Rating the quality of evidence: the GRADE system in systematic reviews/meta-analyses of AKI

Yuke Kong, Xuequan Wei, Lei Duan, Wenge Wang, Zhenmei Zhong, Zhenhua Ming & Rong Zeng

RENAL FAILURE

Taylor & Francis
Taylor & Francis Group

STATE OF THE ART REVIEWS

Rating the quality of evidence: the GRADE system in systematic reviews/meta-analyses of AKI

Yuke Kong[1]*, Xuequan Wei[2]*, Lei Duan[1], Wenge Wang[1], Zhenmei Zhong[3], Zhenhua Ming[3], and Rong Zeng[1]

[1]Department of Nephrology, Second Hospital of Lanzhou University, Lanzhou, Gansu, P.R. China, [2]Department of Nephrology, Linxia People's Hospital, Linxia, Gansu, P.R. China, and [3]The First Clinical Medical College of Lanzhou University, Lanzhou, Gansu, P.R. China

Abstract

Background: The method of systematic reviews/meta-analyses (SRs/MAs) has been widely used in acute kidney injury (AKI) studies. However, it is not quite clear about the quality of the evidence and existing problems. *Objectives*: To grade the evidence quality of published SRs/MAs of AKI by using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system, understand the current situation of evidence rating and analyze the possible problems. *Methods*: Researchers systematically searched for articles about SRs/MAs of AKI published in the following four Chinese databases and four English databases, including Chinese Biomedicine Literature Database, Wanfang Database, China National Knowledge Internet Database, VIP Database, Pubmed, EMBASE, the Cochrane Library and Web of Science. *Results*: Totally, 81 SRs/MAs were included in this study and the overall quality of evidence was not satisfactory. The number of literatures of low and very low evidence quality was 33 (40.7%) and 41 (50.6%), respectively. Limitation was the main factor which caused the quality of research evidence degrading (92.6%), and other degradation factors were inconsistency (56.8%), publication bias (44.4%), indirectness (35.8%) and imprecision (32.1%). The quality of evidence for AKI has been significantly improved after the publication of the GRADE system in 2004. *Conclusions*: Since 2004 when the GRADE system was published, the quality of evidence of AKI has been increased clearly. But quality of AKI evidence of SRs/MAs for intervention is still not satisfactory. Limitation and inconsistency were two major factors leading to degradation.

Introduction

Acute kidney injury (AKI) which is characterized by acute onset and severe condition can cause patients renal failure in a short time. According to foreign reports, the overall fatality rate of AKI is as high as 26.5–45.0%. So, AKI has been a common emergency in clinical departments.[1] Most cases of AKI are reversible in the early stages. It is necessary to take emergency measures to reverse the causes of renal dysfunction, maintain stability of the systemic circulation as far as possible and ensure adequate renal blood flow.[2] Therefore, whatever the causes of AKI are, it is very important to prevent early, diagnose early, correct the fundamental etiology timely and treat early. Although renal replacement therapy at an early phase can remarkably improve survival rates of patients with AKI,[3] many factors may influence the outcomes of AKI patients, such as the first time of dialysis treatment,[4] dose of dialysis,[5] the property of dialysis membrane[6] and so on. As a consequence, how to prevent the occurrence and development of AKI has become a focus and hotspot of current kidney disease researches. At present, there are still many problems existing in diagnosis and treatment of AKI. It is especially important for diagnosis and treatment of AKI that multiple disciplines cooperate closely, but an effective multidisciplinary intervention mode has not developed yet.[7] Currently, systematic reviews/meta-analyses (SRs/MAs) have been widely used in AKI studies and provided high quality evidence of comparison between different interventions for AKI. Nevertheless, it is still unclear that what the quality of evidence of SRs/MAs for AKI present situation and existing problems are.

The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system, developed by a broadly representative formulation team of international guidelines, has been one of the international standards in rating quality of evidence and strength of recommendations currently.[8] It contains specific rules and integrated standards regarding upgrade and downgrade of different levels evidence, and offers a transparent and structured process. GRADE focuses on patients' value and willingness, explicitly interprets strength of recommendations from multiple perspectives, and is suitable for systematic reviews, health technology assessments and guidelines. The GRADE system results in an

---

*These authors are equal contributors.

