

Grading recommendations in clinical practice guidelines: randomised experimental evaluation of four different systems

Carlos A Cuello García,^{1,2} Karla P Pacheco Alvarado,^{1,2} Giordano Pérez Gaxiola³

► Additional appendices are published online only. To view these files please visit the journal online (<http://adc.bmj.com>)

¹Centre for Evidence-Based Practice, Tecnológico de Monterrey, School of Medicine, Monterrey, Mexico

²Quality in Healthcare and Patient Safety Residency Programme, Tecnológico de Monterrey, School of Medicine, Monterrey, Mexico

³Department of Evidence-Based Medicine, Hospital Pediátrico de Sinaloa 'Dr. Rigoberto Aguilar Pico', Culiacán, Mexico

Correspondence to

Carlos A Cuello García, Avda. Morones Prieto 3000 pte, CITES piso 3, Col. Los Doctores, 64710, Monterrey NL, Mexico; carlos.cuello@itesm.mx

Accepted 6 February 2011
Published Online First
19 May 2011

ABSTRACT

Objective To evaluate the effect of presenting a recommendation in a clinical practice guideline using different grading systems to determine to what extent the system used changes the clinician's eventual response to a particular clinical question.

Design Randomised experimental study.

Setting Clinician offices and academic settings.

Participants Paediatricians and paediatric residents in private and public practice in Mexico.

Intervention Case notes of a child with diarrhoea and a question about clinician preference for using racecadotril. The same evidence was provided in a clinical recommendation but with different presentations according to the following grading systems: NICE (National Institute for Health and Clinical Excellence), SIGN (Scottish Intercollegiate Guideline Network), GRADE (Grading of Recommendations Assessment, Development and Evaluation) and CEBM (Centre for Evidence-Based Medicine, Oxford).

Main outcome measure Mean change in direction from baseline response (measured on a 10 cm visual scale and a Likert scale) and among groups.

Results 216 subjects agreed to participate. Most participants changed their decision after reading the clinical recommendations (mean difference 0.7 cm, 95% CI 0.29 to 1.0; $p < 0.001$). By groups, mean change (95% CI) from baseline was 0.04 (−0.68 to 0.77) for NICE, 0.31 (−0.41 to 1.05) for SIGN, 2.18 (1.48 to 2.88) for GRADE and 0.08 (−0.52 to 0.69) for CEBM ($p = 0.007$ between groups). In a final survey, a small difference was noted regarding the clarity of the results presented with the GRADE system.

Conclusion The clinician's decision to use a therapy was influenced most by the GRADE system.

Trial registration number NCT00940290.

INTRODUCTION

Clinical practice guidelines (CPG) are systematically developed statements to assist practitioner and patient decisions about appropriate healthcare in specific clinical circumstances.¹ They are intended to facilitate more consistent, effective and efficient medical practice, and improve health outcomes.²

Guideline panels develop recommendations in an effort to balance the desirable and undesirable consequences of the diagnostic or therapeutic options under consideration. These recommendations should be based on the best evidence available and, ideally, on the results of high-quality systematic reviews of rigorously randomised controlled trials.³

The idea that evidence in the medical literature should be graded was initially proposed

What is already known on this topic

- Clinical practice guidelines are clinical decision support tools to make evidence and best practice recommendations accessible quickly to the end user.
- More than 60 systems for grading evidence and strength of recommendations exist which create confusion among clinicians.

