

A critical review to traumatic brain injury clinical practice guidelines

Bao-shan Di, MD^{a,b}, Min Wei, MD^c, Wen-juan Ma, MMed^d, Qi Zhang, MD^a, An-qing Lu, MD^a, Hu Wang, MD^a, Yang Niu, MD^a, Nong Cao, MD^{b,*}, Tian-kang Guo, MD^{a,*}

Abstract

The aim of this study was to assess the quality of clinical practice guidelines of traumatic brain injury (TBI) and investigate the evidence grading systems.

A systematic search of relevant guideline websites and literature databases (including PubMed, NGC, SIGN, NICE, GIN, and Google) was undertaken from inception to May 2018 to identify and select TBI guidelines. Four independent reviewers assessed the eligible guidelines using the Appraisal of Guidelines for Research and Evaluation (AGREE II) instrument. The degree of agreement was evaluated with intraclass correlation coefficient (ICC).

From 1802 records retrieved, 12 TBI guidelines were included. The mean scores for each AGREE II domain were as follows: scope and purpose (mean \pm SD= 74.2 \pm 9.09); stakeholder involvement (mean \pm SD= 54.6 \pm 11.6); rigor of development (mean \pm SD=70.1 \pm 13.6); clarity and presentation (mean \pm SD=78.4 \pm 11.5); applicability (mean \pm SD= 60.5 \pm 13.6); and editorial independence (mean \pm SD=61.7 \pm 14.8). Ten guidelines were rated as “recommended.” The ICC values ranged from 0.73 to 0.95. Seven grading systems were used by TBI guidelines to rate the level of evidence and the strength of recommendation.

Most TBI guidelines got a high-quality rating, whereas a standardized grading system should be adopted to provide clear information about the level of evidence and strength of recommendation in TBI guidelines.

Abbreviations: AGREE = Appraisal of Guidelines for Research and Evaluation, CPGs = clinical practice guidelines, ICC = intraclass correlation coefficients, IOM = Institute of Medicine, NHMRC = National Health and Medical Research Council, PRISMA = preferred reporting items for systematic reviews and meta- analyses, SIGN = Scottish Intercollegiate Guidelines Network, TBI = traumatic brain injury.

Keywords: clinical practice guidelines, quality assessment, traumatic brain injury

1. Introduction

Traumatic brain injury (TBI) is one of the leading causes of death and disability in both developing and developed countries, with the highest incidence among young people <30 years of age.^[1,2] Clinical practice guidelines (CPGs) have been developed by various organizations from different countries to improve patient's outcomes of TBI; the brain trauma community's approach to guideline development has evolved as the science and application of evidence-based medicine advanced.

During the past 20 years, >30 TBI guidelines have been developed and updated from different organizations.^[3] However,

TBI guidelines vary in quality, comprehensiveness, and grading system, leading to difficulties with standardization of care, adaptation, and implementation. Despite this, a major criticism of the TBI guidelines is that they may not be appropriate for use in all locations due to differences in available resources. Although some previous studies have evaluated quality of existing TBI guidelines, they just have focused on the subsets of TBI severity such as mild TBI only^[3–5] or reviewed only a limited number of TBI guidelines,^[6] and none of them focused on the grading systems of TBI guidelines adopted, which actually is very important to help guideline users, readers, and stakeholders to understand the confidence of estimate of the effects and the strength of recommendations. Moreover, according to Institute of Medicine (IOM) statement of guideline,^[7] some old TBI guidelines (published before 2007) evaluated by these studies had been abandoned in clinical practice as the recommendations were outdated.

Hence, we conducted this study to assess and summarize the quality of all currently available international TBI guidelines by conducting a systematic review using the Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument.^[6] We also compared the codes of evidence quality and strength of recommendation among different TBI guidelines.

2. Materials and methods

2.1. Study design and protocol

This study conducted a comprehensive review of clinical guidelines using the AGREE II instrument. This study was performed in accordance with the guidelines from preferred

Editor: Edirweera Desapriya.

DBS and WM equally contributed to this study.

The authors have no funding and conflicts of interest to disclose.

^a Gansu Province People's Hospital, ^b The First Hospital of Lanzhou University,

^c Anesthesia Department, Traditional Chinese Medicine of Gansu Province,

^d Evidence-Based Medicine Center, Lanzhou University, Lanzhou, China.

* Correspondence: Tian-kang Guo, Gansu Province People's Hospital, Lanzhou 730000, China (e-mail: tiankanguo2016@163.com); Nong Cao, The First Hospital of Lanzhou University, Lanzhou 730000, China (e-mail: 442930413@qq.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medicine (2019) 98:9(e14592)

Received: 30 August 2018 / Received in final form: 21 January 2019 / Accepted: 28 January 2019

<http://dx.doi.org/10.1097/MD.0000000000014592>

reporting items for systematic reviews and meta-analyses (PRISMA).^[8]

2.2. Identification of guidelines

Systematic searches were performed in PubMed database from inception to May 31, 2018, combining the term “traumatic brain injury OR TBI” and a filter to identify guideline documents (Practice Guideline [pt] OR Guideline [pt] OR guideline* [ti] OR statement [ti], recommendations [ti] OR consensus [ti]). We also searched the websites of guideline development organizations, NICE (<https://www.nice.org.uk/guidance>) and SIGN (<http://www.sign.ac.uk/>), and guideline databases such as GIN (<http://www.g-i-n.net/>) and NGC (<https://www.guideline.gov/>). Besides, we searched Google Search Engine as well as the references of all the obtained guidelines to include more potential guidelines.

