

2019 sickle cell disease guidelines by the American Society of Hematology: methodology, challenges, and innovations

M. Hassan Murad,¹ Robert I. Liem,² Eddy S. Lang,³ Elie A. Akl,⁴ Joerg J. Meerpohl,⁵ Michael R. DeBaun,⁶ John F. Tisdale,⁷ Amanda M. Brandow,⁸ Sophie M. Lanzkron,⁹ Stella T. Chou,¹⁰ Starr Webb,¹¹ and Reem A. Mustafa¹²

¹Evidence-Based Practice Center, Mayo Clinic, Rochester, MN; ²Division of Hematology, Oncology & Stem Cell Transplant, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL; ³Department of Emergency Medicine, University of Calgary Alberta Health Services, Calgary, Canada; ⁴Department of Internal Medicine, American University of Beirut, Beirut, Lebanon; ⁵Institute for Evidence in Medicine, Medical Center, University of Freiburg, Freiburg, Germany; ⁶Department of Pediatrics, Vanderbilt-Meharry Center of Excellence in Sickle Cell Disease, Vanderbilt University Medical Center, Nashville, TN; ⁷National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; ⁸Section of Pediatric Hematology/Oncology, Medical College of Wisconsin, Milwaukee, WI; ⁹Division of Hematology, Johns Hopkins University School of Medicine, Baltimore, MD; ¹⁰Children's Hospital of Philadelphia, Philadelphia, PA; ¹¹American Society of Hematology, Washington, DC; and ¹²Department of Medicine, University of Kansas Medical Center, Kansas City, KS

The American Society of Hematology (ASH) convened 5 guideline panels to develop clinical practice recommendations addressing 5 management areas of highest importance to individuals living with sickle cell disease: pain, cerebrovascular complications, pulmonary and kidney complications, transfusion, and hematopoietic stem cell transplant. Panels were multidisciplinary and consisted of patient representatives, content experts, and methodologists. The Mayo Clinic Evidence-Based Practice Center conducted systematic reviews based on a priori selected questions. In this exposition, we describe the process used by ASH, including the GRADE approach (Grades of Recommendations, Assessment, Development and Evaluation) for rating certainty of the evidence and the GRADE Evidence to Decision Framework. We also describe several unique challenges faced by the guideline panels and the specific innovations and solutions used to address them, including a curriculum to train patients to engage in guideline development, dealing with the opioid crisis, and working with indirect and noncomparative evidence.

Background

Sickle cell disease (SCD) affects millions of people worldwide, with a great effect on their morbidity and mortality. There are approximately 100 000 individuals with SCD in the United States.¹ Clinical manifestations vary across the lifespan of affected individuals, and include debilitating pain, increased risk for infection, acute organ injury, end-organ damage, and early death. Despite therapeutic advances focused on disease modification and curative intent, the care of individuals with SCD may be fragmented and negatively affected by poor access to providers with the necessary expertise in SCD.

Developing guidelines on rare diseases such as SCD is hindered by the lack of availability of evidence warranting high certainty that is required for providing informative recommendations.² These challenges, along with several others, are described in this exposition, with solutions, methodologies, and innovations used in the 2019 American Society of Hematology (ASH) SCD guidelines.

Methods

Panel composition

ASH recruited 62 experts and 10 patient representatives to serve on 5 guideline panels. The 5 guidelines addressed 5 management areas of highest importance to individuals living with SCD: pain,

cerebrovascular complications, cardiopulmonary and renal complications, transfusion, and hematopoietic stem cell transplant. These topics were chosen because they were considered to have the greatest effect on the day-to-day care of individuals living with SCD. ASH funded the Mayo Evidence-Based Practice Research Center to conduct systematic reviews and provide methodological support to this project.

There were 6 clinical coauthors (1 panel had 2) and 5 methods coauthors. One methods coauthor was from the Mayo Clinic methodology group; the other 4 were recruited by ASH on the basis of their expertise and research interests related to the topic of the guideline. All panelists were chosen for their clinical and research expertise in the area of focus pertinent to each panel.

The panels were multidisciplinary and consisted of 12 to 16 individuals. In addition to patient representatives, hematologists, and SCD specialists, other specialties were included, such as pain specialists, emergency medicine specialists, and psychologists (pain panel); neurologists, neuroradiologists, and 2 researchers with PhDs in cognitive science education and psychology (cerebrovascular panel); cardiologists, pulmonologists, and nephrologists (cardiopulmonary and renal panel); pathologists/transfusion specialists (transfusion panel); and transplant specialists (transplant panel). The panelists were recruited in 2016 and met virtually and face to face multiple times during 2017 to 2019.

