Akl, E. A. et al. (2017) GRADE Equity Guidelines 2: considering health equity in the GRADE guideline development process. *Journal of Clinical Epidemiology*, 90, pp. 68-75. (doi:10.1016/j.jclinepi.2017.01.017)

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/139873/

GRADE Guidelines: Health equity considerations in rating the certainty of synthesized evidence

Article Type: Original Article
Keywords: Health Equity, GRADE, guidelines, systematic review, meta-analysis, subgroup analysis

Corresponding Author: Dr. Vivian Welch, Bruyère Research Institute, University of Ottawa and Bruyère Continuing Care, Ottawa, Canada
Authors: Vivian Welch, Elie Akl, Kevin Pottie, Mohammed T. Ansari, Matthias Briel, Robin Christensen, Antonio Dans, Leonila Dans, Javier Eslava-Schmalbach, Gordon Guyatt, Monica Hultcrantz, Janet Jull, Srinivasa Vittal Katikireddi, Eddy Lang, Elizabeth Matovinovic, Joerg Meerpohl, Rachael L Morton, Annhild Mosdol, M Hassan Murad, Jennifer Petkovic, Holger Schünemann, Ravi Sharaf, Bev Shea, Jasvinder Singh, Ivan Solà, Roger Stanev, Airton Stein, Lehana Thabane, Thomy Tonia, Mario Tristan, Sigurd Vitols, Joseph Watine, Peter Tugwell

Affiliations:

Elie A. Akl (ea32@aub.edu.lb; Department of Internal Medicine, American University of Beirut, P.O. Box: 11-0236, Riad-El-Solh Beirut 1107 2020, Beirut, Lebanon)

Kevin Pottie (kpottie@uottawa.ca; Departments of Family Medicine and Epidemiology and Community Medicine, Bruyere Research Institute, University of Ottawa WHO and EU ECDC Consultant)

Mohammed T. Ansari (tosansari@gmail.com; School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine, University of Ottawa)

Matthias Briel (matthias.briel@usb.ch; Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital Basel, Switzerland, and Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada)

Robin Christensen, (Robin.Christensen@regionh.dk; Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark)

Antonio Dans (antoniodans@gmail.com, Department of Medicine, University of the Philippines, Manila, Philippines)

Leonila Dans (leonila.dans@gmail.com, Department of Pediatrics, University of the Philippines, Manila, Philippines)

Javier H. Eslava-Schmalbach, MD, MSc, PhD. (jheslavas@unal.edu.co; Group of Equity in Health, Faculty of Medicine, Universidad Nacional de Colombia and Director of the Technology Development Center, Sociedad Colombiana de Anestesiologia y Reanimacion.

Gordon Guyatt (Guyatt@mcmaster.ca; McMaster University Health Sciences Centre, Department of Clinical Epidemiology & Biostatistics and Department of Medicine, McMaster University, 1280 Main Street West, Room HSC-2C12 Hamilton, ON, Canada, L8S 4K1)

Monica Hultcrantz (monica.hultcrantz@sbu.se, 1. Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU), Stockholm, Sweden, 2.Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden)

Janet Jull (jjull013@uottawa.ca, Bruyère Research Institute & University of Ottawa 85 Primrose Avenue, Room 312, Ottawa, ON K1R 7G5)

Srinivasa Vittal Katikireddi (Vittal.Katikireddi@glasgow.ac.uk, MRC/CSO Social & Public Health Sciences Unit, University of Glasgow, Scotland, UK.)

Eddy Lang MD, (eddy.lang@ahs.ca, Department of Emergency Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta)

Elizabeth Matovinovic, (e.matovinovic@gmail.com, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand)

2

Joerg Meerpohl (meerpohl.jj@gmail.com; Centre de Recherche Épidémiologie et Statistique Sorbonne Paris Cité – U1153, Inserm / Université Paris Descartes, Cochrane France, Hôpital Hôtel-Dieu, 1 place du Parvis Notre Dame, 75181 Paris Cedex 04 France; Cochrane Germany, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Germany)

Rachael Morton (rachael.morton@sydney.edu.au; NHMRC Clinical Trials Centre, The University of Sydney, Camperdown, NSW, Australia)

Annhild Mosdøl (annhild.mosdol@fhi.no; Knowledge Centre for the Health Services, Norwegian Institute of Public Health, PO Box 4404, Nydalen, N-0403 Oslo, Norway.)