Two reviewers (DBS and WKP) independently evaluated search results to determine inclusion or exclusion of references and extracted the general characteristics of each guideline. Disagreements were resolved by consensus or by consulting the third expert adjudicator (GTK).

2.3. Selection of guidelines

The inclusion criteria were as follows: complete guideline text is available in English; guideline contains recommendations regarding TBI interventions; and the guideline should be published after 2007. If the guideline had updates, only the most recent version was assessed. For every guideline ultimately included, we thoroughly searched for accompanying technical and supporting documents to better inform our assessments. The following literatures will be excluded: duplicate guidelines, guidelines for patients, editorials, translations of guidelines, secondary or multiple publications, and short summaries. For multiple versions of guidelines, only the newest guidelines were included in the analysis and the older versions were excluded.

2.4. Quality appraisal of guidelines

We used the latest version of the AGREE II instrument to evaluate each TBI CPG meeting our inclusion criteria. According to AGREE II handbook, each CPG was scored on 23 items within 6 domains. Domain 1 (scope and purpose) is divided into 3 items: guideline objectives, health questions, and population application. Domain 2 (stakeholder involvement) is based on 3 items: guideline development group, preferences of target population, and target users. Domain 3 (rigor of development) includes 8 items: systematic methods used to search evidence, criteria for selection, strengths and limitations of the evidence, methods for formulating the evidence, health benefits and side effects of recommendations, explicit links between recommendation and supporting evidence, expert reviewers, and updating guideline for future use. Domain 4 (clarity and presentation) includes 3 items: recommendations are specific and unambiguous, different options for management, and key recommendations. Domain 5 (applicability) includes 4 items: facilitators and barriers, advice/tools to implement recommendations into practice, resources for implications, and auditing criteria. Domain 6 (editorial independence) is based on 2 items: editorial independence from the funding body and conflicts of interest of the guideline development members.

In this study, each TBI guideline was scored by 4 independent reviewers (DBS, WKP, MWJ, and M.W.J) according to AGREE II

user manual. Among the 4 reviewers, DBS and WKP are senior doctors of neurosurgery; MWJ and LY are methodologists in guideline development. Besides, YL had rich experiences in the application of AGREE II and published a study about using AGREE II to assess clinical guidelines.^[9,10] DBS and WKP were trained to use the AGREE II instrument through the online tutorials on the AGREE website.

The user manual defines each item and assists the user in determining a guideline’s score for that item. Items were scored based on a scale ranging from 1 (absence of item) to 7 (item is reported with exceptional quality). Domain scores were calculated by summing item scores within each domain from each reviewer, and then standardizing them as a percentage of the maximum possible score. AGREE II protocol states that no overall score is calculated to determine if a CPG is recommended or not recommended. Instead, guidelines in this study were recommend if the guidelines have >4 domains scoring >50%.^[6]

2.5. Strength of recommendation and level of evidence

The strength of recommendations and level of evidence of each TBI guideline were extracted if these guidelines adopted evidence grading systems.

2.6. Data analysis

We performed a descriptive statistics analysis using the calculation of the total score by each reviewer and the score per domain. The number of recommendations and the percentage distributions among quality of evidence and strength of recommendation classes were determined. Agreement between each reviewer’s scores was tested using a 2-way ANOVA with single-rater 2-way intraclass correlation coefficients (ICCs) with 95% confidence interval (CI) for each domain across all guidelines.^[11] According to a previous study,^[12] the degree of agreement between 0.01 and 0.20 was deemed minor, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 substantial, and 0.81 to 1.00 very good. A value of $P < .05$ denoted statistical significance. All tests were 2-sided. Statistical analyses were conducted using SPSS version 19.0 (SPSS Inc., Chicago, IL).

3. Results

3.1. Study selection

The initial search strategy identified 1802 titles and abstracts, 63 of which were removed for duplicates. From these, 1646 were excluded after reviewing abstracts. A reference and citation analysis was performed on the remaining 93 articles yielding an additional 112 abstracts. Full-text analysis was then performed on a total of 205 articles of which only 12^[13–24] met inclusion criteria (Fig. 1).

3.2. CPG characteristics

Overall, 12 TBI guidelines were included in this analysis (Table 1), representing 12 different organizations and spanning several countries on 4 continents. Of these 12 TBI guidelines, all of them^[22,28,4,29–43] were developed in high-income countries. The CPGs evaluated covered the full scope of adult and pediatric populations with 1 covering pediatric patients,^[20] 5 for adults patients,^[14,15,19,21,24] and 4 covering both populations.^[13,16,18,22] Regarding the severity of TBI, one-fourth of