What this study adds

- Health professionals can change their decision when clear visual aids and numerical projection of the quality of evidence and strength of recommendations are portrayed.
- Among several systems GRADE provokes a major change in direction on the decision made by clinicians, possible due to these factors.

in publications from McMaster University.^{3–4} Since then, a growing number of organisations have employed various systems to grade evidence according to its quality (also referred to as levels of evidence) and the strength of recommendations.^{5–6}

A unique systematic approach to grading evidence and the strength of recommendations can minimise bias and aid interpretation⁷ for both developers and users of CPG. However, more than 60 systems have been described⁸ with wide variations in grading quality of evidence and recommendations, reflecting large differences in currently used approaches. There are plenty of arguments for and against using the same grading system across different types of recommendations.⁶

In a previous report we described the attitudes and preferences of a small group of guideline developers from around the world.⁹ The systems evaluated included those we considered extensively used, that is, the Strength of Recommendation Taxonomy (SORT) scale, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) scheme and the systems developed by the National Institute for Health and Clinical Excellence (NICE), the Canadian Task Force on Preventive Healthcare (CTFPH), the Centre for Evidence-Based Medicine in Oxford (CEBM), the US Preventive Services Task Force

(USPSTF) and the Scottish Intercollegiate Guideline Network (SIGN). Of these, NICE and GRADE were perceived as having a more rigorous development process for the grading of evidence and recommendations, with adequate description of the quality, quantity and consistency of the evidence, although the development process was considered complex and time consuming.

Evidence supporting the use of one system over another is scarce, and there are concerns regarding external and internal consistency as well as validation.⁶ We decided to carry out a randomised trial to determine if use of any of the four most common guideline grading systems (NICE, GRADE, SIGN and CEBM) changed clinician behaviour and decision making regarding a particular clinical question.

METHODS

Participants

Paediatric health professionals invited to participate in the study included paediatric residents from various postgraduate programmes in Monterrey and Culiacán, Mexico and paediatricians and paediatric subspecialists with a public or private practice within the city areas of Monterrey, Culiacán and Mexico City. Paediatricians working in public hospitals under the Health Secretariat of the State of Nuevo León and the Social Security Institute were also asked to participate. We conducted interviews in a relaxed atmosphere free of the distractions of clinical duties (ie, in halls in clinical grand rounds, or in private practice offices). If the clinician rejected the invitation it was noted on the database and the next physician on the list was asked to complete our survey.

This trial is registered at clinicaltrials.gov with the identifier NCT00940290 and was approved by the Institutional Review Board of the Tecnológico de Monterrey School of Medicine.

Interventions

The clinician was given clinical case notes which described a previously healthy child with acute watery diarrhoea (see online supplementary appendix 1). Without previous knowledge of any of the four clinical guideline recommendations on the use of racecadotril (the intervention), the clinician then answered the question “would you recommend racecadotril on this child” on the clinical case on a four point Likert scale as: (A) definitely NOT, (B) probably NOT, (C) probably YES and (D) definitely YES. The answer was indicated on a 10 cm scale using the four ordinal categories of the Likert scale.

After answering the question, the clinician was randomised to one of four groups according to four different clinical guideline grading systems: (A) NICE, (B) CEBM, (C) SIGN and (D) GRADE (see supplementary online appendices 2–5, respectively).

We had previously researched and evaluated the evidence on the topic. CC and GP searched two databases (Cochrane CENTRAL and PubMed) and also used a meta-search engine (Trip Database) but found only two randomised controlled trials^{10 11} and two systematic reviews of the use of racecadotril. Of these, only one systematic review by Emparanza Knörr *et al* was considered for inclusion in our exercise based on the quality of individual studies included and the fact that it was the most recently published study.¹² The best recommended course of action was decided by informal consensus among the three authors. The selected systematic review included two randomised controlled trials which gave a weak recommendation against using racecadotril. To confirm our

recommendation, we also evaluated NICE clinical guideline CG84¹³ on the use of racecadotril and presented a summary of the published guideline.

Participants were randomised into four blocks of 50 subjects each using a web-based tool (www.randomization.com). The randomisation sequence was concealed from the investigator (KP) administering the survey until the last moment when it was disclosed via e-mail or text message according to the number of subjects who had agreed to participate in the survey.

