<td>Field vs. ED agreement</td>
<td>Field and ED assessments of actual patient registry data linked from one Level 1 trauma center in Alabama, USA. Conducted from January 2000 to June 2003</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Eligibility Criteria</td>
<td>Provider Characteristics</td>
<td>Patient Characteristics</td>
<td>Outcomes</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Holmes, et al., 2005</td>
<td>Pediatric patients &lt;18 years with blunt head trauma presenting to the ED. Exclusion: children with trivial head trauma defined by falls from ground level or trauma resulting from walking or running into stationary objects if the only abnormal finding was a scalp laceration or abrasion, and children transferred who had undergone CT scanning before transfer.</td>
<td>None reported</td>
<td>Not reported for the 5% used for interrater reliability</td>
<td>Weighted kappa</td>
</tr>
<tr>
<td>Kerby, et al., 2007</td>
<td>Patients in trauma registry who were &gt;19 years and transported to ED with blunt or penetrating injury. Exclusion: transfers from other hospitals.</td>
<td>None reported</td>
<td>Age (mean, years): 38.9 (SD 15.7) Male: 68.6% White: 56.9% Black: 38.5% Other race: 4.6% Field intubation: 1.7% Blunt trauma: 81.0% Penetrating trauma: 19.0% Alcohol intoxication: 22.9% Positive illicit drugs: 11.4% ISS (mean): 10.6 (SD 9.1) Mortality: 2.7%</td>
<td>Weighed kappa</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Holmes, et al., 2005 | **Weighted kappa (95% CI) across raters**  
                   tGCS: 0.77 (0.38 to 1.00) for ≤2 year olds and 0.91 (0.75 to 1.00) for >2 year olds  
                   mGCS: 0.91 (0.75 to 1.00) for ages combined | High         |
| Kerby, et al., 2007 | **Weighed kappa (95% CI) tGCS vs. mGCS**  
                   Overall: 0.53 (0.48 to 0.58) vs. 0.48 (0.43 to 0.53)  
                   Transport time of <20 minutes: 0.56 (0.50 to 0.61) vs. 0.52 (0.46 to 0.57)  
                   Transport time of ≥20 minutes: 0.42 (0.32 to 0.52) vs. 0.35 (0.25 to 0.46)  
                   Examination of changes in blood pressure/hemodynamic stability show improvement from prehospital to ED suggesting difference may be primarily due to patient improvement. | High         |
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Design Objective</th>
<th>GCS Used Measures Assessed</th>
<th>Type of Assessment or Intervention</th>
<th>Setting of Assessment</th>
<th>Personnel Performing Assessments</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane, et al., 2002</td>
<td>Pre-post study to assess the impact of a teaching video on correct scoring of the tGCS.</td>
<td>tGCS</td>
<td>Ease of use: improvement in correct scoring</td>
<td>Setting NR, providers were from the USA, Mid Atlantic region, provided with simulations of 4 scenarios in 60-second videos. Conducted in 2000, no specific dates reported</td>
<td>EMT Paramedic students</td>
<td>75 (pre-post)</td>
</tr>
<tr>
<td>Lane, et al., 2002</td>
<td>Pre-post study to assess the impact of a teaching video on correct scoring of the tGCS, with randomized assignment to use of GCS reference cards.</td>
<td>tGCS</td>
<td>Ease of use: improvement in correct scoring</td>
<td>Setting NR, providers were from the USA, Mid Atlantic region, provided with simulations of 4 scenarios in 60-second videos. Conducted in 2000, no specific dates reported</td>
<td>EMT Paramedic students</td>
<td>46 (pre-post but 2 cohorts with and without cards)</td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td>Cross-sectional study to determine agreement between the out-of-hospital tGCS and the ED tGCS.</td>
<td>tGCS</td>
<td>Field vs. ED agreement</td>
<td>USA, Wisconsin Urban Children’s Hospital of Wisconsin trauma registry Level 1 trauma center 2000 to 2005</td>
<td>Advanced Life Support paramedic crew leader ED staff (usually physician or nurse)</td>
<td>Screened: 427 Eligible: 377 Included: 196 Analyzed: 185</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Eligibility Criteria</td>
<td>Provider Characteristics</td>
<td>Patient Characteristics</td>
<td>Outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------</td>
<td>--------------------------</td>
<td>-------------------------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane, et al., 2002</td>
<td>Attendees at annual EMS meeting.</td>
<td>Basic EMT: 55%  Paramedic EMT: 35%  RN: 7%  Other level of training: 3%  Male: NR  Race: NR  Years of service (mean): 13.0 (95% CI, 10.8 to 15.09)  Urban setting: 32%  Rural setting: 24%  Suburban setting: 44%  Participated in trauma course within 5 years: 62.7%</td>