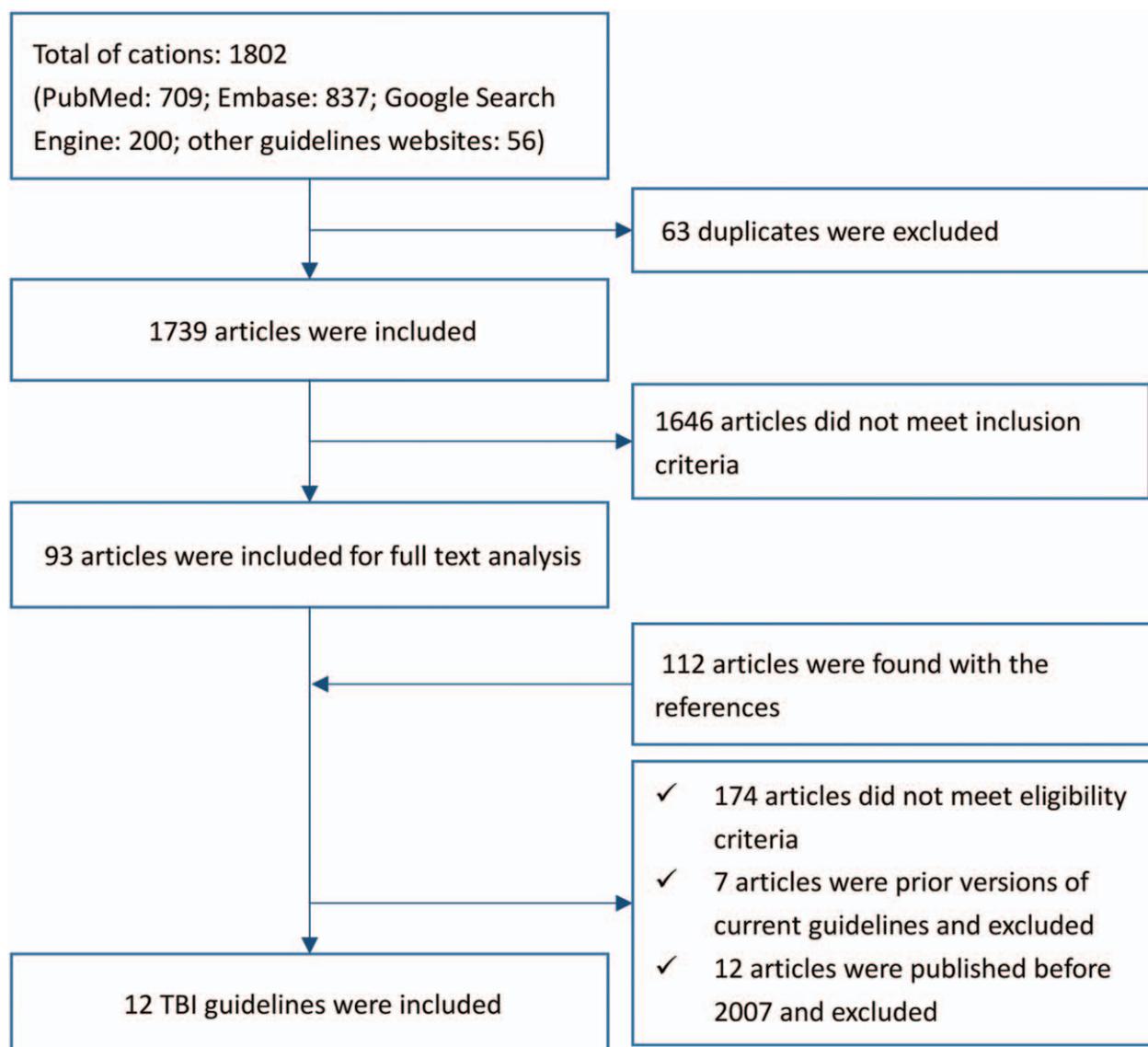


Figure 1. Flow chart of selecting TBI guidelines.

the guidelines were developed for minor or mild TBI,^[15,18,23] another third covered severe TBI,^[17,20,24] and the rest were developed for all levels of TBI severity.^[13,14,16,19,21,22] The majority (7) of CPGs focused on the early management of TBI.^[13,14,16,20–23] Of the 24 assessed CPGs, 4 were developed by professional organizations,^[13,15,23,24] 2 were developed by nonprofit organizations,^[16,20] 2 by international committees,^[18,21] and another 4 by national institutes or government organizations.^[14,17,19,22] The detailed information of included guidelines was shown in Table 1.

3.3. CPG quality assessment (AGREE)

3.3.1. Consistency. The ICC values, which indicate the overall agreement between reviewers, generally received higher reliability scores. The ICC values for TBI guidelines appraisal using the AGREE II ranged from 0.73 to 0.95 (Table 2). The ICCs for the AGREE appraisal conducted by the 4 reviewers were lowest in the “applicability” domain (0.73) but higher in the other 5 domains (all ≥ 0.75), which indicated the intrareviewer item score

agreement was good. Domain scores of the AGREE II quality assessment are illustrated in Table 2.

3.3.2. Domain 1—scope and purpose. This domain focuses on the overall objectives, expected benefits or outcomes and target population of the guidelines. The mean score of TBI guidelines in this domain is 74.2 with a SD of 9.09, and all guidelines scored >50 . The lowest score was 60.1 which was from guidelines for the treatment of minor and severe TBI (RASH, 2008). The highest score was 88.9, from Brain Trauma Foundation and American association of Neurological Surgeons (BTF/AANS, 2017).

3.3.3. Domain 2—stakeholder involvement. This domain contains items on the involvement degree of professional members, consideration of the views and preferences of the target population, and the definition of target users. Scores fluctuated remarkably with a mean score \pm SD of 54.6 ± 11.6 . Five (38%) TBI guidelines scored $<50\%$, of which the lowest was 40.2 from China (China, 2009).

Table 1

The characteristics of included TBI guidelines.