Address correspondence to Rong Zeng, Department of Nephrology, Second Hospital of Lanzhou University, No. 82 Cuiyingmen, Lanzhou 730000, Gansu, P.R. China. Tel: +86 0931 8943909; Fax: +86 0931 8943909; E-mail: zengrong15@126.com

assessment of the quality of a body of evidence as high, moderate, low, or very low, and classifies the strength of recommendations as strong or weak.[9,10] This study aimed to evaluate the quality of evidence of SRs/MAs for AKI using the GRADE system. Therefore, we expected to show the status quo of AKI researches, analyze the possible existing problems and provide objective evidence for clinical physicians to treat disease.

Methods

Search strategy

We systematically searched the following electronic databases for SRs/MAs: PubMed (1966–2013.5), EMBASE (1974–2013.7), the Cochrane Library (–2013.5), Web of Science (–2013.7), Chinese Biomedical Literature Database (1978–2013.5), VIP Database (1994–2013.5), China National Knowledge Internet Database (1989–2013.5), and Wanfang Database (1982–2013.5). The main search terms included ''acute kidney failure'', ''acute renal failure'', ''ARF'', ''acute renal insufficiency'', ''acute kidney insufficiency'', ''acute renal injury'', ''acute kidney injury'', ''acute kidney tubular necrosis'', ''acute tubular necrosis'', ''systematic review'' and ''meta-analysis''. The search strategy is presented in Table 1.

Inclusion/exclusion criteria

The inclusion criterion was as follows: Chinese or English SRs/MAs were about prevention and treatment of AKI high-risk groups or patients.

The exclusion criteria were as follows: (i) Systematic reviews included randomized controlled trials (RCTs) and observational studies at the same time, and data of RCTs cannot be extracted alone; (ii) articles were in the state of protocol; and (iii) a paper which submitted more than once excluded Chinese version.

Screening

Two independent reviewers primarily screened titles and abstracts of the studies according to the inclusion and exclusion criteria designed in advance. Then, full text of potentially proper articles was retrieved by the above two reviewers for further assessment. Disagreements were resolved by consensus, and a third author would act as an adjudicator if needed.

Table 1. Search strategy.

#1. ''acute kidney failure''/exp
#2. ''acute kidney tubule necrosis''/exp
#3 ''acute renal failure'' or ''acute kidney failure'' or ARF
#4. ''acute renal insufficiency'' or ''acute kidney insufficiency''
#5. ''acute renal injury'' OR ''acute kidney injury''
#6. ''acute tubular necrosis''
#7. ''acute kidney tubular necrosis''
#8. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7
#9. ''Meta analysis''/exp
#10. ''systematic review''/exp
#11. ''Meta analysis'' OR ''systematic review''
#12. #9 OR #10 OR #11
#13. #8 AND #12

Data extraction

A data extraction form was designed in advance according to the research objectives and contents, and was revised as per preliminary experiments. Two researchers independently extracted and checked the published data which consisted of general data and GRADE data. General information contains journal name, publication time, time span of retrieval, follow-up period, first author, funding source, competing interests, number of included study, intervention and outcome. GRADE information contains main indicators necessary of evaluation, number of included study, outcome (OR/RR, 95% CI), downgrade factors and quality of evidence. Disagreements were resolved by discussion, and a third reviewer would act as an adjudicator if needed.

Quality assessment

The GRADE system classifies the quality of evidence at one of four levels: high, moderate, low, and very low based on five downgrade factors: limitations, inconsistency, indirectness, imprecision and publication bias.[11–15] ''High'' means we are very confident that the true effect lies close to that of the estimate of the effect. ''Moderate'' means 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. ''Low'' means our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect. ''Very low'' means we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.[10]

Statistical analysis

All statistical analyses were performed using SPSS version 13.0 (SPSS Inc., Chicago, IL) (http://www.spss.com). For continuous data, we used mean $\pm$ SD ($\bar{x} \pm s$) to describe and adopted t test or variance analysis. Dichotomous data were summarized with descriptive statistical analysis (frequency and percentage) and χ^2 test. Statistical significance was set at $p < 0.05$.

