Comment on "Perspective: NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research"

Dear Editor:

We read the Perspective article "NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research," which was published in November 2016 in this Journal (1). We agree it is important to assess the trustworthiness of evidence.

The authors describe the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach as not being applicable to nutrition and nutrition research and suggest a scoring system to overcome this. We are not convinced that this is the right approach and would encourage collaboration in a joint approach with GRADE rather than setting up something separate. GRADE is a community of scientists, physicians, and public health specialists that has been in existence for >17 y, aiming for a common approach across diverse topics. GRADE aimed to clarify the confusion that has led to the development of nearly 100 grading systems without a clear rationale for doing so (2). The GRADE approach has been endorsed and adopted by >100 international organizations and societies, which cover a wide variety of clinical, public health, and methods areas. Although the author's tool based on its name may be perceived as endorsed by the GRADE Working Group, the contrary is the case. Indeed, some aspects of the suggested tool even contradict the conceptual underpinnings of the GRADE approach (3, 4). What we aim to do with this letter is provide some background to GRADE and encourage collaboration and harmonization, which is a fundamental strength of the GRADE approach and the function of the GRADE Working Group.

GRADE is a common and transparent approach to grading certainty (or "quality") of evidence and strength of recommendations. It was developed over more than a decade by the GRADE Working Group (www.gradeworkinggroup.org), consisting of >500 members with different expertise and with involvement of numerous international organizations. GRADE constantly refines and develops its methods and extends its reach through global dialogue and careful, transparent scientific consensus development. For example, there are currently project groups working on GRADE to assess the certainty of evidence in systematic reviews on environmental toxins, qualitative research synthesis, values and preferences, and animal translation models. Each group works within defined frameworks of "project groups," with careful refinement of the methods until these are finalized, approved by the GRADE Working Group, and published. GRADE is open to newcomers and established researchers alike.

In the field of nutrition, GRADE has been applied successfully as part of Cochrane and non-Cochrane systematic reviews (5–7). For example, 118 of 470 nutrition-related *Cochrane Reviews* published in 2015 used GRADE to assess the certainty of evidence (8). Nevertheless, the authors do refer to "several limitations" that arise when applying GRADE; however, it is not clear to us what limitations the authors are actually referring to. For example, lack of blinded randomized controlled trials and the resulting sparse bodies of randomized evidence is not a methodologic shortcoming of the GRADE approach but a limitation of the evidence base. In addition, this issue is not unique to nutrition but applies to other fields such as rare diseases and surgical interventions. Furthermore, GRADE does not classify systematic reviews, but rather, the certainty of bodies of evidence obtained through systematic reviews or other appropriate forms of evidence synthesis.

In terms of the authors' suggestions about the advancements with their scoring system, we would question their appropriateness and validity. The authors are not convincing in their argument as to why randomization would not be critical to balance known and unknown prognostic factors in nutritional studies (9). There is no plausible rationale or supporting evidence to justify their approach to include funding bias as a separate item. In terms of conflict of interest, GRADE captures financial and nonfinancial interests through the existing domains for risk of bias (in particular, selective outcome reporting), indirectness, and publication bias (10). In addition, algorithmic scoring approaches for the assessment of "quality" are inferior given that they imply inevitably assigning "weights" to different items in the scale, and it is difficult to consistently justify the weights assigned (11). We encourage the authors of this article and interested readers to further explore how GRADE works and to join in advancing the methods in a unified approach.

Joerg J Meerpohl
Celeste E Naude
Paul Garner
Reem A Mustafa
Holger J Schünemann

From Cochrane Germany, Medical Center–University of Freiburg, Freiburg, Germany (JJM, e-mail: meerpohl@cochrane.de); the Center for Evidence-based Health Care, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, South Africa (CEN); the Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, United Kingdom (PG); the Division of Nephrology and Hypertension, University of Kansas Medical Center, Kansas City,

KS (RAM); and the Department of Health Research, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada (HJS).