Guideline ID	Origin	Institution/guideline development group	Type of institution/guideline development group	Focus of the guideline	Patient population	Severity of brain injury	Country income
CMA, 2007 ^[13]	Canada	Canadian Medical Association	Professional organization	Early management	Adults and pediatrics	All levels	HIC
RHSA, 2008 ^[14]	Italy	Regional Healthcare Service Agency (requested by Ministry of Health)	Government organization	Early management	Adults	All levels	HIC
ACEP, 2009 ^[15]	USA	American College of Emergency Physicians	Professional organization	Early management and imaging diagnosis	Adults	Mild	HIC
SIGN, 2009 ^[16]	Edinburgh, Scotland	Scottish Intercollegiate Guidelines Network	Nonprofit	Early management	Adults and pediatrics	All levels	HIC
China, 2009 ^[17]	China (Taiwan)	China	Government organization	Early management, and ICU care	Not specify	Severe	HIC
EFNS, 2011 ^[18]	Europe	European Federation of Neurological Societies	International committee	Early management and Imaging diagnosis	Adults and pediatrics	Mild	HIC
NSW MoH, 2011 ^[19]	New South Wales, Australia	NSW Ministry of Health	Government organization	Early management of head injury patients	Adults	All levels	HIC
BTF, 2012 ^[20]	UK	Brain Trauma Foundation (BTF)	Nonprofit	Early management	Pediatrics	Severe	HIC
SCN, 2013 ^[21]	Scandinavia (Norway, Sweden and Finland)	Scandinavian Neurotrauma Committee	International committee	Early management	Adults	All levels	HIC
NIHCE, 2014 ^[22]	UK	National Institute for Health and Clinical Excellence (NIHCE)	National Institute	Early management	Adults and pediatrics	All levels	HIC
EAST, 2015 ^[23]	USA	Eastern Association for the Surgery of Trauma (EAST)	Professional organization	Early management	Not specify	Mild	HIC
BTF/AANS, 2017 ^[24]	USA	Brain Trauma Foundation/American association of Neurological Surgeons	Professional organization	Early management, and ICU care	Adults	Severe	HIC

3.3.4. Domain 3—rigor of development. This domain investigates the method and process of evidence search, grading, summary, and the formulation of the recommendations. The mean score and SD for this domain was 70.1 ± 13.6. All TBI guidelines scored >50%, of which the lowest was 51.3 from China (China, 2009).

3.3.5. Domain 4—clarity of presentation. This domain addresses the presentation and format of guidelines. The mean

score and SD in this domain was 78.4 ± 11.5. The lowest score was 57.4 from Brain Trauma Foundation (BTF, 2012).

3.3.6. Domain 5—application. This domain evaluates the consideration of facilitators or barriers when implementing the guidelines and the monitoring criteria. The mean score and SD of this domain was 60.5 ± 13.6, among which 3 TBI guidelines scored <50, with the lowest score of 39.1 from Brain Trauma Foundation (BTF, 2012).

Table 2

AGREE score by domain of each TBI guidelines.

Guideline ID	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5	Domain 6
CMA, 2007 ^[13]	69.6	50.7	59.4	73.2	56.3	57.4
RHSA, 2008 ^[14]	60.1	45.9	60.1	72.2	53.6	62.5
ACEP, 2009 ^[15]	83.3	45.6	76.7	86.3	61.9	80.2
SIGN, 2009 ^[16]	82.5	61.6	86.8	79.3	78.7	73.1
China, 2009 ^[17]	69.3	40.2	51.3	70.6	48.8	38.5
EFNS, 2011 ^[18]	65.3	36.5	55.9	57.6	40.3	65.4
NSW MoH, 2011 ^[19]	79.6	54.7	76.4	90.1	65.3	40.5
BTF, 2012 ^[20]	60.3	46.8	55.3	57.4	39.1	57.1
SCN, 2013 ^[21]	78.9	60.9	80.7	90.2	77.6	70.8
NIHCE, 2014 ^[22]	79.2	62.1	71.3	83.8	71.1	65.3
EAST, 2015 ^[23]	78.1	63.3	87.2	88.5	60.6	70.2
BTF/AANS, 2017 ^[24]	88.9	78.3	89.6	81.6	79.4	82.7
Mean score ± standard deviation (SD)	74.2 ± 9.09	54.6 ± 11.6	70.1 ± 13.6	78.4 ± 11.5	60.5 ± 13.6	61.7 ± 14.8
ICC (mean ± SD)	0.95 ± 0.02	0.75 ± 0.07	0.86 ± 0.05	0.81 ± 0.03	0.73 ± 0.04	0.81 ± 0.03

Domain 1, scope and purpose; Domain 2, stakeholder involvement; Domain 3, rigor of development; Domain 4, clarity of presentation; Domain 5, applicability; Domain 6, editorial independence.

3.3.7. Domain 6—editorial independence. This domain considers funders and competing interests of experts involved in guideline development. The mean score and SD of this domain was 61.7 ± 14.8 and 3 scored $< 50\%$. The score among these guidelines varied a lot, the lowest score of 38.5 points came from China (China, 2009) and the highest score was 82.7, from Brain Trauma Foundation and American association of Neurological Surgeons (BTF/AANS, 2017).

3.3.8. Overall assessment. This assessment concerns “the rating of body quality of the guidelines and whether the guideline would be recommended for use in practice.” According to the appraisal of the individual domains and overall scores, 10 TBI guidelines had > 4 domains scored > 50 , and rated as “recommended” by the appraisers (Table 2).

3.3.9. Level of evidence and strength of recommendation. Of the 12 included TBI guidelines, all of them were deemed evidence-based; 11 guidelines used 7 grading systems to rate the level of evidence and the strength of recommendation, among which 4^[21–24] adopted GRADE system (EAST, 2015; NIHCE, 2014; SCN, 2013; BTF/AANS, 2017), 2^[17,18] used SIGN system^[11] (China, 2009; SIGN, 2009), 1 used NHMRC system (NSW MoH, 2011),^[19] 1 used^[20] USPSTF system (BTF, 2012), 1^[14] used PNLG system (RHSA, 2008), 1^[18] used AAN system (EFNS, 2011) and 1^[15] used ACEP system (ACEP, 2009). However, the codes of level of evidence and strength of recommendation in different grading systems varied a lot (Table 3).