M. Hassan Murad (Murad.Mohammad@mayo.edu, Mayo Clinic Evidence-based Practice Center, Mayo Clinic 200 1st Street SW, Rochester, MN 55905)

Jennifer Petkovic (jennifer.petkovic@uottawa.ca; Bruyère Research Institute & University of Ottawa, 85 Primrose Avenue, Ottawa, ON K1R 7G5)

Holger Schünemann (holger.schunemann@mcmaster.ca, Chair of Department of Clinical Epidemiology & Biostatistics; Director, Cochrane Canada; Professor of Clinical Epidemiology and Medicine; Michael Gent Chair in Healthcare Research, McMaster University)

Ravi Sharaf (Rsharaf@northwell.edu; Northwell Health/Hofstra University School of Medicine, Long Island Jewish Medical Center, Department of Gastroenterology. Research Building B202, 270-05 76th Avenue, New Hyde Park, NY 11040.)

Beverley Shea (bevshea@uottawa.ca; Bruyère Research Institute and Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Canada)

Jasvinder Singh (Jasvinder.md@gmail.com; 1. Medicine Service and Center for Surgical Medical Acute care Research and Transitions, VA Medical Center, 510, 20th street South, FOT 805B, Birmingham, AL; 2. Department of Medicine at School of Medicine, and Division of Epidemiology at School of Public Health, University of Alabama at Birmingham, 1720 Second Ave. South, Birmingham, AL 35294-0022; 3.

3

Department of Orthopedic Surgery, Mayo Clinic College of Medicine, 200 1st St SW, Rochester, MN 55905)

Ivan Solà (isola@santpau.cat; Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau, CIBER of Epidemiology and Public Health (CIBERESP- IIB Sant Pau), Barcelona, Spain)

Roger Stanev (rstanev@uottawa.ca; Centre for Practice-Changing Research, Ottawa Hospital Research Institute. 501 Smyth Rd., L1218, Ottawa, ON, Canada, K1H6K8. Department of Philosophy, Desmarais Building, room 8101, 55 Laurier Ave. East, Ottawa, ON, Canada, K1N 6N5.)

Airton Tetelbom Stein (airton.stein@gmail.com - Full Professor of Public Health Department Ufcspa and Ulbra, Coordinator of HTA of Conceicao Hospital)

Lehana Thabane (thabanl@mcmaster.ca; Department of Clinical Epidemiology & Biostatistics, McMaser University, Hamilton, Canada)

Thomy Tonia (thomy.tonia@ersnet.org; Institute of Social and Preventive Medicine, University of Bern, Switzerland)

Mario Tristan (mtristan@ihcai.org; IHCAI/ Central American Cochrane Institute in San Jose Costa Rica, IHCAI: International Health Central American Institute)

Sigurd Vitols (sigurd.vitols@ki.se; Department of Medicine, Clinical Pharmacology Unit, Karolinska institute, L7:03, Karolinska University Hospital, SE-171 76 Stockholm, Sweden)

Joseph Watine (watine61@hotmail.com; Laboratory Medicine, Hôpital La Chartreuse, Villefranche-de-Rouergue, France)

Peter Tugwell (tugwell.bb@uottawa.ca; University of Ottawa, Department of Medicine, Faculty of Medicine, K1H 8M5, Ottawa, Canada.)

Word Count: 3,113
Tables: 3

4

Acknowledgments

Abstract

Objective: The aim of this paper is to describe a conceptual framework for how to consider health equity in the GRADE (Grading Recommendations Assessment and Development Evidence) guideline development process.

Study design and setting: Consensus-based guidance conducted with the GRADE working group and colleagues.

Results: We developed consensus-based guidance to help address health equity when rating the certainty of synthesized evidence (i.e., quality of evidence). We propose five methods of assessing health equity explicitly: 1) Include health equity as an outcome; 2) Consider patient-important outcomes relevant to health equity; 3) Assess differences in the relative effect size of the treatment; 4) Assess differences in baseline risk and hence the differing impacts on absolute effects; and 5) Assess indirectness of evidence to disadvantaged populations and/or settings.

