Summary of Findings Tables for Joanna Briggs Institute Systematic Reviews

Originally developed by the Joanna Briggs Institute Levels of Evidence and Grades of Recommendation Working Party April 2014*

Updated: January 2016

Contents

Introduction

There has been a push internationally to adopt the approach of the GRADE (Grading of Recommendations Assessment, Development and Evaluation) working group[1] who have developed a grading of evidence and recommendation system that has been endorsed by many evidence-based healthcare organizations, including Cochrane, WHO, AHRQ, NICE, BMJ Clinical Evidence and SIGN, amongst others.

The approach of GRADE is not to classify findings based solely on study design but to consider other factors as well. The GRADE working group have developed a process to establish confidence in the synthesized results of quantitative research through considering issues related to risk of bias, publication bias, inconsistency, indirectness, imprecision of evidence, effect sizes, dose-response relationships, and confounders of findings.[2] The evidence is then ranked into one of four levels (High, Moderate, Low, Very Low). This process begins with findings being assigned a pre-ranking based on their design (High = randomized controlled trials (RCTs), Low = observational studies), and then downgraded or upgraded based on the aforementioned factors. A new, more nuanced ranking can then be assigned to an individual finding or outcome. In this way, evidence from observational studies can be ranked above that of RCTs where appropriate. This score is then applied to the major results of a quantitative systematic review. Key findings and important supporting information is presented in a 'Summary of Findings' table (or evidence profile) within the systematic review. These 'Summary of Findings' tables have been shown to improve understanding and accessibility of the results of systematic reviews.[3-6]

The Joanna Briggs Institute (JBI) and its collaborating entities have recently decided to adopt the GRADE approach for systematic reviews of effectiveness. However, the JBI do not conduct reviews pertaining to effectiveness alone, and have developed methodology for conducting reviews of qualitative research and text and opinion. To date, there has been no widely accepted approach to assist health care professionals and policy makers in establishing confidence in the synthesized findings of qualitative systematic reviews and to develop summary of findings for these types of reviews. To address this, a working party was established within the JBI to develop a system for establishing the confidence in the synthesized findings of qualitative research and to present this in a Summary of Findings table.

This document outlines the methods for creating a Summary of Findings table for effectiveness, qualitative and text and opinion reviews.

PLEASE NOTE: As of January 2016, all new JBI systematic reviews of effectiveness or qualitative evidence should include a summary of findings table. These should appear underneath the executive summary in JBI systematic reviews in the appropriate format, following the implications for research.

Effectiveness Reviews

Systematic reviews should be accompanied by a Summary of Findings table.[6] The Summary of Findings table should include the question being investigated, the population, intervention and comparison, the outcomes assessed, estimated risk or odds for categorical data or weighted means for continuous data, relative effect, sample size as well as the number of studies which contributed to the sample, the GRADE quality of evidence for each finding, and any comments (including decisions as to why the reviewers assigned the final GRADE ranking). These Summary of Findings tables can be created using the web application GRADEPro GDT (http://gradepro.org/) and should appear underneath the executive summary in JBI systematic reviews in the appropriate format, following the implications for research.

To determine a GRADE quality of the evidence, the GRADE approach begins by assigning findings to one of two starting levels of quality depending on the study design. Initially, randomized trials are high quality, while observational studies are low quality. Additionally, two other levels exist; moderate and very low. This gives four levels: High, Moderate, Low and Very low. Studies can then be up- or downgraded based on certain factors.[1]

Factors that should lead to downgrading are: Risk of bias (as determined by the JBI MASTARI critical appraisal form; -1 if serious risk of bias, -2 if very serious risk of bias), Inconsistency or heterogeneity of evidence (-1 if serious inconsistency, -2 if very serious inconsistency), Indirectness of evidence (-1 if serious, -2 if very serious), Imprecision of results (-1 if wide confidence interval, -2 if very wide confidence interval) and Publication bias (-1 if likely, -2 if very likely).[6, 7 8 9]

Factors that should lead to upgrading are: Large magnitude of effect (+1 level if a large effect, +2 if a very large effect), Dose response (+1 level if there is evidence of a gradient), All plausible confounding factors would reduce the demonstrated effect (+1 level) or create a spurious effect where results suggest no effect (+1 level).[6-9]