Results

Search

In the initial search, we identified 1629 potentially relevant articles including 48 Chinese ones and 1581 English ones. By means of reading titles and abstracts, 1478 were excluded due to duplication, non-clinical trial or purpose unconformity. In the further assessment via reading full text according to the inclusion and exclusion criteria, 70 papers were excluded. Finally, we incorporated 81 articles including 79 English and 2 Chinese ones, as the flow chart of the literature selection shows in Figure 1.

Baseline information

Among all the included literatures which were published between 2001 and 2013, 13 (16.0%) and 68 (84.0%) were published. Respectively. before and after the release of GRADE system. There are 56 (69.1%) SRs/MAs concerning

AKI prevention and the rest (25, 30.9%) were about AKI therapy. As to titles, 43 (53.1%) and 11 (13.6%) articles were, respectively, named after meta-analysis and systematic review, while 17 (21.0%) and 10 (12.3%), respectively, named after both systematic review and meta-analysis or neither. Literature authors come from 14 countries and first authors of 31 (38.3%) articles are from the USA. The percentage of articles that reported funding source was 38.3% (31/81). Only 17 (21.0%) SRs/MAs of AKI had follow-up records.

Quality of evidence

Eventually, 81 articles were classified into four levels according to the evidence quality grade. The number of articles assessed as high, moderate, low and very low were 0 (0%), 7 (8.6%), 33 (40.7%) and 41 (50.6%). At present, the overall quality of SRs/MAs of AKI is low.

Factors rating down the quality of evidence

Data have revealed that among the five factors rating down the quality of evidence, study limitation is the principal factor (75/81, 92.6%) and inconsistency is the next one (46/81, 56.8%). The rate of imprecision that can downgrade the quality of evidence is the least (26/81, 32.1%) (Table 2).

Analysis of quality of evidence before and after release of the GRADE system

To investigate whether the publication of the GRADE system was associated with an improvement in evidence quality (pre-GRADE vs. post-GRADE), the period $\leq$2004 was compared with >2004. Table 3 shows that, following release of the GRADE system, the quality of evidence has improved significantly. By contrast, there were fewer studies and no high or moderate literatures which were assessed as very low before 2004. We found that inconsistency resulted in fewer articles downgrading after 2004 when compared with those published prior to or in 2004 (48.5% vs. 100%, $p < 0.01$); whereas among the former articles, there existed literature downgrading two grades. For the rest of the four factors, no
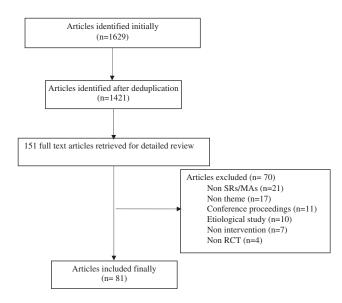
statistical difference was showed between SRs/MAs published before 2004 and those after 2004 (Table 4).

Analysis of AKI prevention and treatment

As presented in Table 3, high, moderate, low and very low articles of AKI prevention accounted for 0 (0%), 4 (7.1%), 24 (42.9%) and 28 (50.0%), respectively. Meanwhile, those figures of AKI treatment articles were 0 (0%), 3 (12.0%), 9 (36.0%) and 13 (52.0%), respectively. As to factors that lead to rating down the quality of evidence, limitation and inconsistency are still the main factors affecting AKI prevention and treatment SRs/MAs. In addition, more articles of AKI prevention downgrade because of indirectness (48.2%) (Table 4).