4. Discussion

This critical review investigated the quality of TBI guidelines published after 2006. Although there exist some TBI guidelines published before 2006,^[25–28] and still were not updated, the recommendations in the guidelines had been outdated and could not be used in practice according to IOM statements of clinical guidelines.^[7] Hence, we did not include these guidelines in this review.

Our study included 12 TBI guidelines; across all TBI guidelines, the highest mean scores were achieved in clarity and presentation, scope and purpose, and rigor of development, whereas the main weaknesses across TBI guidelines were stakeholder involvement, applicability, and editorial independence. Brain Trauma Foundation/American association of Neurological Surgeons (BTF/AANS, 2017), Scandinavian Neurotrauma Committee (SCN, 2013), and Eastern Association for the Surgery of Trauma (EAST, 2015) were the 3 CPGs with best results. All of the TBI guidelines evaluated in this study were developed by high-income countries (HICs) and are therefore minimally applicable in resource-limited settings. Besides, the distribution of level of evidence and strength of recommendations varied significantly among different TBI guidelines.

Overall, the strong scores in the clarity and presentation, scope and purpose, and rigor of development domains have been reported in other systematic reviews evaluating TBI CPGs.^[4,29–30] This is likely attributed to the scientific rigor of developing a guideline, which typically involves a highly methodical approach.^[31] In general, the guidelines that were more recently developed or updated, and those that had undergone numerous updates, most consistently demonstrated the highest quality by AGREE II scores.

Our analysis indicates an overall improvement in the above domains in the most current CPGs, consistent with other studies.^[21–24] In a 2011 critical review of mild TBI guidelines by Tavender et al,^[4] the NSW 2006 TBI guideline got a worse score of AGREE II in all domains with the exception of scope and purpose domain when compared with the updated 2011 version.^[19] It is noteworthy to mention that new version of TBI guidelines also has the advantage of newer and more rigorous evidence-based medicine in addition to the availability of designing guidelines around the AGREE II format. Nevertheless, a frequently criticized area in our results, within the rigor of development domain, was the lack of procedures for updating the guidelines for quality improvement. Given the trend toward improved guideline quality with newer revisions, development of a quality improvement list may help to ensure quality of future TBI guidelines.

The AGREE II result may also associate with reporting of key information in guidelines, which indicated more attention should be paid to improve the reporting quality of guidelines. In 2016, the AGREE working group developed a new checklist for improving the reporting quality of CPGs,^[32] which might be referred by TBI guidelines developers in the future.

Older reviews have demonstrated limited stakeholder involvement in TBI guideline development, a trend that persisted in new versions.^[4,33] Although there have been progressive improvement in guideline development, the domains of stakeholder involvement, applicability, and editorial independence remain weak, specifically when it comes to piloting interventions, addressing potential costs, barriers to implementation, and auditing for quality improvement. Recent studies suggested that successful implementation of guidelines could improve patient outcomes^[34–36]; however, applicability of guidelines to a given locale based on factors, such as availability and cost of resources, provider skills, and population needs and values, is critically important for successful implementation of guidelines in a manner that will improve patient care. Consideration of stakeholder involvement and applicability are imperative considering these domains are intrinsically associated with CPG implementation and translation to other settings such as LMICs.

It has been suggested that adaptation of existing guidelines to local situations may be a more valid and cost-effective means of achieving high-quality guidelines worldwide.^[37] However, the various codes of evidence level and strength of recommendation could bring challenges to reach the target. As we know, most guidelines, especially evidence-based guidelines, applied grading systems to communicate clear message, quickly and concisely so as to help guideline users, readers, and stakeholders to understand the confidence of estimate of the effects and the strength of recommendations easily. The confidence of estimate of the effects reflects the extent to which confidence in an estimate of the effect is adequate to support a particular recommendation. And the strength of guideline recommendation reflects the extent of collective confidence that adherence to the recommendation will do more good than harm.^[38,39] However, in our study, different grading systems with various systems of codes were used to communicate grades of evidence and recommendations in TBI guidelines, which could confuse the guideline users when using these guidelines. Therefore, a standardized grading system should be established to provide a clear information about the level of evidence and the strength of recommendation for TBI guidelines users, and the good news is that we find some guideline organizations such as Scottish Intercollegiate Guidelines

Table 3

Grading systems used in the included guidelines.