Conclusion: The most important priority for research is to identify and document examples where health equity has been considered explicitly in guidelines. Although there is a weak evidence base for assessing health equity, this should not discourage

6

the explicit consideration of how guidelines and recommendations affect the most

vulnerable members of society.

Keywords: guidelines, health equity, systematic reviews, indirectness, GRADE guidelines,

Running title: GRADE Equity: Rating certainty of synthesized evidence

Key points
- Considering health equity in rating the certainty in synthesized evidence requires *a priori* elaboration of the disadvantaged populations and settings of interest, and methods to assess both relative and absolute effects for these populations
- GRADE judgments about directness require transparent reporting of how judgments were made

1. Background

Health inequities are differences in health that are not only unnecessary and avoidable but are also considered unfair and unjust [1] . As described in the introductory paper in this series, we use the acronym PROGRESS Plus (Place of residence, Race/ethnicity/culture/language, Occupation, Gender/sex, Religion, Education, Socioeconomic status, or Social capital + personal, relational and time-dependent characteristics) to identify individual and context-specific characteristics across which health inequities may occur [2].

Guideline panels need to decide early on whether they plan to develop equity-sensitive recommendations (as described in the introductory paper in this series). Using explicit prompts may be helpful in this process [3]. In principle, considering health equity is important for two main types of guidelines: 1) universal interventions where health inequity is a concern [4-7]; and 2) targeted or dedicated interventions aimed at one or more disadvantaged populations that have experienced health inequities. An example of the latter is the Canadian immigrant health guidelines [8], developed to raise

7

awareness of migrant health needs and improve access to effective preventive screening.

This paper provides guidance to address health equity when rating the certainty in synthesized evidence using the GRADE approach. This paper is the third paper in a four-part series on health equity and GRADE, with the introduction [Welch et al], overall process [Akl et al] and evidence to decision methods [Pottie et al].

2. Existing guidance

As discussed in the earlier two papers in this series, several authors have assessed how guidelines consider health inequity concerns[Welch et al this series, Akl et al this series. None of these papers focus on rating the certainty of synthesized evidence (i.e., quality of evidence) using the GRADE approach.

3. GRADE certainty in synthesized evidence and health equity

The GRADE approach of presenting the evidence by outcome and the associated certainty (i.e., quality of evidence) involves the production of summary tables. These tables include evidence profiles (with details on the rating of certainty for each outcome) and summary of findings (SoF) tables that are intended for the public, patients, purchasers, payers, practitioners, product-makers (e.g. manufacturers, industry) and policy-makers [9].

Five methods can be used to assess health equity with the GRADE approach:
  a) Include health equity as an outcome
  b) Consider patient-important outcomes relevant to health equity
  c) Assess differences in the magnitude of effect in relative terms between disadvantaged and more advantaged individuals or populations
  d) Assess differences in baseline risk and hence the differing impacts on absolute effects for disadvantaged individuals or populations
  e) Assess indirectness of evidence to disadvantaged populations and/or settings;

3a) Consider including health equity as an outcome for the SoF tables

If health inequity is considered an important concern by relevant stakeholders, then health equity could be included as an outcome in the PICO (Population, Intervention, Comparison, Outcome) questions, analytic framework and SoF table. In doing so, guideline developers must recognize that health equity is primarily assessed with a

8

subgroup analysis. The developers should also note that this may risk excluding other patient important outcomes, if SoF tables are limited to only seven outcomes as recommended by GRADE. For example, the NICE guideline on maternal and child nutrition identified impact on health inequalities as one of its key priorities, and framed its key question as: "What nutritional interventions are effective in improving the health of pre-conceptual, pregnant and postpartum mothers and children (up to 5 years) and reducing nutrition-related health inequalities" [10]. By including health equity as an outcome in the SoF table, it is easier for guideline panels to find the information (or lack thereof) about health equity, and consider it in their deliberations.