Discussion

The quality of SRs/MAs of AKI was generally low

SRs/MAs of AKI are one of the most important clinical evidence sources for a physician to practice making evidence-based decisions. So, rating quality of evidence, methodological quality and reporting quality of SRs/MAs of this filed will affect recommendation strength and utility of the clinical evidence. In this study, we systematically searched English and Chinese SRs/MAs for AKI published before May 2013, and evaluated the quality of evidence based on the GRADE system. The results showed that quality of SRs/MAs about AKI intervention was not satisfactory. The ratio of quality of evidence for 81 SRs/MAs assessed as high, moderate, low and very low were 0:1:5:6. It was the main reason for resulting in low quality of evidence that SRs/MAs did not include complete original data, and the included RCTs contained some methodological defects.

Limitation and inconsistency were the main degradation factors

Limitation was the most principal factor that downgrades quality of evidence of SRs/MA for AKI. Main outcome indicators of 92.7% SRs/MAs were downgraded because of limitation. What is more, the proportion of SRs/MAs of AKI for treatment degraded by limitation was as high as 96.0%. According to the analysis, main reasons were that the number of original material for AKI was less, and blind method or allocation concealment was inadequate. Another important degradation factor of SRs/MAs of AKI for treatment was inconsistency, which caused 56.8% literatures downgrading. The factors leading to high inconsistency were that main outcome indicators and index of evaluating validity for SRs/MAs in AKI field were not unified, and the basis of research object was different.

Figure 1. Flow chart of articles selection procedure.

Table 2. Factors rating down quality of evidence of AKI systematic reviews *n* (%).

Factors	No severe	Severe	Very severe
Limitation	6 (7.4)	68 (84.0)	7 (8.6)
Inconsistency	35 (43.2)	44 (54.3)	2 (2.5)
Indirectness	52 (64.2)	29 (35.8)	0 (0)
Imprecision	55 (67.9)	24 (29.6)	2 (2.5)
Publication bias	45 (55.6)	35 (43.2)	1 (1.2)

Table 3. Quality of evidence of AKI systematic reviews.

Quality of Evidence	Published year		Category		Total
	≤2004	>2004	Prevention	Treatment	
High	0 (0)	0 (0)	0 (0)	0 (0)	0 ((0)
Moderate	0 (0)	7 (10.3)	4 (7.1)	3 (12.0)	7 (8.6)
Low	1 (7.7)	32 (47.1)	24 (42.9)	9 (36.0)	33 (40.7)
Very low	12 (92.3)	29 (42.6)	28 (50.0)	13 ((52.0)	41 (50.6)
Research quantity	13 (100)	68 (100)	56 (100)	25 (100)	81 (100)

Table 4. Subgroup analysis of factors rating down quality of evidence *n* (%).

	≤2004			>2004			Prevention			Treatment		
	No severe	Severe	Very severe	No severe	Severe	Very severe	No severe	Severe	Very severe	No severe	Severe	Very severe
Limitation	1 (7.7)	11 (84.6)	1 (7.7)	5 (7.4)	57 (83.8)	6 (8.8)	5 (8.9)	49 (87.5)	2 (3.6)	1 (4.0)	19 (76.0)	5 (20.0)
Inconsistency	0 (0)	13 (100)	0 (0)	35 (51.5)	31 (45.6)	2 (2.9)	24 (42.9)	32 (57.1)	0 (0)	11 (44.0)	12 (48.0)	2 (8.0)
Indirectness	7 (53.8)	6 (46.2)	0 (0)	45 (66.2)	23 (33.8)	0 (0)	29 (51.8)	27 (48.2)	0 (0)	23 (92.0)	2 (8.0)	0 (0)
Imprecision	9 (69.2)	4 (30.8)	0 (0)	46 (67.7)	20 (29.4)	2 (2.9)	41 (73.2)	15 (26.8)	0 (0)	14 (56.0)	9 (36.0)	2 (8.0)
Publication bias	5 (38.5)	8 (61.5)	0 (0)	40 (58.8)	27 (39.7)	1 (1.5)	28 (50.0)	27 (48.2)	1 (1.8)	17 (68.0)	8 (32.0)	0 (0)

Indirectness was one of the most important degradation factors of SRs/MAs of AKI for prevention

In addition to limitation and inconsistency, analysis showed that indirectness was another major degradation factor for the SRs/MAs of prevention. Major outcome indicators of AKI for prevention were that level of serum creatinine rose about 26.5 mmol/L (0.3 mg/dL) within 48 h or increased 50% than original level, or on the basis of volume of urine reduced to less than 0.5 mL/(kg h) for over 6 h.[16] Replacing clinical manifestation with biochemical test results as outcome indicator has significantly increased the risk of degradation by indirectness for the prevention group.