Grading systems	Institution/development group	Origin	Codes of evidence and recommendation				
			Level of evidence	Definition of evidence level	Strength of recommendation	Definition of recommendation strength	
GRADE	GRADE working group	International organization	A, B, C, D	<p>A. High quality: we are very confident that the true effect lies close to that of the estimate of the effect</p> <p>B. Moderate quality: we are moderately confident in the effect estimate—the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different</p> <p>C. Low quality: our confidence in the effect estimate is limited—the true effect may be substantially different from the estimate of the effect</p> <p>D. Very low quality: we have very little confidence in the effect estimate—the true effect is likely to be substantially different from the estimate of effect</p>	1, 2	<p>1. Strong recommendation: it is clear that the benefits outweigh the drawbacks. This means that all or virtually all, patients will want the recommended treatment.</p> <p>2. Weak recommendation: it is less clear whether the benefits outweigh the drawbacks. It still means that most patients will want the recommended treatment. There is however a greater opportunity for variation in individual preferences.</p>	EAST, 2015 NHCE, 2014 SCN, 2013 BTF/AANS, 2017
SIGN	Scottish Intercollegiate Guidelines Network	UK	1, 2, 3, 4	<p>1++ High-quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias</p> <p>1+ Well-conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias</p> <p>1- Meta-analyses, systematic reviews, or RCTs with a high risk of bias</p> <p>2++ High-quality systematic reviews of case control or cohort studies. High-quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal</p> <p>2+ Well-conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal</p> <p>2- Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal</p> <p>3- Non-analytic studies, e.g. case reports, case series</p> <p>4- Expert opinion</p>	A, B, C, D, E	<p>A. At least one meta-analysis, systematic review, or RCT rated as 1+++, and directly applicable to the target population; or a body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results</p> <p>B. A body of evidence including studies rated as 2+++, directly applicable to the target population, and demonstrating overall consistency of results; or extrapolated evidence from studies rated as 1++ or 1+</p> <p>C. A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++</p> <p>D. Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+</p>	China, 2009 SIGN, 2009
NHMRC	National Health and Medical Research Council	Australia	I, II, III-1, III-2, III-3, IV	<p>I. Evidence obtained from a systematic review of all relevant randomized control trials.</p> <p>II. Evidence obtained from at least one properly-designed randomized control trial</p> <p>III-1. Evidence obtained from well-designed pseudo-randomized controlled trials</p> <p>III-2. Evidence obtained from comparative studies with concurrent controls and allocation not randomized, cohort studies, case-control studies, or interrupted time series with a control group</p> <p>III-3. Evidence obtained from comparative studies with</p>	A, B, C, D	<p>A. Body of evidence can be trusted to guide practice in most situations</p> <p>B. Body of evidence can be trusted to guide practice in most situations</p> <p>C. Body of evidence provides some support for recommendation (s) but care should be taken in its application</p> <p>D. Body of evidence is weak and recommendation must be applied with caution</p>	NSW MoH, 2011

(continued)

Table 3
(continued).

Codes of evidence and recommendation							
Grading systems	Institution/development group	Origin	Level of evidence	Definition of evidence level	Strength of recommendation	Included guidelines	
USPSTF	US Preventive Services Task Force	USA	I, II, III	<p>historical control, 2 or more single arm studies or interrupted time series without a parallel control group</p> <p>IV. Evidence obtained from a case-series, either post-test or pre-test/post-test</p> <p>I. Good quality randomized controlled trial (RCT)</p> <p>II. Moderate quality RCT or Good quality cohort and case control study</p> <p>III. Case series, moderate or poor quality cohort and case control study</p>	A, B, C, D, I	<p>A. Strongly Recommended: The USPSTF strongly recommends that clinicians provide [the service] to eligible patients.</p> <p>B. Recommended: The USPSTF recommends that clinicians provide [the service] to eligible patients.</p> <p>C/D. No recommendation: the USPSTF makes no recommendation (C) for or against (D) routine provision of [the service].</p> <p>I. Insufficient Evidence to make a Recommendation: The USPSTF concludes that the evidence is insufficient to recommend for or against routinely providing [the service].</p>	BTF, 2012
PNLG	Programma Nazionale per le Linee Guida	Italy	I, II, III, IV, V, VI	<p>I. Evidence obtained from several randomized controlled clinical studies and/or systematic reviews of randomized studies</p> <p>II. Evidence obtained from a single suitably designed randomized study.</p> <p>III. Evidence obtained from coordinated non-randomized studies with simultaneous or historical controls or their meta-analysis</p> <p>IV. Evidence obtained from retrospective case-control type studies or their meta-analysis.</p> <p>V. Evidence obtained from case studies ("series of cases") without control group</p> <p>VI. Evidence based on the opinion of authoritative experts or committees of experts as indicated in guidelines or conference consensus, or based on the opinions of members of the working group responsible for these guidelines</p>	A, B, C, D, E	<p>A. The execution of that particular procedure or diagnostic test is strongly recommended.</p> <p>B. It indicates a particular recommendation supported by good quality scientific evidence, even if not necessarily of type I or II. Doubts exist regarding the fact that particular procedure or operation must always be recommended, but it is thought that its execution must be considered carefully.</p> <p>C. Substantial uncertainty exists in favor of or against the recommendation to execute the procedure or operation Execution of the procedure is not recommended</p> <p>D. Execution of the procedure is strongly advised against</p>	RHSA, 2008
ACEP	American College of Emergency Physicians	USA	I, II, III	<p>I. Randomized, controlled trial or meta-analyses of randomized trials; prospective cohort using a criterion standard; population prospective cohort</p> <p>II. Nonrandomized trial; retrospective observational; retrospective cohort and case control</p> <p>III. Case series; case report; consensus; review</p>	A, B, C	<p>A. Generally accepted principles for patient management that reflect a high degree of clinical certainty</p> <p>B. Recommendations for patient management that may identify a particular strategy or range of management strategies that reflect moderate clinical certainty</p> <p>C. Other strategies for patient management that are based on Class III studies, or in the absence of any published literature, based on panel consensus.</p>	ACEP, 2009

(continued)

Table 3
(continued).