The direction and size of the effect on health equity is influenced by decisions such as the reference comparator group, use of relative or absolute measures, and whether the outcome is a desirable or undesirable event [11, 12]. For example, the choice of absolute or relative effects can change the conclusions about health inequalities. This is illustrated by gender disparity in stomach cancer mortality rates in the United States between 1930 and 2000 has decreased when looking at absolute differences (the rates for both men and women have declined). However, the relative risk for men compared to women has increased (increased disparity, male/female ratio)[13].

A lack of evidence about a critical health equity outcome should not be a reason to omit this from the SOF table. Indeed, this should be explicitly identified as an empty row, highlighting the need for further research to answer questions about health equity.

Example 1: The Community Guide Water fluoridation guideline [14] included "health disparities" as an outcome in the analytic framework and the SoF table because the Community Task Force placed a high value on reducing socioeconomic disparities in dental caries. Socioeconomic disparities were measured as the difference in absolute terms of a continuous outcome (caries). The evidence review found three studies that provided insufficient evidence about socioeconomic disparities to draw conclusions, highlighting a gap in the evidence-base (**Table 1**).

Example 2: "Equity impact" was the primary outcome of a systematic review on interventions to reduce smoking in adults [15]. Equity impact was assessed as the difference in the magnitude of a dichotomous outcome in absolute terms, defined as a difference in absolute effect on prevalence in lower socioeconomic status compared to higher socioeconomic status. This review showed that while increases in price or taxes reduced health inequities in smoking, mass media campaigns were more likely to worsen health inequities. This type of review

9

provides evidence that could be used to include impact on health equity as an outcome of interventions.

3b) Consider patient-important outcomes relevant to health equity

As described in the previous paper in our series[Akl et al in this series], the evidence synthesis process should consider the relative importance of different outcomes, determined with input from stakeholders representing disadvantaged groups. The evidence-base for these outcomes should then be assessed. Examples of patient importance and health equity were provided in the previous paper in this series such as the importance of inconvenience of a subcutaneous chelation pump for people with sickle cell disease [described in Akl et al in this series].

3c) Assess differences in the magnitude of effect in relative terms between disadvantaged and more advantaged individuals or populations

Average effects obscure differences between subpopulations- that is, subgroup effects may exist. Examining whether effects differ across socioeconomic status or other variables relating to health inequity requires investigating heterogeneity in the treatment effect- for example, using statistical approaches such as meta-regression or subgroup analysis. However, such results may not be available in the literature. There is evidence that systematic reviews under-report subgroup analyses from primary studies [16, 17]. Furthermore, many primary studies fail to assess possible subgroup effects related to disadvantaged populations.

Relative effects are usually similar across diverse populations and settings, and spurious subgroup effects are common [18]. Thus, if analysis suggests an apparent subgroup effect, it is important to assess the credibility of the apparent effect [19].   Sun and colleagues describe several criteria to help do this such as determining *a priori* which subgroup analysis to conduct, finding a low p-value associated with a statistical test for interaction, and providing results from within study comparisons [20, 21]. Sun et al also showed that subgroup analyses reported in the literature rarely meet these criteria. Evidence synthesis that involves subgroup analyses should therefore consider the full set of credibility issues, using an appropriate checklist, and avoid making conclusions based on chance findings (**Table 2**).

If applying the criteria in Table 2 leads to a conclusion that the subgroup effect is credible, the guideline panel should provide different estimates of relative and absolute effect for the subgroups. The panel should then consider making different recommendations for patients in these subgroups, or consider whether

10

recommendations that apply to the overall population need to be adapted to enhance equity. When the credibility of subgroup effects is low, the guideline panel may suggest that further research is needed. Few subgroup analyses meet all of these criteria; however, when most criteria are met, decision-making must consider the likely existence of subgroup effects.

Example: Hypertension and ethnicity

The Eighth Joint National Committee (JNC 8) guideline on management of hypertension recommends a calcium channel blocker or thiazide-type diuretic as initial therapy in the black hypertensive population (whereas an angiotensin-converting enzyme inhibitor, angiotensin receptor blocker, calcium channel blocker, or thiazide-type diuretic is recommended for others with hypertension) [22]. This recommendation was based on a pre-specified subgroup analysis of the ALLHAT trial (n=18,102 participants, 35% black [23]) that showed stroke was 51% (95% CI 1.22-1.86) greater for blacks treated with an ACE inhibitor first compared to those treated with a calcium channel blocker. The guideline panel rated this subgroup effect as moderate quality evidence. Had the panel not identified this subgroup effect, use of an ACE inhibitor as a first line agent would have increased health disparities between black and white ethnic groups.