Table 1: Summary of Findings Template

Summary of findings:

Very Tight glycemic control in both diabetic and nondiabetic patient compared to for

Patient or population:

Setting:

Intervention: Very Tight glycemic control in both diabetic and nondiabetic patient

Comparison:

Outcomes	Anticipated absolute effects[*] (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)
	Risk with	**Risk with Very Tight glycemic control in both diabetic and nondiabetic patient**			
All cause mortality (Mortality) assessed with: Number of death follow up: range 6-60 days to	**Study population**		**OR 0.59** (0.37 to 0.96)	1729 (5 RCTs)	⊕◯◯◯ VERY LOW [1,2,3]
	52 per 1000	**32 per 1000** (20 to 50)			
Length of stay in ICU(in days)	The mean length of stay in ICU(in days) was **0**	The mean length of stay in ICU(in days) in the intervention group was 0.07 more (0.01 fewer to 0.15 more)	-	590 (3 studies)	⊕◯◯◯ VERY LOW [1,2,3]

*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **OR:** Odds ratio; **MD:** Mean difference

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
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
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
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

1. No blinding of investigators
2. indicated substantial heterogeneity
3. wide confidene intervals

JBI endorses GRADEPro GDT for the development of Summary of Findings tables. All Summary of Findings tables created for JBI effectiveness reviews must use the GRADEPro GDT web application.

When logging in to GRADEPro GDT, there are different format options. JBI reviews must use the Summary of Findings (SoF) table option (Figure 1).
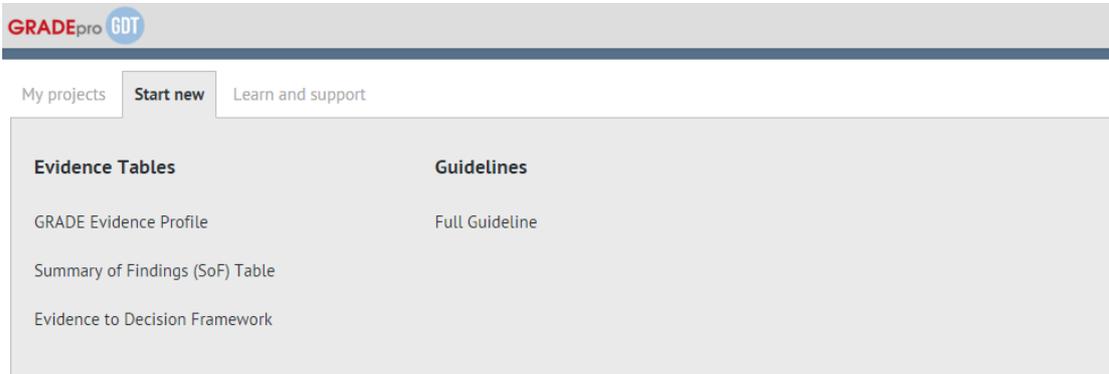
Figure 1: Options within GRADEPro GDT