ASH vetted each panelist and systematic review team member for conflicts of interest according to ASH policies. A majority of the panel members did not have direct financial conflicts with for-profit companies that could be directly affected by the guidelines. The systematic review team did not have any direct financial conflicts of interest.

Developing guideline questions

Each panel met face to face to develop a list of 10 questions. The questions were defined using the PICO framework (patient, intervention, comparison, and outcomes). The questions were chosen on the basis of dilemmas faced in everyday clinical practice, regardless of the preconceived knowledge of the evidence base. Therefore, some questions were chosen, realizing that the supporting evidence would likely be sparse or of very low certainty. The chosen outcomes were patient-important outcomes,³ defined as outcomes that patients would feel, recognize, and view as important. All panel members rated the importance of outcomes on a scale of 1 to 9 (1-3 of limited importance, 4-6 important, and 7-9 critical).

Guideline framework

The development of the ASH guidelines on SCD followed the GRADE approach. In this approach, a first and critical step is to determine the certainty of the evidence for health effects. In a second step, the guideline panel moves from the evidence to making recommendations. Certainty of the evidence (also called quality of evidence) is a construct that represents the trustworthiness of the whole body of evidence that supports a particular outcome and represents our certainty that the treatment effect is within a certain range.⁴ The certainty of evidence is determined on the basis of study design (with bodies of evidence consisting of randomized trials starting at a high certainty rating and bodies of evidence consisting of nonrandomized studies starting at

a low rating). This initial rating is then modified⁵ on the basis of other certainty domains such as risk for bias, imprecision, indirectness, inconsistency, and the likelihood of publication bias. Other factors such as having a large effect size can increase certainty in evidence derived from nonrandomized studies.

Moving from evidence to making recommendations follows a framework called Evidence to Decision framework.⁶ In this framework, other factors and considerations are incorporated with the certainty in evidence to make a recommendation. These factors include the balance of benefits and harms, patient's values, feasibility of the recommended action, acceptability of the recommended action, effect of resources, and effect on health equity.⁶ All these factors will determine the direction of the recommendation (ie, favoring the intervention over the comparator or vice versa) and its strength (ie, strong or conditional).

Interpretation of strong and conditional recommendations

The strength of a recommendation is expressed as either strong ("the guideline panel recommends...") or conditional ("the guideline panel suggests..."). The strength of a recommendation has important implications to patients, clinicians, policymakers, and researchers (Table 1).

Overview of the literature search and study selection and abstraction

One of the main principles of evidence-based medicine and a criterion for a trustworthy guideline is that evidence is collected systematically after explicit inclusion and exclusion criteria are determined a priori.⁵ For all 5 SCD guidelines, systematic searches were designed by the methodology team at the Mayo Clinic Evidence-based Practice Center, which included 1 of the methodological coauthors, medical reference librarians, and other investigators with expertise in evidence synthesis. Searches for primary studies included at a minimum Medline and Embase. Controlled vocabulary supplemented with keywords was used to search for SCD and the other concepts. The searches leveraged existing systematic reviews when available, including those done to support the National Heart, Lung, and Blood Institute guidelines⁸ by updating previous search strategies if they were judged to be sufficiently comprehensive. Guideline panelists with SCD-specific expertise helped by identifying additional studies missed by electronic searches, monitoring the literature for key studies published after the search, and informing the methodology team about recently presented abstracts and anticipated upcoming studies. Abstracts not published as full manuscripts were considered when deemed to provide critical information, realizing that their data are not peer reviewed and appraisal of their risk for bias is limited.

Study selection and data extraction were performed by independent pairs of investigators and disagreements were resolved by consensus. Risk for bias assessment tools were chosen on the basis of the type of the question and study design (eg, the Cochrane Collaboration Tool for Randomized Controlled Trials and the modified Newcastle Ottawa Scale for nonrandomized studies).