3d) Assess differences in baseline risk and hence the differing impacts on absolute effects for disadvantaged individuals or populations

A higher baseline risk of adverse events in any population may lead to greater absolute harm from an intervention and conversely a higher baseline prevalence of the outcome of interest may lead to greater absolute benefit [24]. The SoF table should present the baseline risks and risk differences for each relevant population, and provide supporting evidence. Since disadvantaged populations have a disproportionate burden of almost all health conditions, it is particularly important to consider the baseline risk for these populations. Baseline risk of adverse event rates or for the outcomes of interest for specific populations are best assessed using the most robust observational data on the actual population rather than from randomized trials. GRADE guidance regarding assessing certainty of estimates of risk from broad populations is available [25, 26].

Example 1: WHO guidelines on vitamin A supplementation in children 6-59 months

In 2011 WHO recommended vitamin A supplementation for children aged 6 months to 5 years in countries where vitamin A deficiency is a public health

problem (strong recommendation) [27]. This was based on findings of a Cochrane review with a relative risk for all-cause mortality of 0.76 (95% CI: 0.69 to 0.83). The baseline risk of all-cause mortality was estimated at 0/1000 in low risk populations and 90/1000 in high risk populations (with vitamin A deficiency), based on control group event rates in the trials. Thus the absolute effects in terms of numbers of deaths prevented with vitamin A compared to the control group were 0/1000 for low risk and 22/1000 for high risk populations.

Example 2: National guide to a preventive health assessment for Aboriginal and Torres Strait Islander people

In Australia, a guideline panel sought to determine the optimal age at which to begin a series of preventive interventions in the Australian Aboriginal and Torres Strait Islander population. The panel recommended preventive interventions at an earlier age than the general population on the basis of higher prevalence of preventable diseases in Aboriginal and Torres Strait Islander populations. For example, type II diabetes is 3-4 times more common than in the general Australian population at all ages, leading to a recommendation for screening starting from age 18, instead of age 40 for the general population [28].

3e) Assess indirectness of evidence to disadvantaged populations

GRADE quality (or certainty) "*reflects our confidence that the estimates of the effect are correct. In the context of recommendations, quality reflects our confidence that the effect estimates are adequate to support a particular recommendation. "Quality" as used in GRADE means more than risk of bias and so may also be compromised by imprecision, inconsistency, indirectness of study results, and publication bias.*" [29]. Qualitative evidence may also be important when considering health equity. Certainty for qualitative evidence synthesis can be rated using the CerQUAL tool [30] in which the domain "relevance" is most closely aligned with directness.

Indirectness refers to the comparability between the population, the intervention, or the outcomes measured in research studies and those under consideration in a guideline or systematic review [31]. The GRADE approach evaluates the lack of directness as "Indirectness". Direct evidence may be lacking because some populations may not represent a large proportion of trial populations (e.g. migrants and refugees) and data are unlikely to be disaggregated for specific subgroups. Direct evidence may also be lacking because some populations are explicitly excluded from trials, such as pregnant women and people with multiple morbidities [32-35]. Since multiple morbidities are more

common in socioeconomically disadvantaged people [36], this may result in disproportionate exclusion of disadvantaged populations from trials. When direct evidence for the relevant disadvantaged population is not available, guideline developers will have to evaluate the indirectness of evidence obtained from other populations [37].

As a rule, certainty of the evidence should not be rated down for indirectness for population differences unless there are compelling reasons to anticipate differences in effect due to biology/physiology, sociocultural influences or setting-specific resource issues that impact the effectiveness or harms of the intervention. In other words, one anticipates a different subgroup effect in either relative or absolute impact of treatment, though evidence is not available to make a formal assessment. (If it were, it should be formally assessed, as in 3c and 3d). Guideline panels need to consider that rating down for indirectness could in itself increase inequities if this leads to less use of an effective intervention by disadvantaged groups. In other words, lower certainty in effect estimates may lead to a weak recommendation and therefore under-use of a beneficial treatment. Rating down for indirectness should therefore be done cautiously since effective interventions are needed even more in some populations that are often excluded from trials, such as those with multiple morbidities.