When submitting to the JBISRIR, SoF tables should be in exported in portrait view with the default SoF format (Figure 2)
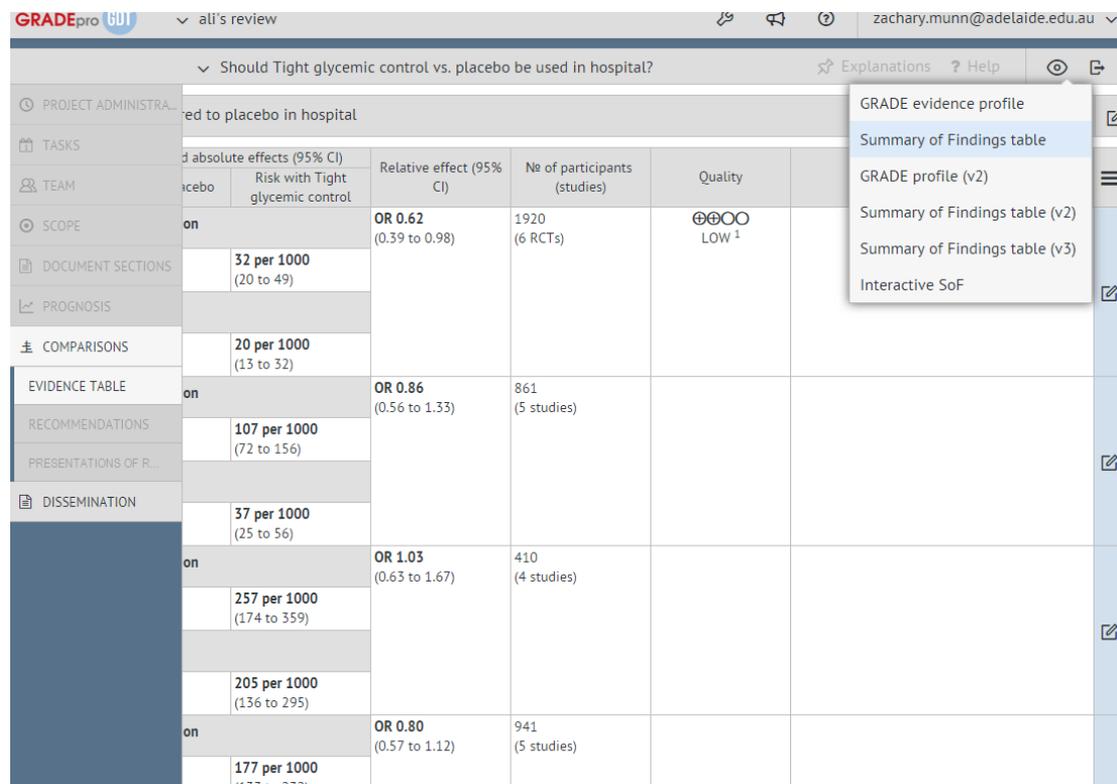
Figure 2: Summary of findings option

Detailed guidance and help for creating a Summary of Findings table is available within the GRADEPro software. (http://gradepro.org/)

Further information regarding the development of Summary of Findings tables and GRADE can be found at the following websites:

http://www.gradeworkinggroup.org/index.htm

http://tech.cochrane.org/revman/other-resources/gradepro/download

Qualitative and Text and Opinion Reviews

Within GRADE, findings are given a pre-ranking of high (for RCTS) or low (for observational studies). As it is not appropriate to distinguish between different qualitative study designs (for example a phenomenological study or an ethnographic study) via a hierarchy, in this system all qualitative research studies start off as 'high' (on a ranking scale of High, Moderate, Low to Very Low). Expert opinion is pre-ranked at low.

This ranking system then allows synthesized findings to be downgraded based on their dependability and credibility. Downgrading for dependability may occur when the appraisal criteria relevant to dependability are not met (a subset of criteria from the JBI-QARI/ NOTARI critical appraisal checklist).

Five questions of the JBI-QARI checklist [10] are viewed as specifically relating to the concept of dependability in qualitative research. These are:

- Is there congruity between the research methodology and the research question or objectives?
- Is there congruity between the research methodology and the methods used to collect data?
- Is there congruity between the research methodology and the representation and analysis of data?
- Is there a statement locating the researcher culturally or theoretically?
- Is the influence of the researcher on the research, and vice-versa, addressed?

Five questions of the JBI-NOTARI checklist were viewed as specifically relating to the concept of dependability in text and opinion.

- Does the source of the opinion have standing in the field of expertise?
- Is the opinion's basis in logic/experience clearly argued?
- Is the argument developed analytical?
- Is there reference to the extant literature/evidence and any incongruence with it logically defended?
- Is the opinion supported by peers?

If 4-5 of the responses to these questions are yes, the synthesized finding remains at the level it is currently. If 2-3 of these responses are yes, it moves down one level (i.e. from High to Moderate). If 0-1 of these responses are yes, it moves down two levels (from High to Low, or Moderate to Very Low).

Downgrading for credibility may occur when not all the findings included in a synthesis are considered unequivocal.

Unequivocal: findings accompanied by an illustration that is beyond reasonable doubt and; therefore not open to challenge

Credible: findings accompanied by an illustration lacking clear association with it and therefore open to challenge

Unsupported: findings are not supported by the data.

