

Appendix B

Guideline grading systems

1. Evidence and recommendation grading system (excluding complementary and alternative medicine)

Studies were graded using a two-tier system that is detailed in the *Handbook for the Preparation of Explicit Evidence-Based Clinical Practice Guidelines*, published in November 2001 by NZGG.³⁷⁰ This system has been adapted from other grading systems currently in use, in particular the SIGN system.⁸¹

The searches for this guideline concentrated on finding high grade evidence to answer the identified clinical questions, such as systematic reviews, randomised controlled trials and, where these were not available, observational studies such as well designed cohort and case control studies. Only these types of study design were graded. Where these types of study were not available, less rigorous study designs such as cross-sectional studies and case studies were considered but were not formally graded.

The two-tier system follows this process.

1. Critical appraisal of individual relevant studies (identified from the searching) and assigning of **a level of evidence** for the first section of the GATEFRAME checklist that is incorporated into the evidence tables. A random sample of appraisals in the guideline were performed independently by two assessors and the results compared.
2. Joint consensus by the Guideline Development Team on the issues of volume, consistency, clinical relevance and applicability of the body of evidence in the evidence table (filling out the NZGG Considered Judgement form for each clinical question) and the development of **graded recommendations** that attempt to answer the clinical questions posed.

Levels of evidence

There are three levels of evidence that can be assigned to the Validity section of the GATEFRAME (Section 1):

- + strong study where all or most of the validity criteria are met
- ~ fair study where not all the validity criteria are met, but the results of the study are not likely to be influenced by bias
- x weak study where very few of the validity criteria are met and there is a high risk of bias.

Developing recommendations

Recommendations were formulated by joint meetings of the multidisciplinary Guideline Development Team. The group considered the entire body of evidence (summarised in the evidence tables) and filled out Considered Judgement forms for each clinical question that was identified as being relevant to the guideline (see www.nzgg.org.nz). The following aspects were discussed: volume of evidence, applicability to the New Zealand setting, consistency and clinical impact, with the aim of achieving consensus. Consensus was sought and achieved over the wording of the recommendation and grading. In this guideline, where a recommendation is based on the clinical experience of members of the Guideline Development Team, this is referred to as a good practice point.

Grading of recommendations

The NZGG grades of recommendation are as follows:

RECOMMENDATIONS	GRADE
The recommendation is supported by good evidence (where there are a number of studies that are valid, consistent, applicable and clinically relevant).	A
The recommendation is supported by fair evidence (based on studies that are valid, but there are some concerns about the volume, consistency, applicability and clinical relevance of the evidence that may cause some uncertainty but are not likely to be overturned by other evidence).	B
The recommendation is supported by international expert opinion.	C

Grades indicate the strength of the supporting evidence, rather than the importance of the recommendations – refer to Appendix B for grading details.

GOOD PRACTICE POINT	
Where no evidence is available, best practice recommendations are made based on the experience of the Guideline Development Team, or feedback from consultation within New Zealand.	✓

This is the opinion of the Guideline Development Team, or feedback from consultation within New Zealand where no evidence is available.

2. Complementary and alternative medicines grading system

A grading system has been developed by NZGG to assess both study design and quality for complementary and alternative medicines (CAMs). This is described in Table 1. This system is compatible with the other grading systems used by NZGG and also maps to other international systems (see Table 2).

Due to the emerging nature of the evidence for CAMs, many studies are non-randomised or uncontrolled. Often no Level 1 or Level 2 evidence is available. Sometimes Level 1 or Level 2 studies cannot be carried out because it would involve a safety risk for participants. Sometimes it would be too difficult to carry out a Level 1 or Level 2 study large enough to measure rare effects. In these instances evidence is based on lower level studies.

Lower level evidence is subdivided into Level 3 and Level 4. This serves to illustrate a progression that may occur when investigating CAMs from Level 4 through Level 3 and Level 2 to Level 1 evidence.

Although possible harms and adverse events are important aspects of any CAM, they are frequently only reported from lower level studies. Higher level evidence is not often available for the reasons stated above.

A range of expert opinion also exists. In other grading systems, this is usually included in a fifth level. This level of evidence has not been reviewed for the CAM chapter in this guideline. Note also that the numbers are omitted from the level of evidence in the chapter, and that only the words are used.

TABLE 1: LEVELS OF EVIDENCE USED IN THE TBI GUIDELINE

LEVEL OF EVIDENCE	WHERE THE EVIDENCE COMES FROM
<p>1 Evidence with a high degree of reliability</p>	<p>Studies that use well tested methods to make comparisons in a fair way and where the results leave very little room for uncertainty.</p> <p>Trial design: usually Level 1 studies are systematic reviews or large, high-quality randomised controlled studies.</p>
<p>2 Evidence with reliability but open to debate</p>	<p>Studies that use well tested methods to make comparisons in a fair way but where the results leave room for uncertainty (for example, due to the size of the study, losses to follow-up or the method used for selecting groups for comparison).</p> <p>Trial design: usually Level 2 studies are systematic reviews without consistent findings, small randomised controlled trials, randomised controlled trials in which large numbers of participants are lost to follow-up, or cohort studies.</p>
<p>3 Some evidence without a high degree of reliability</p>	<p>Studies where the results are doubtful because the study design does not guarantee that fair comparisons can be made.</p> <p>Trial design: usually Level 3 studies are systematic reviews of case-control studies or individual case-control studies.</p>
<p>4 Some evidence but based on studies without comparable groups</p>	<p>Studies where there is a high probability that results are due to chance (for example because there is no comparison group or because the groups compared were different at the outset of the study).</p> <p>Trial design: usually cohort or case-control studies where the groups were not really comparable, or case-series studies.</p>

TABLE 2: SYSTEMS FOR GRADING THE QUALITY OF INDIVIDUAL STUDIES

NZGG* CAM	NZGG/ GATE†	SIGN‡	GRADE§	USPTF**	OXFORD CEBM††	NHMRC 2000**	CCS 2000 CONSENSUS§§
Level 1	Good / +	++	High	Good	Level 1 abc	Level I	Level I
Level 2	Fair / ~	+	Moderate	Fair	Level 2 abc	Level II	Level II
Levels 3 and 4	Poor / -	-	Low (very low)	Poor	Level 3 ab, and 4	Level III (1, 2, 3) and IV	Level III, IV and V

* New Zealand Guidelines Group

† Graphic Appraisal Tool for Epidemiology

‡ Scottish Intercollegiate Guidelines Network

§ Grading of Recommendations Assessment

Development and Evaluation

** US Preventable Services Task Force

†† Oxford Centre for Evidence-based Medicine

‡‡ National Health and Medical Research Council 2000

§§ Canadian Cardiovascular Society 2000 Consensus

There are methods for considering the evidence from multiple studies that address a specific question, incorporating trial designs and for weighing competing factors in forming a recommendation. For more information on study design and for guidance on balancing the benefits and harms of an intervention the reader is referred to the following websites:

- www.nzgg.org.nz
- www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/
- www.cebm.net/levels_of_evidence.asp
- www.ahrq.gov/clinic/ajpmsuppl/harris3.htm#table7
- www.gradeworkinggroup.org/links.htm
- www.sign.ac.uk/methodology/index.html
- <http://gacguidelines.ca/article.pl?sid=03/01/29/1642226&mode=thread>