Table 1. Implications of the strength of recommendation

Stakeholder	Strong recommendation	Conditional recommendation
Patient	Most individuals in this situation would want the recommended course of action; only a small proportion would not	The majority of individuals in this situation would want the suggested course of action, but many would not
Clinician	Most individuals should follow the recommended course of action; decision aids are unlikely to be needed	Different choices are appropriate and clinicians must help each patient arrive at a management decision consistent with his or her values and preferences. Decision aids are likely useful
Policymaker	The recommendation can be adopted as policy in most situations (may be a quality criterion or performance indicator)	Debate and involvement of various stakeholders is needed to develop policy (recommendations are not candidate to be a quality measure)
Researcher	Most of the time, the supporting evidence is of high certainty, and additional research is not needed.* Rarely, the evidence supporting a strong recommendation would warrant lower certainty, in which case, future research to produce more trustworthy estimates would be needed	This recommendation is likely to be strengthened (for future updates or adaptation) by additional research

Adapted from Schünemann et al.⁷

*Additional research may not be needed to support this specific effect estimate. However, additional research can be needed to extend findings to other populations or settings or to support implementation.

Challenges and innovations

Patient engagement

Although patient engagement in guideline development is supported by a strong moral and ethical rationale, the quality and depth of this engagement varies, and guideline developers often struggle with achieving adequate engagement.⁹ Tokenistic engagement, that is, offering engagement as a gesture, is not unusual.¹⁰ For the SCD guidelines, 2 patient representatives were recruited for each of the 5 guideline panels. To empower patient representatives and facilitate meaningful engagement, we developed a curriculum based on content from the GRADE Working Group. The curriculum was delivered through a webinar, followed by a face-to-face half-day workshop.¹¹ The curriculum targeted patients' knowledge, skills, and attitudes and was followed by a postcurriculum survey that demonstrated increased knowledge about guideline development, improved self-efficacy, and confidence in their ability to participate in the process. Patients developed a script to use during panel deliberations (eg, what to do when the conversation includes too much jargon). They also developed an instruction sheet for panelists on how to empower and engage patients.¹¹

The opioid crisis

The pain panel was challenged to develop guidelines for acute and chronic pain management in a unique environment and changing national policies. On average, 130 Americans die every day from an opioid overdose, and in 2017, the number of overdose deaths involving opioids was 6 times higher than in 1999.¹² Many of these deaths are attributable to illicitly obtained high-potency synthetic opioids. Yet many are a result of prescription opioids. The opioid crisis has become a major public health emergency that prompted action by various governmental agencies on a federal and state level. Prescribing guidelines published by the Centers for Disease Control and Prevention in 2016 restricted the use of opioids for patients with chronic noncancer pain. A subsequent letter from the Centers for Disease Control and Prevention to ASH clarified that the guideline was not intended to deny clinically appropriate opioid therapy to individuals with SCD¹³, which was followed by changes of prescribing requirements in some states, adding SCD to the exception group.

The pain panel considered the opioid crisis when they chose the guideline questions and when they applied the Evidence to Decision framework to make decisions. The panel developed 10 critical questions that centered on commonly encountered pain management issues in both the outpatient and inpatient settings. The existence of relevant research evidence was not the driver for choosing these questions. Longstanding practices in pain management related to transfusion therapy, basal opioid infusions, and administration of parenteral hydration during acute pain episodes were subjected to scrutiny.

The Evidence to Decision framework was contextualized within a complex landscape of issues ranging from stigma, discrimination, risks for opioid use disorders, and the current deadly opioid crisis. In terms of values, the panel placed a high value on functional status as the major factor driving the care strategy when individualizing pain management plans. In terms of the balance of benefits and harms for chronic opioid therapy for the treatment of chronic pain, the panel differentiated 3 situations. Harms may exceed benefits for initiation of chronic opioid therapy as first-line treatment for newly diagnosed chronic pain. Benefits of continuing chronic opioid therapy in a high-functioning patient suffering from chronic pain may exceed harms. Finally, harms of continuing a poorly functioning patient with chronic pain on chronic opioid therapy may exceed benefit. In terms of feasibility, the panel considered the effect of the recommendations on access to pain care for individuals with SCD who suffer from severe and disabling pain.