Example 1: Canadian migrant guidelines not rated down for indirectness

The quality of the evidence was not rated down for indirectness in the Canadian migrant guideline addressing screening for latent TB; the panel considered the evidence not to be indirect for migrants. Although no migrants were included in studies of intervention effectiveness, the developers did not expect different relative effects [38].

Example 2: CDC Guidelines for brief alcohol counselling for people with HCV infection rated down for indirectness

The Centers for Disease Control and Prevention recommended brief alcohol screening and counselling for all person with HCV infection, based on a systematic review of 22 randomized trials which found a reduction of alcohol consumption of 38.42% (95% CI = 30.91–65.44) more than the control groups after 1 year. This evidence was rated down for indirectness by the guideline panel because none of the trials included persons with HCV infection [39].

13

4. Methodologic challenges

In developing this guidance, we identified a number of methodologic challenges. Firstly, assessing effects on health equity is not a linear process. There may be a need to revisit the focus of the guideline during the evidence review process, including the consideration of important disadvantaged groups. NICE does this explicitly by revisiting their key questions regarding health equity throughout the process.

Secondly, there are often limitations in the underlying evidence-base including poor reporting of sociodemographic characteristics [40, 41], under-reporting of subgroup analyses that are not statistically significant [21, 41], and use of multivariable models that may be over-adjusted for effect mediators, and / or include unnecessary collinear variables [42]. Lack of evidence on whether the effects are consistent or different for disadvantaged populations makes it difficult to judge indirectness and rate certainty of evidence. When the evidence-base is insufficient to assess effects on health equity, guideline panels need to make these limitations explicit and transparently report how they made judgments.

Thirdly, epidemiologic evidence addressing baseline risk for specific disadvantaged groups may be difficult to obtain for the population or geographic region for which the recommendations are being developed. Health systems at local, regional and national levels do not have consistent or reliable methods for reporting health status across all sociodemographic indicators of interest. Guideline panels should transparently report how they determined baseline risk estimates.

Fourthly, assessing directness of evidence depends on the clinical and methodological expertise and judgment of SoF developers. The GRADE Guideline Development Tool (GDT) includes an explicit checklist when producing SoFs to ask whether the evidence is direct across population, intervention, comparison and outcome, and document the decision for rating down, if performed.

5. Research agenda

The most important research priority is to systematically identify further examples of how guideline panels have assessed health equity considerations and incorporated these assessments into recommendations using transparent methods. For example, all WHO guidelines make their evidence to recommendation tables and SoFs publicly available for research such as this. These assessments could provide examples of whether, and how, the five issues (a-e) above have been considered for different

situations, such as assessing the credibility of subgroup analyses and judging indirectness for disadvantaged populations.

In conclusion, the GRADE process provides a structured approach to assess effects on health equity. Health equity considerations warrant increased use of these methods in systematic reviews and guidelines. The findings of assessing health equity using these five steps in guideline development provides a basis for judging "impact on equity" which is part of the DECIDE framework, and details about this process are covered in the fourth paper of this series [Pottie et al].

Table 1. Effect of Community Water Fluoridation on socioeconomic health inequities in caries [14]

Outcome	Measure	Evidence
Health equity as measured by socioeconomic disparities in caries	% caries reduction	Inconsistent results on socioeconomic disparities (3 studies)
	dmft/DMFT	No data on socioeconomic disparities

DMFT/dmft – Decayed, Missing or Filled Teeth
Upper case refers to permanent teeth; lower case to primary teeth

Table 2: Checklist for assessing credibility of subgroup analyses [43]

Design

    Is the subgroup variable a characteristic measured at baseline or after randomization?

    Is the effect suggested by comparisons within rather than between studies?

    Was the hypothesis specified a priori?

    Was the direction of the subgroup effect specified a priori

    Was the subgroup effect one of a small number of hypothesized effects tested?

Analysis

    Does the interaction test suggest a low likelihood that chance explains the apparent subgroup effect?

    Is the significant subgroup effect independent?

Context

    Is the size of the subgroup effect large?

    Is the interaction consistent across studies?