Grading systems	Institution/development group	Origin	Codes of evidence and recommendation				
			Level of evidence	Definition of evidence level	Strength of recommendation		
AAN	American Academy of Neurology	USA	I, II, III, IV	<p>I. An adequately powered prospective, randomized, controlled clinical trial with masked outcome assessment in a representative population or an adequately powered systematic review of prospective randomized controlled clinical trials with masked outcome assessment in representative populations.</p> <p>II. Prospective matched-group cohort study in a representative population with masked outcome assessment that meets A–E above or randomized controlled trial in a representative population that lacks one criteria</p> <p>III. All other controlled trials (including well-defined natural history controls or patients serving as own controls) in a representative population, where outcome assessment is independent of patient treatment</p> <p>IV. Evidence from uncontrolled studies, case series, case reports, or expert opinion</p>	<p>A. rating (established as effective, ineffective, or harmful) requires at least one convincing class I study or at least 2 consistent, convincing class II studies</p> <p>B. rating (probably effective, ineffective, or harmful) requires at least one convincing class II study or overwhelming class III evidence</p> <p>C. (possibly effective, ineffective, or harmful) rating requires at least 2 convincing class III studies</p>	EFNS, 2011	
Not adopted grading system							CMA, 2007

Network (SIGN) and National Health and Medical Research Council (NHMRC) begin to adopt GRADE system instead of old systems in their new version of guideline development handbooks.^[40,41]

4.1. Strengths and limitations

Our overall findings have some strengths. First, our authors were from different background consisting of clinical experts and methodologists with extensive experience in evaluating clinical guidelines, which improved the reliability of our findings. Second, different domains have been appropriately weighed to derive overall assessment and recommendation. Nonetheless, our study also has limitations. Exclusion of guidelines published in languages other than English, or other forms (i.e., books, booklets, or government documents), might have resulted in under-representation of guidelines from less developed countries. Third, AGREE II instrument focused on methods of guideline development and the transparency of reporting, but could not assess potential impacts of recommendations on patient's outcomes.^[42,43] Furthermore, our study could not establish the causality between the poor performance and the characteristics of TBI guidelines, matching current guidelines to future guideline updates (in a cohort study) would allow for better assessment of guideline quality than did our cross-sectional assessment.

5. Conclusions

Most TBI guidelines got a high-quality rating. The high-quality domains were achieved in clarity and presentation, scope and purpose, and rigor of development. Our findings called for a standardized grading system to provide a clear information about the level of evidence and the strength of recommendation in TBI guidelines.

Acknowledgments

We thank Jinhui Tian and Kehu Yang (Evidence-Based Medicine Center of Lanzhou University) for supporting the methodology and Liang Yao (Hong Kong Baptist University) for providing assistance with editing the final article.

Author contributions

LAQ and WKP had full access to all of the data in the study and taken responsibility for the integrity of the data and the accuracy of the data analysis. CN and GTK contributed to drafting and critical revision of the article; DBS, WKP, and MWJ contributed to the data collection, AGREE process, and data analysis; DBS and GTK contributed to the study design and critical review. All authors contributed to the interpretation of study data and critically reviewed and approved the article before submission.

Conceptualization: Anqi Lu.

Data curation: Baoshan Di, Kongping Wei, Wenjuan Ma.

Methodology: Tiankang Guo.

Project administration: Yang Niu.

Resources: Qi Zhang.

Software: Kongping Wei, Wenjuan Ma.

Supervision: Anqi Lu, Hu Wang, Nong Cao, Tiankang Guo.

Writing – original draft: Baoshan Di.

Writing – review and editing: Baoshan Di, Kongping Wei, Wenjuan Ma, Tiankang Guo.