<td>Scenarios used, patient characteristics NR.  Correct tGCS and mGCS score used  Scenario 1: 15 and 6  Scenario 2: 8 and 4  Scenario 3: 5 and 3  Scenario 4: 15 and 6</td>
<td>Accuracy: percent correct pre- and post-video</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane, et al., 2002</td>
<td>Paramedic class participants.</td>
<td>Group 1 (with reference cards) vs. group 2 (without reference cards)  Basic EMT: 100% vs. 100%  Years of service (mean): 5.2 (95% CI, 2.92 to 7.39) vs. 3.2 (95% CI, 2.21 to 4.29)  Participated in trauma course within 5 years: 100% vs. 84.6%</td>
<td>Scenarios used, patient characteristics NR.  Correct tGCS and mGCS score used  Scenario 1: 15 and 6  Scenario 2: 8 and 4  Scenario 3: 5 and 3  Scenario 4: 15 and 6</td>
<td>Accuracy: percent correct pre- and post-video</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td>Patients 5 to 18 years old with blunt TBI, transported by the Milwaukee County EMS System to the Children's Hospital of Wisconsin, with data on out-of-hospital and ED GCS scores. Exclusion: children without both scores documented, preexisting neurological illness, history of bleeding disorder, penetrating head trauma, and those transferred from other centers.</td>
<td>None reported</td>
<td>Age (mean, years): 11 (SD 4.0)  Male: 69%  White: 33.5%  Black: 48.6%  Hispanic: 11.9%  Asian: 2.7%  Other race/ethnicity: 3.2%  Trauma type: 100% TBI  GCS score ≤13: 27%</td>
<td>Out-of-hospital and ED GCS scores agreement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane, et al., 2002</td>
<td><strong>Accuracy pre vs. post for tGCS</strong>&lt;br&gt;All 4 scenarios together: 14.7% vs. 64.0%; RR 4.36 (95% CI, 2.46 to 7.73)&lt;br&gt;Scenario 1: 76.0% vs. 98.7; RR 1.30 (95% CI, 1.14 to 1.48)&lt;br&gt;Scenario 2: 36.0% vs. 74.7; RR 2.07 (95% CI, 1.49 to 2.88)&lt;br&gt;Scenario 3: 45.3% vs. 94.7%; RR 2.09 (95% CI, 1.62 to 2.69)&lt;br&gt;Scenario 4: 64.0% vs. 89.3%; RR 1.40 (95% CI, 1.16 to 1.68)</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane, et al., 2002</td>
<td><strong>Multivariate analysis for associations with correct scores for post video assessments, standardized beta</strong>&lt;br&gt;Participation in prehospital trauma course: 0.429, p=0.001&lt;br&gt;Years of training and level of service were not significant&lt;br&gt;No significant associations noted for pre video assessments</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td><strong>Kappa (95% CI) tGCS:</strong> 0.69 (0.57 to 0.81)&lt;br&gt;<strong>Weighed kappa (95% CI) tGCS:</strong> 0.74 (0.63 to 0.85)&lt;br&gt;Pearson correlation 0.841&lt;br&gt;Concordance correlation 0.839&lt;br&gt;ED scores tended to be higher than prehospital but the difference was very small (0.4371) on average and there was no difference in the medians.</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Design</td>
<td>Objective</td>
<td>GCS Used Measures Assessed</td>
<td>Type of Assessment or Intervention</td>
<td>Setting of Assessment</td>
<td>Personnel Performing Assessments</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>Takahashi, et al., 2011</td>
<td>Cross-sectional study evaluating agreement between raters and accuracy of each scale of the GCS.</td>
<td>Evaluating agreement between raters and accuracy of each scale of the GCS.</td>
<td>tGCS only</td>
<td>Interrater reliability</td>
<td>10 medical facilities including 4 university hospitals in Japan provided with actual patient data. Conducted from April 2007 to April 2008</td>
<td>Physicians, Nurses, Residents, Paramedics, Medical students</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Eligibility Criteria</td>
<td>Provider Characteristics</td>
<td>Patient Characteristics</td>
<td>Outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Takahashi, et al., 2011</td>
<td>Patients ages 5 to 99 years. Exclusion: patients for whom the evaluation posed a risk of aggravating their condition due to time loss.</td>
<td>None reported other than profession</td>
<td>Not reported separately for TBI Overall Age (mean, years): 58.6 (SD 22.4) Male: 52.7% Race: NR</td>
<td>Weighted kappa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Results</td>
<td>Risk of Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Takahashi, *et al.*, 2011 | Weighted kappa (95% CI) across raters  
TBI patients only: 0.74 (0.71 to 0.76) | Moderate     |