A synthesized finding can be made up of unequivocal, credible, or not supported findings, or a mixture of all. If synthesized findings come from only unequivocal findings, it can remain where it is on the ranking system. For a mix of unequivocal/credible findings, the synthesized finding can be downgraded one (-1). For credible findings, the synthesized finding can be downgraded two (-2). For credible/unsupported findings, it can be downgraded three (-3), and for not-supported findings, it can be downgraded four (-4).

The proposed system would then give an overall score of High, Moderate, Low to Very Low. This ranking can be considered a rating of 'confidence' in the qualitative synthesized finding, a process we have called 'ConQual' for short.

The Summary of Findings table includes the major elements of the review and details how the ConQual score is developed. Included in the table is the title, population, phenomena of interest and context for the specific review. Each synthesized finding from the review is then presented along with the type of research informing it, a score for dependability, credibility, and the overall ConQual score. The type of research column (i.e. qualitative) has been included to stress to users who are more familiar with quantitative research that this is coming from a different source. The Summary of Findings table has been developed to clearly convey the key findings to a reader of the review in a tabular format, with the aim being to improve the accessibility and usefulness of the systematic review (Table 1).

Table 1: ConQual Summary of Findings Example

Systematic review title: The patient experience of high technology medical imaging: a systematic review of the qualitative evidence

Population: Persons who had undergone high technology medical imaging

Phenomena of interest: The meaningfulness of a patients experience of undergoing diagnostic imaging using high technology

Context: Male and Female Adult Patients presenting to a medical imaging department

Synthesized Finding	Type of research	Dependability	Credibility	ConQual Score	Comments
People undergoing imaging often expect a health issue to be found during their scan, which can then lead to anxiety and worry	Qualitative	Downgrade 1 level*	Downgrade 1 level **	**Low**	*Downgraded one level due to dependability of primary studies **Downgraded one level due to equivocal findings

Further information:

Munn Z, Porrit K, Lockwood C, Aromataris E, Pearson A. Establishing confidence in the output of qualitative research synthesis: the ConQual approach. BMC Med Res Methodol. 2014;14:108. http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-14-108

How to cite this document:

The Joanna Briggs Institute Levels of Evidence and Grades of Recommendation Working Party*. Summary of Findings Tables for Joanna Briggs Institute Systematic Reviews. The Joanna Briggs Institute. 2016. (website address)

The Levels of Evidence and Grades of Recommendation Working party consists of:

Dr Zachary Munn, Senior Research Fellow, Implementation Science, The Joanna Briggs Institute, School of Translational Health Science, The University of Adelaide.

Dr Kylie Porritt, Research Fellow, Implementation Science, The Joanna Briggs Institute, School of Translational Health Science, The University of Adelaide.

Associate Professor Edoardo Aromataris, Director Synthesis Science, The Joanna Briggs Institute, School of Translational Health Science, The University of Adelaide.

Associate Professor Craig Lockwood, Director Implementation Science, The Joanna Briggs Institute, School of Translational Health Science, The University of Adelaide.

Dr Micah Peters, Research Fellow, Synthesis Science, The Joanna Briggs Institute, School of Translational Health Science, The University of Adelaide.

References

1   GRADE Working group. Education and debate - grading quality of evidence and strength of recommendations. BMJ. 2004;328(1490).

2   Goldet G, Howick J. Understanding GRADE: an introduction. Journal of evidence-based medicine. 2013;6(1):50-4.

3   Langendam MW, Akl EA, Dahm P, Glasziou P, Guyatt G, Schunemann HJ. Assessing and presenting summaries of evidence in Cochrane Reviews. Systematic reviews. 2013;2(1):81.

4   Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. J Clin Epidemiol. 2010;63(6):620-6.

5   Treweek S, Oxman AD, Alderson P et al. Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE): protocol and preliminary results. Implementation science : IS. 2013;8:6.

6   Guyatt G, Oxman AD, Akl EA et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383-94.

7   Balshem H, Helfand M, Schunemann HJ et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64(4):401-6.

8   Guyatt GH, Oxman AD, Santesso N et al. GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. J Clin Epidemiol. 2013;66(2):158-72.

9   Guyatt GH, Thorlund K, Oxman AD et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. J Clin Epidemiol. 2013;66(2):173-83.

10   Hannes K, Lockwood C, Pearson A. A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. Qualitative health research. 2010;20(12):1736-43.