Indirectness of evidence

Indirect evidence is often needed to develop recommendations. However, this challenge was most prominent when developing recommendations for pain management in SCD. Despite the high prevalence of both acute and chronic pain in patients with SCD, there was a near-complete absence of robust research in this area. Although some studies were available on acute pain in SCD, hardly any were available on chronic pain. Therefore, evidence synthesis was conducted in 2 phases. The first phase focused on direct evidence (ie, studies of pain management in individuals with SCD). The second phase focused on indirect evidence (ie, studies of pain management in individuals without SCD). The panel conducted an iterative process to obtain panel consensus on pain conditions other than SCD from which to derive the indirect evidence. Through

this process, the panel concluded that evidence for pain management in patients with fibromyalgia among other chronic noncancer pain conditions would be leveraged for select questions. The multidisciplinary panel made the judgment that such evidence would inform the recommendations and acknowledged that certainty in the evidence would be lowered because of indirectness.¹⁴ The extent of this indirectness varied. It is plausible that some of this evidence is sufficiently direct, whereas some warrants a rating of very serious indirectness.

Lack of comparative evidence

The 2014 National Heart, Lung, and Blood Institute Expert Panel Report on SCD did not address stem cell transplantation¹⁵; however, evidence has been accumulating and the question of who should receive transplant, and when, has become relevant to the daily practice of SCD care. The guideline panel dealing with hematopoietic stem cell transplant faced the challenge of the absence of comparative studies. To ascertain net benefit supporting the transplant panel recommendations, outcomes of patients who received hematopoietic stem cell transplant in a group of studies were compared indirectly with outcomes of patients managed with transfusion and hydroxyurea in other studies. To be informative and comparable, this indirect comparison required standardization of event rates (ie, risk for stroke) and presenting them as events per 100 patient years. To alleviate the concern that data from a single study with patients receiving standard treatment might not be representative, different data sources were considered to establish a comparator group risk: data from a recent systematic review about outcomes in patients treated with hydroxyurea therapy, data from a recent systematic review about outcomes in patients receiving regular transfusion therapy, data from recent large studies of hydroxyurea, and studies of transfusion therapy considered by the panel to be most representative of current best practice. These different comparator group risks were considered in sensitivity analyses when discussing recommendations.

However, transplant has been historically only offered to the most severely affected patients, rendering these comparisons unbalanced. Given the nonadjusted nature of the comparison, the likelihood of ecological bias (when inferences about individuals are deduced from inferences about the groups to which they belong), the statistical assumptions needed for analysis, and the assumptions about comparability of prognosis of the compared group, an additional source of data was needed. The panel was able to extract registry data from the Center for International Blood and Marrow Transplant Research, where all transplant data are housed. These data answered several of the guideline questions more directly and provided longer follow-up than reported studies. Efforts were taken to identify overlap between registry data and the published literature.

Making decisions about screening

The cardiopulmonary and renal panel faced the particular challenge of making recommendations about screening in asymptomatic individuals. The GRADE approach has laid out guidance about judging the certainty of evidence and making recommendations about health care-related tests and diagnostic strategies.¹⁶⁻¹⁹ However, this approach addressed decision making when only diagnostic studies were available (ie, data were available on test accuracy results and not on health outcomes) and not specifically for screening.^{19,20}

To allow for consistent decisions among the different screening questions, we developed a framework that supported developing recommendations based on multiple domains that can justify screening. This framework was informed by criteria proposed in 1968 to the World Health Organization by Wilson and Jungner, the US Preventive Services Taskforce manual, the GRADE diagnosis Evidence to Decision framework, and other literature about screening tests.^{19,21-23} Figure 1 depicts the framework explaining when screening is justified.

Applying the screening framework

The cardiopulmonary and renal panel was unable to identify direct supporting evidence for screening individuals with asymptomatic SCD with echocardiography, pulmonary function testing, or polysomnography. Test accuracy data were either lacking or, when available, had serious limitations hindering decision making. There were also no direct, head-to-head comparisons of benefits and harms among individuals who underwent versus those who did not undergo screening. As a consequence, literature reviews targeted other data that could potentially support developing recommendations. These additional data were mostly in the form of observational studies that reported some patient important outcomes among patients who received the screening tests and had specific results.

Using this framework is critical for transparency and to show guideline users how decisions were made (the effect of values on the decision, and which criteria drove the decision). The cardiopulmonary and renal panel placed high value on meeting 2 criteria for screening tests to be considered: the panel must have high certainty that individuals with a positive screening test would get a different management than those with a negative test, and the panel must have high certainty that there is an effective treatment/management for the condition that improves outcomes if administered earlier than when the condition is clinically apparent. Placing higher value on these 2 criteria led to conditional recommendations against screening with echocardiography, pulmonary function testing, or polysomnography in asymptomatic patients.