Suggestions for future researches

First of all, the methodological quality of RCT about AKI should be improved in the future. Researchers should register in a clinical trial registry, report the results according to the Consolidated Standards of Reporting Trials statement, reduce bias as far as possible, and improve the authenticity of AKI clinical trials. Second, when choose main outcome indicators, researchers should reduce the substitution of biochemical test results for outcome indicator, so as to decrease the risk of degradation caused by indirectness. Third, SRs/MAs should be strictly designed and produced according to the manual of Cochrane systematic reviews to improve quality.

Advantages and disadvantages

Our study showed the following two advantages. First, this is the first research to rate quality of evidence for SRs/MAs of AKI by using the GRADE system. It revealed the current situation of AKI studies and provided a reference basis for clinical doctors to prevent and treat diseases. Second, GRADE is a scientific and effective method to assess the quality of evidence. This research made a scientific evaluation about the quality of evidence for each study based on the requirements of GRADE guideline. There were some limitations in our study at the same time. First, it only included English and Chinese articles and evaluated RCTs, which cannot fully represent the overall quality of evidence. Besides, it rated quality of evidence mainly about major measurement indicators of SRs/MAs for AKI, while the secondary measurement indicators were not assessed. Then, last but not least, it only performed qualitative evaluation, so SRs/MAs of the same level in quality of evidence may still contain differences.

References

1. Liu S. *Analysis of Clinical and Pathological Features in 236 Cases Drug Associated Acute Kidney Injury.* Changchun, P.R. China: Jilin University; 2011.
2. Ma XC, Li X. Advance in definition, prevention and treatment of acute kidney injury. *Chinese J Pract Int Med.* 2012;32(6):408–411.
3. Duan SB, Zhang H, Peng YM. The pathogenesis and prevention of acute kidney injury. *Chinese J Blood Purif.* 2010;9:349–335.
4. Gettings LG, Reynolds HN, Scalea T. Outcome in post-traumatic acute renal failure when continuous renal replacement therapy is applied early vs. late . *Intensive Care Med.* 1999;25:805–813.
5. Paganini EPTM, Goormastic M, Halstenberg W, et al. Establishing a dialysis therapy/patient outcome link in intensive care unit acute dialysis for patients with acute renal failure. *Am J Kidney Dis.* 1999;28(Suppl 3):S81–S89.
6. Schiffl H, Lang SM, Konig A, et al. Biocompatible membranes in acute renal failure: prospective case-controlled study. *Lancet.* 1994; 344:570–572.
7. Chen N, Zhang W. Clinical research status of acute kidney injury in China. *Clin Nephrol.* 2013;13(1):4–6.
8. http://www.gradeworkinggroup.org/publications/JCE_series.htm. Accessed Aug 19, 2015.

9. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 4. GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64:383–394.

10. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011; 64(4):401–406.

11. Guyatt GH, Oxman AD, Vist GE, et al. GRADE guidelines: 4. Rating the quality of evidenced study limitations (risk of bias). *J Clin Epidemiol*. 2011;64:407–415.

12. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol*. 2011;64(12):1294–1302.

13. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence·indirectness. *J Clin Epidemiol*. 2011;64(12):1303–1310.

14. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 6. Rating the quality of evidence—imprecision (random error). *J Clin Epidemiol.*. 2011;64(12):1283–1293.

15. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol.*. 2011;64(12):1277–1282.

16. Molitoris BA, Levin A, Warnock DG, et al. Improving outcomes of acute kidney injury: report of an initiative. *Nat Clin Pract Nephrol*. 2007;3(8):439–442.