Outcome measures

The primary outcome was the change (before compared to after reading the guideline) in the decision made by the clinician regarding their use (or not) of racecadotril when treating a patient as described in the clinical scenario. The outcome was set a priori to objectively measure differences between the groups before compared to after reading the clinical recommendation. We measured this response as continuous data (mean±SD) in centimetres from 0 to 10 on the visual analogue scale, and as ordinal data on the Likert scale to measure proportions of respondents (see online supplementary appendix 1).

As secondary outcomes, (A) the mean final response was measured in each group and the differences between them, and (B) a seven item questionnaire was administered at the end of the exercise to measure the clinicians overall perception of the CPG recommendation. The questions specifically enquired about the clinician's opinion concerning the rigor of the process of obtaining the evidence, the clarity, quality and quantity of the evidence, similarities (consistency) between different studies included in the synthesis, the directness of the recommendation, and if the recommendation had considered the costs, risks and benefits of the therapy. This final survey was an exploratory endpoint and not a formal evaluation and did not deem the physicians experts in CPG methodology. All parameters were evaluated on a 10 cm visual analogue scale from ‘completely disagree’ to ‘completely agree’ (see online supplementary appendix 6).

To avoid bias, clinicians were unaware of the recommendations of other three CPGs, did not know which system was used for other participants and were not allowed to see other clinicians' answers.

Statistical analysis

All variables were tested for their normal distribution with the Shapiro and Kolmogorov–Smirnov tests. Continuous measures were described as means and medians; SD, IQR and 95% CI were used as measures of statistical dispersion. We did not have prior information about how a clinician would respond or change a decision; hence we considered this a pilot study and intended to calculate the power of the study after completion.

For comparison among groups we used the analysis of variance statistic for continuous normally distributed variables, or the Kruskal–Wallis test for non-Gaussian distribution. For categorical variables, the χ^2 test was used. Proportions of differences before compared to after reading the CPG were compared within groups using the McNemar test. Before compared to after continuous data comparisons in each group were analysed using the Student paired t test or the Wilcoxon test for normal or non-normal distributions, respectively. We also plotted the mean differences before compared to after the intervention with 95% CIs and although the same scale was used, we calculated standardised mean differences with

95% CI. Statistical analysis was performed using SPSS for Windows, v 13.0.

RESULTS

Of 237 paediatric health professionals asked to participate in the study, 21 refused consent. Therefore, the final sample consisted of 216 specialists, subspecialists and paediatric residents distributed among the four groups (figure 1).

Baseline characteristics were similar among the four groups (table 1). A small difference was noted in the CEBM group baseline response before reading the recommendation, where a higher proportion of physicians responded as 'definitely NOT' compared to the other three groups (table 2 and figure 2).

Overall, the group of health professionals changed their decision after reading any of the clinical recommendations from a mean of 5.6 (SD 3.3) cm (on a scale of 0 to 10) to 4.9 (SD 3.2) cm (mean difference 0.7 cm (95% CI 0.29 to 1.0); $p<0.001$).

Reference to the GRADE system resulted in the biggest change from baseline compared to the other systems, when evaluating the mean change and the change in the Likert scale categories (figure 2) and also when assessing the difference between groups at the end of the exercise (table 3).

Standardised mean differences (95% CI) in the CEBM, GRADE, NICE and SIGN groups were 0.02 (−0.24 to 0.28), 0.68 (0.41 to 0.94), 0.01 (−0.25 to 0.27) and 0.10 (−0.16 to 0.36), respectively.

Because of baseline differences in the first response of the CEBM group, we performed a multiple regression analysis

with 'response after reading the CPG' as the dependent continuous variable and adjusting for baseline response, age, sex, years of practice and systems evaluated. After adjustment, belonging to the GRADE group persisted as the variable with the strongest association ($p<0.001$). The power of the study was calculated using the paired *t* test, giving a δ of 2.18, $n=52$, σ (SD)=3.2 and $\alpha=0.05$; the study has 99% power at 5% significance for detecting the difference of 2.18 points.