Please see Appendix C. Included Studies for full study references.

AEMT= advanced emergency medical technician; CCP= critical care paramedic; CI=confidence interval; CT= computed tomography; ED= emergency department; EMS= emergency medical services; EMT= emergency medical technician; GCS= Glasgow Coma Scale; ISS=injury severity score; mGCS= motor Glasgow coma scale; n= number; NR=not reported; RCT= randomized controlled trial; RN= registered nurse; RR= relative risk; SD= standard deviation; TBI=traumatic brain injury; tGCS= total Glasgow Coma Scale; vs.= versus
### Appendix K. Quality Assessment of Studies of Predictive Utility

#### Table K-1. Quality assessments of predictive utility studies

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Participation</th>
<th>Study Attrition</th>
<th>Prognostic Factor Measurement</th>
<th>Outcome Measurement</th>
<th>Study Confounding</th>
<th>Statistical Analysis and Reporting</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acker, et al., 2014</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>Low, 110/2,341 excluded due to GCS missing data.</td>
<td>Unclear, details on methods for measuring outcomes not reported.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Al-Salamah, et al., 2004</td>
<td>Low</td>
<td>Low, 20/815 excluded due to missing data from trauma registry.</td>
<td>Low, 20/815 excluded due to missing data from trauma registry.</td>
<td>Unclear, details on methods for measuring outcomes not reported.</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Beskind, et al., 2014</td>
<td>Low</td>
<td>Low</td>
<td>Moderate, approximately 25% missing GCS data, multiple imputation performed.</td>
<td>Low</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Low</td>
</tr>
<tr>
<td>Brown, et al., 2014</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>Moderate, approximately 40% missing GCS data, multiple imputation performed.</td>
<td>Low (NTDB)</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Caterino and Raubenolt, 2012</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>High, approximately 25% missing GCS data, excluded from analysis.</td>
<td>Low</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Caterino, Raubenolt, and Cudnik, 2011</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Cicero and Cross, 2013</td>
<td>Low</td>
<td>Low, between 2% and 4% missing data for outcomes.</td>
<td>High, approximately 51% excluded due to missing GCS data.</td>
<td>Low (NTDB)</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Corrigan, et al., 2014</td>
<td>Low</td>
<td>Unclear, approximately 50% missing data (either GCS or outcomes).</td>
<td>Unclear, approximately 50% missing data (either GCS or outcomes).</td>
<td>Low (NTDB)</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Davis, et al., 2006</td>
<td>Low</td>
<td>Moderate, 30% missing outcome data.</td>
<td>Unclear. Missing GCS data not reported.</td>
<td>Unclear, details on methods for measuring outcomes not reported.</td>
<td>Not applicable, study report discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Eken, et al., 2009</td>
<td>Low</td>
<td>Low, &lt;2% missing data for outcomes.</td>
<td>Unclear, missing GCS data not reported.</td>
<td>Unclear, details on methods for measuring outcomes not reported.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Gill, et al., 2005</td>
<td>Low</td>
<td>Low, &lt;1% missing data for outcomes.</td>
<td>Low, 216/8,648 excluded due to missing data.</td>
<td>Low</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Low</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Participation</td>
<td>Study Attrition</td>
<td>Prognostic Factor Measurement</td>
<td>Outcome Measurement</td>
<td>Study Confounding</td>
<td>Statistical Analysis and Reporting</td>
<td>Risk of Bias</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>-----------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>-----------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Gill, et al., 2006</td>
<td>Low</td>
<td>Low, &lt;1% missing data for outcomes.</td>
<td>Low, 1415/8,648 excluded due to missing GCS data.</td>
<td>Low</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Low</td>
</tr>
<tr>
<td>Haukoos, et al., 2007</td>
<td>Low</td>
<td>Low</td>
<td>Low, 583/21,753 excluded due to missing GCS data.</td>
<td>Low</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Low</td>
</tr>
<tr>
<td>Healey, et al., 2003</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>Low, 1,926 missing GCS scores for 204,181 patients, excluded from analysis.</td>
<td>Low (NTDB)</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Low</td>
</tr>
<tr>
<td>Holmes, et al., 2005</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>Unclear, missing GCS data not reported.</td>
<td>Low</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Not applicable, study reports discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Johnson and Krishnamurthy, 1996</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>Unclear, missing GCS data not reported.</td>
<td>Low</td>
<td>Not applicable, study reports diagnostic accuracy only.</td>
<td>Not applicable, study reports diagnostic accuracy only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Kupas, et al., 2016</td>
<td>Low</td>
<td>Moderate, 33% missing ICU admission, other outcomes &lt;3% missing data</td>
<td>Unclear, missing GCS data not reported.</td>
<td>Low</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Leijbesdorf, et al., 2014</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Majdan, et al., 2015</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear, not reported if data was validated.</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td>Unclear</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Unclear</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Reisner, et al., 2014</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Ross, et al., 1998</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>High, 56% missing GCS data, excluded from analysis.</td>
<td>Unclear, details on methods for measuring outcomes not reported.</td>
<td>Not applicable, study reports diagnostic accuracy only.</td>
<td>Not applicable, study reports diagnostic accuracy only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Thompson, et al., 2011</td>
<td>Low</td>
<td>Low, &lt;1% missing data for outcomes.</td>
<td>Unclear, missing GCS data not reported.</td>
<td>Low</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Not applicable, study reports diagnostic accuracy and discrimination only.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Van de Voorde, et al., 2008</td>
<td>Low</td>
<td>Unclear, attrition not reported.</td>
<td>Unclear, missing GCS data not reported.</td>
<td>Unclear, details on methods for measuring outcomes not reported.</td>
<td>Not applicable, study reports diagnostic accuracy only.</td>
<td>Not applicable, study reports diagnostic accuracy only.</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