The cerebrovascular panel addressed the issues of screening for abnormal transcranial Doppler measurements, which are a risk factor in children with Hb SS or SB⁰ thalassemia for strokes within 12 months, silent cerebral infarcts, and cognitive and developmental delays. These sequelae met the criteria of high prevalence (at least 35% for silent cerebral infarcts), grave consequences (high rate of recurrent cerebral infarcts after an initial silent cerebral infarct, and cognitive morbidity such as the loss of IQ points), and available interventions (eg, transfusion for silent cerebral infarcts in children, school-based interventions for students, and federal legislation to support adults with cognitive impairment [disabilities]). The panel recommended screening for cognitive impairment and developmental delays with standardized questionnaires and at least 1 nonsedated magnetic resonance imaging scan of the brain to detect silent cerebral infarct during childhood, and again during adulthood.

It is plausible that some health care providers or patients may disagree with the screening recommendations made by these 2 panels and emphasize different criteria in the framework. Yet, using an explicit framework that demonstrates how the decisions were made and which values or criteria were emphasized provides the process with critical transparency and may affect implementation.

Figure 1. When is screening justified?

Importance	<ul style="list-style-type: none"> The condition should be an important health problem (either sufficiently prevalent or has severe consequences)
Natural History	<ul style="list-style-type: none"> The condition being screened for should have a natural history that is understood and a recognized latent
Difference in management	<ul style="list-style-type: none"> Individuals with a positive screening test would get a different management than those with a negative test
Available treatment	<ul style="list-style-type: none"> There should be an effective treatment/management available for the condition that improves outcomes if administered earlier than when the condition is clinically apparent
Difference in Outcomes	<ul style="list-style-type: none"> Improvement in outcomes based on management according to screening results outweighs harms of screening (overdiagnosis, overtreatment of false positives, anxiety, stigma, etc.)
Accuracy	<ul style="list-style-type: none"> High or moderate certainty evidence for a sufficient accuracy of the test (acceptable low rates of being incorrectly diagnosed; i.e., low rates of false positives and false negatives)
Other Considerations	<ul style="list-style-type: none"> Screening should be cost effective Screening should be acceptable to patients Screening should be feasible to implement

Implementation remarks

To make the guidelines more helpful to end-users, several recommendations were followed by implementation considerations (also called technical remarks). These remarks contain practical information needed to implement and apply the recommendation. Such remarks were not subjected to systematic reviews, can be viewed to have conditional strength, and were derived from extrapolated evidence and the clinical expertise of the panel. For example, recommendations about screening with transcranial Doppler and primary stroke prevention in children with SCD required implementation advice for practitioners. This advice was that the threshold for treatment would be: 2 nonimaging transcranial Doppler ultrasound measurements with a time-averaged mean of a maximum velocity of at least 200 cm/s, or a single measurement of more than 220 cm/s. A recommendation for shared decision making about using tissue plasminogen activator in patients with SCD presenting with acute ischemic stroke was followed by remarks that included guidance on patient selection. A recommendation for screening for developmental delay and cognitive impairment was followed by signaling questions that can be used by practitioners to screen individuals with SCD. Moreover, conditional recommendations against screening asymptomatic individuals with SCD with echocardiography, pulmonary function testing, or formal polysomnography were accompanied by remarks guiding users about symptoms of pulmonary hypertension, abnormal lung function, or sleep-disordered breathing for which diagnostic evaluation would be warranted.

References

1. National Center on Birth Defects and Developmental Disabilities. Data & statistics on sickle cell disease. <https://www.cdc.gov/ncbddd/sicklecell/data.html>. Accessed 3 October 2019.
2. Pai M, Yeung CHT, Akl EA, et al. Strategies for eliciting and synthesizing evidence for guidelines in rare diseases. *BMC Med Res Methodol*. 2019;19(1):67.
3. Gandhi GY, Murad MH, Fujiyoshi A, et al. Patient-important outcomes in registered diabetes trials. *JAMA*. 2008;299(21):2543-2549.

Conclusions

The 5 ASH guideline panels faced a number of challenges related to the creation of sound, explicit, and transparent recommendations to address pressing issues in SCD care. A number of methodologic innovations ranging from patient and family engagement strategies to the consideration of noncomparative studies fostered the creation of a wide range of carefully developed and patient-centered recommendations that leverage shared decision-making wherever possible. Evidence to decision considerations intertwined with an emerging and increasingly robust evidence base facilitated the creation of what we anticipate will be unique and useful guidance.