In the final survey on perceptions towards the different grading systems, only a small difference was noted regarding the clarity of the results presented with the GRADE system (figure 3) when compared to the others.

DISCUSSION

Translating evidence-based recommendations into improved clinical outcomes has still to be accomplished in the healthcare arena. Clinical recommendations should be based on the best evidence available along with clinical experience and patient preferences.¹⁴ Even with access to adequate unbiased and current knowledge, external and internal barriers can affect a physician's ability to carry out recommendations, and include lack of awareness, physician attitudes, lack of agreement, the feeling of practicing 'cook-book medicine' and inertia.¹⁵

Clinical guidelines can provide a timely summary of the evidence needed at the point of care to answer questions faced on a daily basis and have an impact on patient outcomes. However, clinical guideline developers around the world are inconsistent in how they rate quality of evidence and grade the strength of recommendations,⁶ thus making it difficult for guideline users to understand the message that CPG developers are trying to communicate.

Evidence-based recommendations at the point of care can change clinician decisions and clinical outcomes as recently reported by Albano *et al.*¹⁶ Their work demonstrated that clinical guidelines can efficiently and significantly change the clinician's point of view and thus impact on clinical outcomes. Althabe *et al* demonstrated a multifaceted behavioural intervention (which included clinical audits and evidence-based tutorials) that increased the use of prophylactic oxytocin during the third stage of labour and reduced the use of episiotomies in different hospitals.¹⁷

Our study demonstrated that clinicians may change decisions on a specific topic or question after reading a recommendation from a CPG, and this change is influenced by the grading system that is applied. Of the four grading systems

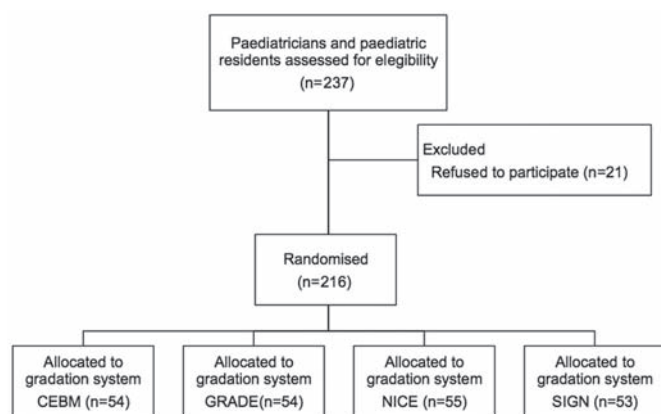


Figure 1 Participant flowchart.

Table 1 Baseline characteristics

	Study group (grading system)*			
	CEBM (n=54)	GRADE (n=54)	NICE (n=55)	SIGN (n=53)
Age (years), median (IQR)	38 (28–49)	30 (28–47)	31 (28–40)	34 (28–47)
Female	25 (46.3)	25 (46.3)	25 (45.5)	27 (50.9)
Position				
Resident	21 (38.9)	28 (51.9)	26 (47.3)	21 (39.6)
Paediatrician	12 (22.2)	16 (29.6)	13 (23.6)	20 (37.7)
Subspecialist	21 (38.9)	10 (18.5)	16 (29.1)	12 (22.6)
Years of clinical practice, median (IQR)	9 (2–20)	4 (2–19)	4 (2–12)	5 (2–18)
Private practice*, mean (±SD)	56 (39)	46 (42)	44 (39)	54 (40)

Values are number (%) unless otherwise specified.

*Percentage of time that the participant was considered to be practicing in for profit hospitals or ambulatory clinics.

CEBM, Centre for Evidence Based Medicine, Oxford; GRADE, Grading of Recommendations Assessment, Development and Evaluation; NICE, National Institute for Health and Clinical Excellence; SIGN, Scottish Intercollegiate Guideline Network.