References

1. Schwingshackl L, Knuppel S, Schwedhelm C, Hoffmann G, Missbach B, Stelmach-Mardas M, Dietrich S, Eichelmann F, Kontopanteils E, Iqbal K, et al. NutriGrade: a scoring system to assess and judge the meta-evidence of randomized controlled trials and cohort studies in nutrition research [perspective]. Adv Nutr 2016;7:994–1004.

2. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence. In: Evidence report/technology assessment. Vol. 47. Rockville (MD): Agency for Healthcare Research and Quality; 2002.

3. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383–94.

4. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401–6.

5. Kristjansson E, Francis DK, Liberato S, Benkhalti Jandu M, Welch V, Batal M, Greenhalgh T, Rader T, Noonan E, Shea B, et al. Food supplementation for improving the physical and psychosocial health of socio-economically disadvantaged children aged three months to five years. Cochrane Database Syst Rev 2015;3:CD009924.

6. Te Morenga L, Mallard S, Mann J. Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies. BMJ 2012;346:e7492.

7. Santesso N, Akl EA, Bianchi M, Mente A, Mustafa R, Heels-Ansdell D, Schunemann HJ. Effects of higher- versus lower-protein diets on health outcomes: a systematic review and meta-analysis. Eur J Clin Nutr 2012;66:780–8.

8. Naude CE, Durao S, Harper A, Volmink J. Scope and quality of Cochrane reviews of nutrition interventions: a cross-sectional study. Nutr J 2017;16:22.

9. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.

10. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. J Clin Epidemiol 2011;64:1277–82.

11. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. Biostatistics 2001;2:463–71.

Reply to JJ Meerpohl et al.

Dear Editor:

The statements "[…] that any potential bias, inconsistency, indirectness, imprecision or reliance on study type other than randomized trials can result in downgrading of the quality of evidence means that formally identifying effects which are regarded as important and based on high quality evidence using the GRADE system may be unattainable in the context of nutritional determinants of chronic disease […]" and that "This needs to be taken into account when developing nutritional recommendations" (1), which were published in a highly cited meta-analysis in the *British Medical Journal* in 2013, highlight the need for developing and evaluating new tools to assess the quality of meta-evidence in the field of nutrition research. In conceiving our article, "NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research" (2), we reasoned that the development of a scoring or grading system to assess the quality of evidence in nutrition research represents a scientific contribution that is part of a continuous process to improve global efforts in an important field of public health. With regard to developing criteria for a grading system of meta-evidence in the field of nutrition research, our concept differs from the well-established and widely used Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach in the following aspects:

- Better balanced evaluation of prospective observational and intervention study designs
- Assessment of nutrition-specific aspects, such as dietary assessment methods and their validation, calibration of FFQs, or the assessment of diet-associated biomarkers
- Consideration of the conflict of interest and funding bias as a separate item
- Introduction of a scoring system

In our article (2), we outlined the rationale for each of our decisions, including those criteria for which GRADE has a different view. The different views may be explained by the distinct perspectives of the groups. Our group is mainly composed of scientists with expertise in the field of nutrition, whereas GRADE is historically composed of mostly clinical research scientists. Other scientists from related disciplines have already found that processing evidence in the clinical research compared with the public health research areas follows slightly different approaches. For example, the US government established task forces for both evidence in clinical settings as well as evidence in public health (3). We feel that there is still a need for a scientific debate as to whether both areas could be combined in a unified system that is not dominated by one perspective.

In their letter, Meerpohl et al. "encourage[d] the authors of this article and interested readers to further explore how GRADE works and to join in advancing the methods in a unified approach." We acknowledge this kind offer from the GRADE group to open up the process of developing timely solutions, particularly in areas in which the groups have different views and also in light of the recent exponential increase in the number of meta-analyses published in the field of nutrition research. In this way, we also appreciate the actual invitation of a scientific debate and potential collaboration with GRADE as proposed by Meerpohl et al., which also includes a wider discussion of our own positions. Overall, NutriGrade should not be considered as a competitor to GRADE but rather as an approach suggested by nutrition specialists that adapts the GRADE philosophy of assessing