References

- [1] Hyder A, Wunderlich C, Puvanachandra P, et al. The impact of traumatic brain injuries: a global perspective. *NeuroRehabilitation* 2007;22:341–53.
- [2] Bruns J, Hauser W. The epidemiology of traumatic brain injury: a review. *Epilepsia* 2003;44:2–10.
- [3] Patel A, Vieira MMC, Abraham J, et al. Quality of the development of traumatic brain injury clinical practice guidelines: a systematic review. *PLoS One* 2016;11:e0161554.
- [4] Tavender E, Bosch M, Green S, et al. Quality and consistency of guidelines for the management of mild traumatic brain injury in the emergency department. *Acad Emerg Med* 2011;18:880–9.
- [5] Rusnak M, Mauritz W, Lecky F, et al. Evaluation of traumatic brain injury guidelines using AGREE instrument. *Bratisl Lek Listy* 2008;109:374–80.
- [6] Brouwers M, Kho M, Browman G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ* 2010;182:E839–42.
- [7] Steinberg E, Greenfield S, Wolman DM, et al. *Clinical Practice Guidelines We Can Trust*. Washington, DC: National Academies Press; 2011.
- [8] Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009;6:e1000097.
- [9] Chen YL, Yao L, Xiao XJ, et al. Quality assessment of clinical guidelines in China: 1993–2010. *Chin Med J* 2012;125:3660–4.
- [10] Yao L, Chen Y, Wang X, et al. Appraising the quality of clinical practice guidelines in traditional Chinese medicine using AGREE II instrument: a systematic review. *Int J Clin Pract* 2017;71:e12931.
- [11] Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
- [12] Alonso-Coello P, Irfan A, Sola I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care* 2010;19:e58.
- [13] Hebb M, Clarke D, Tallon J. Development of a provincial guideline for the acute assessment and management of adult and pediatric patients with head injuries. *Can J Surg* 2007;50:187–94.
- [14] Rusticali B, Villani R. Treatment of minor and severe traumatic brain injury. National reference guidelines. *Minerva Anestesiol* 2008;74:583–616.
- [15] Jagoda A, Bazarian J, Bruns J, et al. Clinical policy: neuroimaging and decisionmaking in adult mild traumatic brain injury in the acute setting. *J Emerg Nurs* 2009;35:e5–40.
- [16] Scottish Intercollegiate Guidelines Network. *Early Management of Patients With a Head Injury: A National Clinical Guideline*. 1st ed. Edinburgh: Scottish Intercollegiate Guidelines Network; 2009. Available at: <http://www.sign.ac.uk/pdf/sign110.pdf>. Accessed April 4, 2018.
- [17] Liao K, Chang C, Chang H, et al. Clinical practice guidelines in severe traumatic brain injury in Taiwan. *Surg Neurol* 2009;72:S66–73.
- [18] V Vos PE, Alekseenko Y, Battistin L, et al. Mild traumatic brain injury. *Eur J Neurol* 2012;19:191–8.
- [19] NSW Ministry of Health Adult Trauma Clinical Practice Guidelines: Initial Management of Closed Head Injury in Adults. 2nd ed. Sydney: NSW Institute of Trauma and Injury Management; 2011. 1–129.
- [20] Kochanek P, Carney N, Adelson P, et al. Guidelines for the acute medical management of severe traumatic brain injury in infants, children, and adolescents—second edition. *Pediatr Crit Care Med* 2012;13(Suppl. 1):S1–2.
- [21] Undén J, Ingebrigtsen T, Romner B. Scandinavian guidelines for initial management of minimal, mild and moderate head injuries in adults: an evidence and consensus-based update. *BMC Med* 2013;11:50.
- [22] National Institute for Health and Clinical Excellence: *Guidance. Head Injury Triage, Assessment, Investigation and Early Management of Head Injury in Infants, Children and Adults*. 1st ed. London: National Institute for Health and Clinical Excellence: Guidance; 2014 Available at: http://www.ncbi.nlm.nih.gov/books/NBK53036/pdf/Bookshelf_NBK53036.pdf. Accessed April 4, 2018.
- [23] Seamon MJ, Haut ER, Van Arendonk K, et al. An evidence-based approach to patient selection for emergency department thoracotomy: a practice management guideline from the Eastern Association for the Surgery of Trauma. *J Trauma Acute Care Surg* 2015;79:159–73.
- [24] Carney N, Totten AM, O'Reilly C, et al. Guidelines for the management of severe traumatic brain injury, fourth edition. *Neurosurgery* 2017;80:6–15.
- [25] Schutzman S, Barnes P, Duhaime A, et al. Evaluation and management of children younger than two years old with apparently minor head trauma: proposed guidelines. *Pediatrics* 2001;107:983–93.

- [26] Kamerling S, Lutz N, Posner J, et al. Mild traumatic brain injury in children: practice guidelines for emergency department and hospitalized patients. *Pediatr Emerg Care* 2003;19:431–40.
- [27] Piek J. Guidelines for the pre-hospital care of patients with severe head injuries. *Intensive Care Med* 1998;24:1221–5.
- [28] The management of minor closed head injury in children. Committee on Quality Improvement, American Academy of Pediatrics. Commission on Clinical Policies and Research, American Academy of Family Physicians. *Pediatrics* 1999;104:1407–15.
- [29] Berrigan L, Marshall S, McCullagh S, et al. Quality of clinical practice guidelines for persons who have sustained mild traumatic brain injury. *Brain Injury* 2011;25:742–51.
- [30] Grimmer K, Dizon J, Milanese S, et al. Efficient clinical evaluation of guideline quality: development and testing of a new tool. *BMC Med Res Methodol* 2014;14:63.
- [31] Qaseem A. The development of clinical practice guidelines and guidance statements of the American College of Physicians: summary of methods. *Ann Intern Med* 2010;153:194.
- [32] Irving G. The AGREE Reporting Checklist is useful for assessing the quality of clinical practice guideline development. *BMJ* 2016;353:i202711.
- [33] Alarcon J, Rubiano A, Chirinos M, et al. Clinical practice guidelines for the care of patients with severe traumatic brain injury. *J Trauma Acute Care Surg* 2013;75:311–9.
- [34] Lee J, Rittenhouse K, Bupp K, et al. An analysis of Brain Trauma Foundation traumatic brain injury guideline compliance and patient outcome. *Injury* 2015;46:854–8.
- [35] Fakhry S, Trask A, Waller M, et al. Management of brain-injured patients by an evidence-based medicine protocol improves outcomes and decreases hospital charges. *J Trauma* 2004;56:492–500.
- [36] Griesdale D, Örtengren V, Norena M, et al. Adherence to guidelines for management of cerebral perfusion pressure and outcome in patients who have severe traumatic brain injury. *J Crit Care* 2015;30:111–5.
- [37] Fervers B, Burgers J, Haugh M, et al. Adaptation of clinical guidelines: literature review and proposition for a framework and procedure. *Int J Qual Health Care* 2006;18:167–76.
- [38] Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
- [39] Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
- [40] Harbour RT, Forsyth L. Scottish Intercollegiate Guidelines Network-SIGN 50: A Guideline Developer's Handbook. Scottish Intercollegiate Guidelines Network 2014; Available at: <http://www.sign.ac.uk/guidelines/fulltext/50/>. Accessed April 3, 2018
- [41] NHMRC Guideline Handbook. Available at: <https://www.nhmrc.gov.au/guidelines-publications/information-guideline-developers/resources-guideline-developers>. Accessed April 3, 2018.
- [42] Vlayen J, Aertgeerts B, Hannes K, et al. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *Int J Qual Health Care* 2005;17:235–42.
- [43] Watine J, Friedberg B, Nagy E, et al. Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners. *Clin Chem* 2006;52:65–72.