Original article

Table 2 Results (see figure 2)

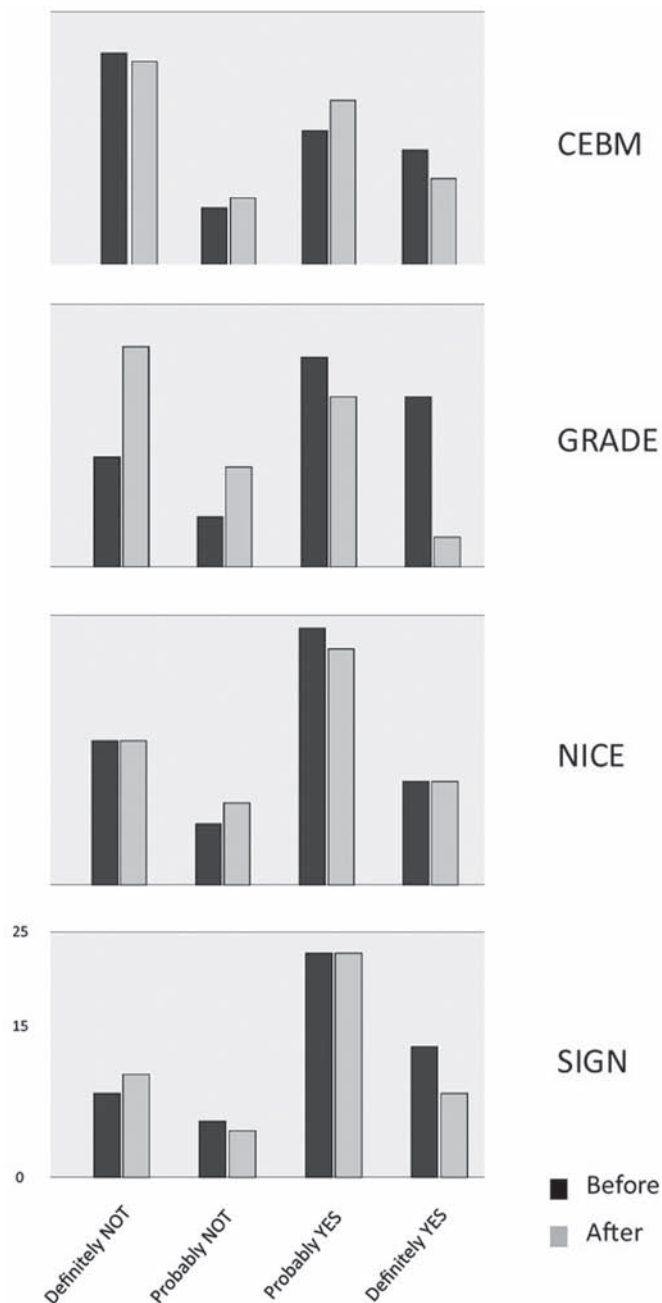
Answer	CEBM (n=54)		GRADE (n=54)		NICE (n=55)		SIGN (n=53)		p Value*	
	Grading system									
	(B)	(A)	(B)	(A)	(B)	(A)	(B)	(A)	(B)	(A)
Definitely YES	12 (22.2)	9 (16.7)	17† (31.5)	3† (5.8)	10 (18.2)	10 (18.2)	14 (26.4)	9 (18.4)	0.41	0.21
Probably YES	14 (25.9)	17 (31.5)	21 (38.9)	17 (32.7)	25 (45.5)	23 (41.8)	24 (45.3)	24 (49)	0.12	0.22
Probably NOT	6 (11.)	7 (13)	5 (9.3)	10 (19.2)	6 (10.9)	8 (14.5)	6 (11.3)	5 (10.2)	0.98	0.61
Definitely NOT	22 (40.7)	21 (39)	11† (20.4)	22† (42.3)	14 (25.5)	14 (25.5)	9 (17)	11 (22.4)	0.02	0.07

Values are number (%).