Authorship

Contribution: All authors contributed content to this article; M.H.M. drafted the first version of this manuscript; and critical revisions were made by all authors.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: M.H.M., 0000-0001-5502-5975; E.S.L., 0000-0003-0850-4337; J.J.M., 0000-0002-1333-5403; M.R.D., 0000-0002-0574-1604.

Correspondence: M. Hassan Murad, Evidence-Based Practice Center, Mayo Clinic, 200 1st St SW, Rochester, MN 55905; e-mail: murad.mohammad@mayo.edu.

4. Hultcrantz M, Rind D, Akl EA, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol*. 2017;87:4-13.
5. Murad MH. Clinical practice guidelines: a primer on development and dissemination. *Mayo Clin Proc*. 2017;92(3):423-433.
6. Alonso-Coello P, Schünemann HJ, Moher J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ*. 2016;353:i2016.
7. Schünemann HJ, Jaeschke R, Cook DJ, et al; ATS Documents Development and Implementation Committee. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med*. 2006;174(5):605-614.
8. Yawn BP, Buchanan GR, Afenyi-Annan AN, et al. Management of sickle cell disease: summary of the 2014 evidence-based report by expert panel members. *JAMA*. 2014;312(10):1033-1048.
9. Shippee ND, Domecq Garces JP, Prutsky Lopez GJ, et al. Patient and service user engagement in research: a systematic review and synthesized framework. *Health Expect*. 2015;18(5):1151-1166.
10. Domecq JP, Prutsky G, Elraiyah T, et al. Patient engagement in research: a systematic review. *BMC Health Serv Res*. 2014;14(1):89.
11. Daraz L, Webb S, Kunkle R, Murad MH, Lang E. Training curriculum to help patient representatives participate meaningfully in the development of clinical practice guidelines [published online ahead of print 19 June 2003]. *BMJ Evid Based Med*. doi:10.1136/bmjebm-2019-111186.
12. Centers for Disease Control and Prevention. Wide-ranging online data for epidemiologic research (WONDER). <http://wonder.cdc.gov>. Accessed 6 November 2019.
13. American Society of Hematology. CDC issues key clarification on guideline for prescribing opioids for chronic pain. <https://www.hematology.org/Newsroom/Press-Releases/2019/9537.aspx>. Accessed 6 November 2019.
14. Zhang Y, Alonso-Coello P, Guyatt GH, et al. GRADE Guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences-Risk of bias and indirectness. *J Clin Epidemiol*. 2019;111:94-104.
15. Savage WJ, Buchanan GR, Yawn BP, et al. Evidence gaps in the management of sickle cell disease: a summary of needed research. *Am J Hematol*. 2015;90(4):273-275.
16. Schünemann HJ, Mustafa RA. Decision making about healthcare-related tests and diagnostic test strategies. Paper 1: a new series on testing to improve people's health. *J Clin Epidemiol*. 2017;92:16-17.
17. Mustafa RA, Wiercioch W, Cheung A, et al. Decision making about healthcare-related tests and diagnostic test strategies. Paper 2: a review of methodological and practical challenges. *J Clin Epidemiol*. 2017;92:18-28.
18. Schünemann HJ, Mustafa RA, Brozek J, et al; GRADE Working Group. GRADE guidelines: 22. The GRADE approach for tests and strategies-from test accuracy to patient-important outcomes and recommendations. *J Clin Epidemiol*. 2019;111:69-82.
19. Schünemann HJ, Mustafa R, Brozek J, et al; GRADE Working Group. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *J Clin Epidemiol*. 2016;76:89-98.
20. Mustafa RA, Wiercioch W, Ventresca M, Brozek J, Schünemann HJ; DU-Diagnosis expert group. Decision making about healthcare-related tests and diagnostic test strategies. Paper 5: a qualitative study with experts suggests that test accuracy data alone is rarely sufficient for decision making. *J Clin Epidemiol*. 2017;92:47-57.
21. Wilson J, Jungner G. The principles and practice of screening for disease. Public Health Papers no. 34. Geneva, Switzerland: World Health Organization; 1968.
22. Procedure Manual. US Preventive Services Taskforce. <https://www.uspreventiveservicestaskforce.org/Page/Name/procedure-manual>. Accessed 6 November 2019.
23. Prorok PC, Chamberlain J, Day NE, Hakama M, Miller AB. UICC Workshop on the evaluation of screening programmes for cancer. *Int J Cancer*. 1984; 34(1):1-4.