K-20
Please see Appendix C. Included Studies for full study references.

GCS= Glasgow Coma Scale; NTDB= National trauma Data Bank
Appendix L. Quality Assessments of Studies of Reliability and Ease of Use

Table L-1. Quality assessments of reliability and ease of use studies

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Assessment Type</th>
<th>Patient Selection</th>
<th>Index Tests</th>
<th>Reference Standard</th>
<th>Flow and Timing</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbabi, et al., 2004</td>
<td>Field vs. ED agreement</td>
<td>High</td>
<td>Unclear</td>
<td>Moderate</td>
<td>Unclear</td>
<td>High</td>
</tr>
<tr>
<td>Blesdsoe, et al., 2015</td>
<td>Ease of use</td>
<td>Low</td>
<td>Not applicable</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Dinh, et al., 2013</td>
<td>Field vs. ED agreement</td>
<td>Low</td>
<td>Unclear</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Feldman, et al., 2015</td>
<td>Ease of use</td>
<td>Low</td>
<td>Not applicable</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Heim, et al., 2009</td>
<td>Ease of use</td>
<td>High</td>
<td>Not applicable</td>
<td>Unclear</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Holmes, et al., 2005</td>
<td>Interrater reliability</td>
<td>High</td>
<td>Unclear</td>
<td>Moderate</td>
<td>Unclear</td>
<td>High</td>
</tr>
<tr>
<td>Kerby, et al., 2007</td>
<td>Field vs. ED agreement</td>
<td>High</td>
<td>High</td>
<td>Moderate</td>
<td>Unclear</td>
<td>High</td>
</tr>
<tr>
<td>Lane, et al., 2002</td>
<td>Ease of use</td>
<td>High</td>
<td>Not applicable</td>
<td>Unclear</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Nesiama, et al., 2012</td>
<td>Field vs. ED agreement</td>
<td>Moderate</td>
<td>High</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Takahashi, et al., 2011</td>
<td>Interrater reliability</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Please see Appendix C. Included Studies for full study references.

ED= emergency department; vs.= versus