†p<0.001, McNemar test (within groups, before compared to after).

* χ^2 test, comparison among the four groups.

A, after; B, before; CEBM, Centre for Evidence Based Medicine, Oxford; GRADE, Grading of Recommendations Assessment, Development and Evaluation; NICE, National Institute for Health and Clinical Excellence; SIGN, Scottish Intercollegiate Guideline Network.

**Figure 2** Before compared to after responses among groups.

evaluated, GRADE was the most successful in provoking change in the paediatrician's decision to not prescribe racecadotril to the child in the hypothetical clinical case. In seeking an explanation for this difference, first we hypothesised that visual presentation of results, inherent to the GRADE system, influenced the clinician's point of view. The use of visual aids can increase the chances of change in the clinician's perception.^{18 19} Second, the GRADE system presents an unambiguous final recommendation, that is, either a weak or a strong recommendation is presented with no middle ground. In our case, the final recommendation was 'a weak recommendation against the use of racecadotril' or 'probably do not do it'. Third, clinicians could have perceived greater complexity and thoroughness in elaborating the recommendation as indicating a more reliable source of information. This is supported by the results of the final survey of our study.

It is important to note that only GRADE and NICE clearly distinguish recommendations from the quality of evidence; however, comparison of only these two systems shows that the biggest change results from reference to the former.

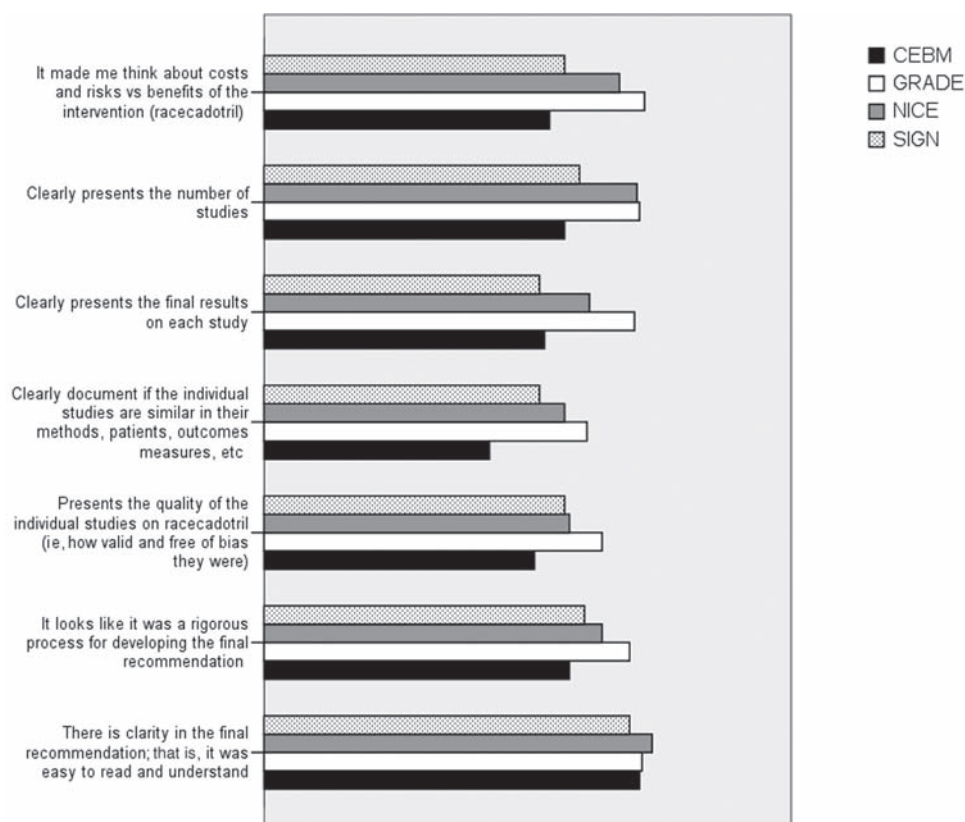
Our work has several limitations. Except for the recommendation presented by NICE, we prepared and presented the other guidelines as excerpts, and so our statements may have differed from the text other authors would have written. As stated above, the final recommendation was the same (against the use of racecadotril) among the four different grading systems according to the results and recommendations of the most recent systematic review and one existing CPG (from NICE); nevertheless, presentation bias could have taken place without our knowledge. Our population included only Mexican paediatricians, hence making external validity an issue. We did not measure prior physician knowledge or preferences for any of the grading systems. The final survey was less an analysis and more an exploratory and pragmatic evaluation of the clinicians' perceptions of differences among the systems. We did not obtain information from physicians who refused to participate as we considered them to be too few to have influenced the outcome. The experiment was based on written case notes, which did not include patient preference, past medical history or other factors that usually influence clinical decision-making, and the study was applied in a relaxed environment without the distractions of clinical duties; this situation could have influenced the physicians to change their minds. We did not include other systems (such as the US Preventive Services Task Force and American Heart Association) commonly used in institutions