    Is the interaction consistent across closely related outcomes within the study?

    Is there indirect evidence that supports the hypothesized interaction (biological rationale)?

6. References

1. Whitehead, M., *The concepts and principles of equity and health.* Int J Health Serv, 1992. **22**(3): p. 429-45.

2. O'Neill, J., et al., *Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health.* J Clin Epidemiol, 2014. **67**(1): p. 56-64.

3. Welch, V., et al, *GRADE Equity: Introduction: Why health equity and GRADE guidelines?* Journal of Clinical Epidemiology, 2016.

4. Excellence, N.I.f.H.a.C., *Positively equal: a guide to addressing equality issues in developing NICE clinical guidelines.* , 2012, National Institute for Health and Clinical Excellence: London, UK. http://www.nice.org.uk.

5. Mitchell, J.T., M.; Degnan, B.; Gabert, T.; Haight, B.; Kessler, D.; Mack, N.; Mallen, E.; Novak, H.; Rossmiller, D.; and L.S. Setterlund, K.; Valentino, N.; Vincent, S. , *Depression, Adult in Primary Care*, 2013, Institute for Clinical Systems Improvement (ICSI), available: https://www.icsi.org/_asset/fnhdm3/Depr.pdf.

6. Eslava-Schmalbach, J.e.a., *Colombia guidelines on complications of pregnancy and childbirth.* 2013.

7. *Suspected Cancer in Primary Care: Guidelines for investigation, referral and reducing ethnic disparities*, 2009, New Zealand Guidelines Grup. Available: http://www.health.govt.nz/system/files/documents/publications/suspected-cancer-guideline-sep09.pdf: Wellington, NZ.

8. Pottie, K., et al., *Evidence-based clinical guidelines for immigrants and refugees.* CMAJ, 2011. **183**(12): p. E824-925.

9. Concannon, T.W., et al., *A new taxonomy for stakeholder engagement in patient-centered outcomes research.* Journal of General Internal Medicine, 2012. **27**(8): p. 985-91.

10. *Maternal and child nutrition. NICE public health guidance 11*, 2014, NICE. Available at: http://www.nice.org.uk/guidance/ph11/resources/guidance-maternal-and-child-nutrition-pdf: London, UK.

18

11. Harper, S., N.B. King, and M.E. Young, *Impact of selective evidence presentation on judgments of health inequality trends: an experimental study.* PLoS One, 2013. **8**(5): p. e63362.

12. Keppel, K., et al., *Methodological issues in measuring health disparities.* Vital Health Stat 2, 2005(141): p. 1-16.

13. Harper, S. and J. Lynch, *Methods for Measuring Cancer Disparities: Using Data Relevant to Healthy People 2010 Cancer-Related Objectives*, in *Cancer Surveillance Monograph Series, Number 6*2005, National Cancer Instititute. Cancer Surveillance Monograph Series, Number 6: Bethesda, MD.

14. *Preventing Dental Caries: Community Water Fluoridation.* The Community Guide. Community Preventive Services Task Force, 2013.

15. Brown, T., S. Platt, and A. Amos, *Equity impact of population-level interventions and policies to reduce smoking in adults: a systematic review.* Drug Alcohol Depend, 2014. **138**: p. 7-16.

16. Tugwell, P., et al., *Is health equity considered in systematic reviews of the Cochrane Musculoskeletal Group?* Arthritis Rheum, 2008. **59**(11): p. 1603-10.

17. Humphreys, D.K. and D. Ogilvie, *Synthesising evidence for equity impacts of population-based physical activity interventions: a pilot study.* Int J Behav Nutr Phys Act, 2013. **10**: p. 76.

18. Schmid, C.H., et al., *An empirical study of the effect of the control rate as a predictor of treatment efficacy in meta-analysis of clinical trials.* Stat Med, 1998. **17**(17): p. 1923-42.

19. Burke, J.F., et al., *Three simple rules to ensure reasonably credible subgroup analyses.* BMJ, 2015. **351**: p. h5651.

20. Sun, X., et al., *Credibility of claims of subgroup effects in randomised controlled trials: systematic review.* BMJ, 2012. **344**: p. e1553.