Table 3 Results measured on a visual analogue scale from 0 to 10 cm

	Grading system				p Value
	CEBM (n=54)	GRADE (n=54)	NICE (n=55)	SIGN (n=53)	
Answer before reading the recommendation (baseline)	4.75 (3.77 to 5.73)	6.03 (5.20 to 7.02)	5.61 (4.75 to 6.46)	6.09 (5.34 to 7.06)	0.20
Answer after reading the recommendation	4.66 (3.73 to 5.59)	3.84 (3.0 to 4.69)	5.56 (4.70 to 6.43)	5.77 (4.85 to 6.68)	0.007
Before compared to after, mean difference (95% CI)	0.08 (−0.52 to 0.69)	2.18 (1.48 to 2.88)	0.04 (−0.68 to 0.77)	0.31 (−0.41 to 1.05)	–

Values are mean (95% CI).

CEBM, Centre for Evidence Based Medicine, Oxford; GRADE, Grading of Recommendations Assessment, Development and Evaluation; NICE, National Institute for Health and Clinical Excellence; SIGN, Scottish Intercollegiate Guideline Network.

**Figure 3** Final survey results.

and organisations around the world. We recognise that a change in the answers to one single clinical question is not sufficient evidence that clinical guidelines formulated with the GRADE system are going to change points of view in every clinical situation. More studies are needed to compare the different grading systems with other sets of questions in real life clinical practice and in other countries and medical disciplines.

Grading systems are continuously changing and improving. For example, at the time of writing experts at the Centre for Evidence Based Medicine in Oxford are updating their levels of evidence interpretation to include items such as directness precision and individual study quality.²⁰

To our knowledge this is the first study to address whether different grading systems used to give a clinical recommendation in a CPG have an influence on the clinician's point of view when facing a therapeutic decision.

The GRADE system significantly changed the decisions made by a group of paediatricians compared to three other systems. This difference was probably influenced by visual aids and an awareness of the meticulousness of the system

in formulating the final recommendation. Further research is needed to assess the influence of different systems on a large scale and on different scenarios in clinical practice. The current number of grading systems generates confusion and possibly promotes poor adherence to the recommendations. The endorsement of a single system could pave the way for the adequate implementation of CPGs at the point of care.

Competing interests None.

Ethics approval This study was conducted with the approval of the Institutional Review Board of the Tecnológico de Monterrey School of Medicine.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Beck C, Cody M, Souder E, *et al*. Dementia diagnostic guidelines: methodologies, results, and implementation costs. *J Am Geriatr Soc* 2000;**48**:1195–203.
2. Woolf SH, Grol R, Hutchinson A, *et al*. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999;**318**:527–30.
3. Cook DJ, Guyatt GH, Laupacis A, *et al*. Clinical recommendations using levels of evidence for antithrombotic agents. *Chest* 1995;**108**:227–30S.