21. Sun, X., et al., *The influence of study characteristics on reporting of subgroup analyses in randomised controlled trials: systematic review.* BMJ, 2011. **342**: p. d1569.

22. James, P.A., et al., *2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8).* JAMA, 2014. **311**(5): p. 507-20.

19

23. Leenen, F.H., et al., *Clinical events in high-risk hypertensive patients randomly assigned to calcium channel blocker versus angiotensin-converting enzyme inhibitor in the antihypertensive and lipid-lowering treatment to prevent heart attack trial.* Hypertension, 2006. **48**(3): p. 374-84.

24. Manuel, D.G., et al., *Modelling preventive effectiveness to estimate the equity tipping point: at what coverage can individual preventive interventions reduce socioeconomic disparities in diabetes risk?* Chronic Dis Inj Can, 2014. **34**(2-3): p. 94-102.

25. Spencer, F.A., et al., *Uncertainties in baseline risk estimates and confidence in treatment effects.* BMJ, 2012. **345**: p. e7401.

26. Iorio, A., et al., *Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients.* BMJ, 2015. **350**: p. h870.

27. WHO, *Guideline: Vitamin A supplementation in infants and children 6-59 months of age*, 2011, World Health Organization: Geneva, Switzerland, Available at: http://apps.who.int/iris/bitstream/10665/44664/1/9789241501767_eng.pdf?ua=1&ua=1.

28. NACCHO/RACGP, *National guide to a preventive health assessment for Aboriginal and Torres Strait Islander people, 2nd ed.*, 2012, National Aboriginal Community Controlled Health Organisation (NACCHO) and The Royal Australian College of General Practitioners (RACGP) South Melbourne.

29. Balshem, H., et al., *GRADE guidelines: 3. Rating the quality of evidence.* J Clin Epidemiol, 2011. **64**(4): p. 401-6.

30. Lewin, S., et al., *Using Qualitative Evidence in Decision Making for Health and Social Interventions: An Approach to Assess Confidence in Findings from Qualitative Evidence Syntheses (GRADE-CERQual).* PLoS Med, 2015. **12**(10): p. e1001895.

31. Guyatt, G.H., et al., *GRADE guidelines: 8. Rating the quality of evidence--indirectness.* J Clin Epidemiol, 2011. **64**(12): p. 1303-10.

32. Bartlett, C., et al, *The causes and effects of socio-demographic exclusions from clinical trials.* Health Technol Assess, 2005. **9**(38): p. iii-iv, ix-x, 1-152.

20

33. Braun, J., et al, *Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial.* Lancet, 2002. **359**(9313): p. 1187-93.

34. O'Dell, J.R., et al, *Therapies for active rheumatoid arthritis after methotrexate failure.* N Engl J Med, 2013. **369**(4): p. 307-18.

35. Charytan, D. and R.E. Kuntz, *The exclusion of patients with chronic kidney disease from clinical trials in coronary artery disease.* Kidney Int, 2006. **70**(11): p. 2021-30.

36. Barnett, K., et al., *Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study.* Lancet, 2012. **380**(9836): p. 37-43.

37. Welch, V., et al., *Systematic reviews need to consider applicability to disadvantaged populations: inter-rater agreement for a health equity plausibility algorithm.* BMC Med Res Methodol, 2012. **12**: p. 187.

38. Greenaway, C., et al., *Tuberculosis: evidence review for newly arriving immigrants and refugees.* CMAJ, 2011. **183**(12): p. E939-51.

39. Smith, B.D., et al *Recommendations for the identification of chronic hepatitis C virus infection among persons born during 1945–1965. .* MMWR, 2012(61): p. RR-4.

40. Furler, J., et al., *Participant demographics reported in "Table 1" of randomised controlled trials: a case of "inverse evidence"?* Int J Equity Health, 2012. **11**: p. 14.

41. Welch, V., et al., *How effects on health equity are assessed in systematic reviews of interventions.* Cochrane Database Syst Rev, 2010(12): p. MR000028.

42. Morton, R.L., et al., *The impact of social disadvantage in moderate-to-severe chronic kidney disease: an equity-focused systematic review.* Nephrol Dial Transplant, 2015.

43. Sun, X., et al., *Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses.* BMJ, 2010. **340**: p. c117.