Original article

4. **Sackett DL**. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1989;**95**:2S–4S.
5. Canadian Task Force on the Periodic Health Examination. The periodic health examination. *Can Med Assoc J* 1979;**121**:1193–254.
6. **Schünemann HJ**, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Res Policy Syst* 2006;**4**:21.
7. **Atkins D**, Best D, Briss PA, *et al*. Grading quality of evidence and strength of recommendations. *BMJ* 2004;**328**:1490.
8. **West S**, King V, Carey TS, *et al*. Systems to rate the strength of scientific evidence. *Evid Rep Technol Assess (Summ)* 2002;1–11.
9. **Cuello-García CA**, Dávalos-Rodríguez ML. Perceptions and attitudes towards different grading systems from clinical guidelines developers. *J Eval Clin Pract* 2009;**15**:1074–6.
10. **Salazar-Lindo E**, Santisteban-Ponce J, Chea-Woo E, *et al*. Racecadotril in the treatment of acute watery diarrhea in children. *N Engl J Med* 2000;**343**:463–7.
11. **Cézard JP**, Duhamel JF, Meyer M, *et al*. Efficacy and tolerability of racecadotril in acute diarrhea in children. *Gastroenterology* 2001;**120**:799–805.
12. **Empanaza Knörr JI**, Ozcoidi Erro I, Martínez Andueza MC, *et al*. [Systematic review of the efficacy of racecadotril in the treatment of acute diarrhoea]. *An Pediatr (Barc)* 2008;**69**:432–8.
13. National Institute for Health and Clinical Excellence (NICE). Diarrhoea and Vomiting Caused by Gastroenteritis in Children Under 5. NICE Guideline, 2009. <http://guidance.nice.org.uk/CG84> (accessed 26 July 2010).
14. **Sackett DL**, Rosenberg WM, Gray JA, *et al*. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;**312**:71–2.
15. **Cabana MD**, Rand CS, Powe NR, *et al*. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;**282**:1458–65.
16. **Albano F**, Lo Vecchio A, Guarino A. The applicability and efficacy of guidelines for the management of acute gastroenteritis in outpatient children: a field-randomized trial on primary care pediatricians. *J Pediatr* 2010;**156**:226–30.
17. **Althabe F**, Buekens P, Bergel E, *et al*. A behavioral intervention to improve obstetrical care. *N Engl J Med* 2008;**358**:1929–40.
18. **Akl EA**, Maroun N, Guyatt G, *et al*. Symbols were superior to numbers for presenting strength of recommendations to health care consumers: a randomized trial. *J Clin Epidemiol* 2007;**60**:1298–305.
19. **Schünemann HJ**, Best D, Vist G, *et al*. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *CMAJ* 2003;**169**:677–80.
20. **Howick J**. Levels of Evidence. Centre for Evidence-Based Medicine (CEBM). Oxford: University of Oxford, 2010. <http://www.cebm.net/index.aspx?o=5653> (accessed 26 July 2010).



Grading recommendations in clinical practice guidelines: randomised experimental evaluation of four different systems

Carlos A Cuello García, Karla P Pacheco Alvarado and Giordano Pérez Gaxiola

Arch Dis Child 2011 96: 723-728 originally published online May 19, 2011

doi: 10.1136/adc.2010.199307

Updated information and services can be found at:

<http://adc.bmj.com/content/96/8/723.full.html>

These include:

Data Supplement

"Web Only Data"

<http://adc.bmj.com/content/suppl/2011/05/23/adc.2010.199307.DC1.html>

References

This article cites 17 articles, 6 of which can be accessed free at:

<http://adc.bmj.com/content/96/8/723.full.html#ref-list-1>

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

[Child health](#) (1690 articles)

[Guidelines](#) (31 articles)

[Diarrhoea](#) (83 articles)

Notes

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>