Title

Update on the Methods of the U.S. Preventive Services Task Force: Methods for Understanding Certainty and Net Benefit When Making Recommendations

Journal

American Journal of Preventive Medicine, 54(1)

ISSN

0749-3797

Authors

Krist, Alex H
Wolff, Tracy A
Jonas, Daniel E
et al.

Publication Date

2018

DOI

10.1016/j.amepre.2017.09.011

Peer reviewed

American Journal of
Preventive Medicine

SPECIAL ARTICLE

Update on the Methods of the U.S. Preventive Services Task Force: Methods for Understanding Certainty and Net Benefit When Making Recommendations

Alex H. Krist, MD, MPH,[1] Tracy A. Wolff, MD, MPH,[2] Daniel E. Jonas, MD, MPH,[3,4] Russell P. Harris, MD, MPH,[3,4] Michael L. LeFevre, MD, MSPH,[5] Alex R. Kemper, MD, MPH, MS,[6] Carol M. Mangione, MD, MSPH,[7,8] Chien-Wen Tseng, MD, MS, MPH,[9,10] David C. Grossman, MD, MPH[11,12]

Since the 1980s, the U.S. Preventive Services Task Force (USPSTF) has developed and used rigorous methods to make evidence-based recommendations about preventive services to promote health and well-being for all Americans. Recommendations are based on the evidence of magnitude of net benefit (benefits minus harms). Expert opinion is not substituted when evidence is lacking. Evidence gaps are common. Few preventive services are supported by high-quality studies that directly and comprehensively determine the overall magnitude of benefits and harms in the same study. When assessing the body of evidence, studies may not have been conducted in primary care settings, studies may not have sufficiently included populations of interest, and long-term outcomes may not have been directly assessed. When direct evidence is not available, the USPSTF uses the methodologies of applicability to determine whether evidence can be generalized to an asymptomatic primary care population; coherence to link bodies of evidence and create an indirect evidence pathway; extrapolation to make inferences across the indirect evidence pathway, extend evidence to populations not specifically studied, consider service delivery intervals, and infer long-term outcomes; and conceptual bounding to set theoretical lower or upper limits for plausible benefits or harms. The USPSTF extends the evidence only so far as to maintain at least moderate certainty that its findings are preserved. This manuscript details with examples of how the USPSTF uses these methods to make recommendations that truly reflect the evidence.

Am J Prev Med 2018;54(1S1):S11–S18.

INTRODUCTION

The mission of the U.S. Preventive Services Task Force (USPSTF) is to provide evidence-based recommendations on preventive services. Recommendations are based on a systematic review of the evidence concerning both the benefits and harms, assessment of the certainty and magnitude of net benefit, and assignment of a letter grade. This process is based on

From the [1]Department of Family Medicine and Population Health, Virginia Commonwealth University, Richmond, Virginia; [2]Center for Evidence and Practice Improvement, Agency for Healthcare Research and Quality, Rockville, Maryland; [3]Cecil G. Sheps Center for Health Services Research, University of North Carolina, Chapel Hill, North Carolina; [4]RTI International–University of North Carolina at Chapel Hill Evidence-based Practice Center, Chapel Hill, North Carolina; [5]Department of Family and Community Medicine, University of Missouri School of Medicine, Columbia, Missouri; [6]Ambulatory Pediatrics Division, Nationwide Children's Hospital, Columbus, Ohio; [7]Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, California; [8]Department of Health Policy and Management, UCLA Fielding School of Public Health, University of California, Los Angeles, California; [9]Department of Family Medicine and Community Health, University of Hawaii John A. Burns School of Medicine, Honolulu, Hawaii; [10]Pacific Health Research and Education Institute, Honolulu, Hawaii; [11]Kaiser Permanente Washington Health Research Institute, Seattle, Washington; and [12]Department of Health Services and Pediatrics, University of Washington, Seattle, Washington

Address correspondence to: Alex H. Krist, MD, MPH, One Capital Square, 6th Floor, Room 631, 830 East Main St., Richmond VA 23219. E-mail: ahkrist@vcu.edu.

This article is part of a supplement issue titled Advancing the U.S. Preventive Services Task Force Methods: Important Considerations in Making Evidence-Based Guidelines.

0749-3797/$36.00

https://doi.org/10.1016/j.amepre.2017.09.011

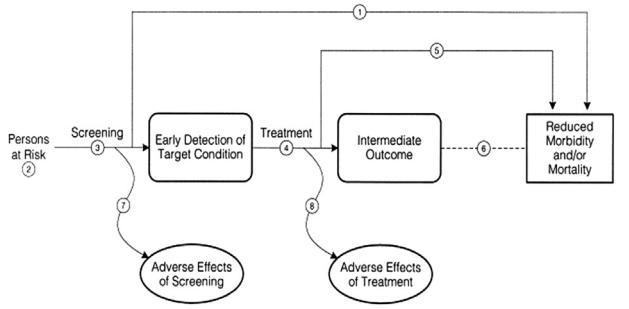

Figure 1. U.S. Preventive Services Task Force analytic framework demonstrating direct and indirect evidence path. *Note:* Key question 1 (KQ1) represents the direct evidence pathway. KQ1: Is there evidence that randomizing a cohort of persons to screening or not screening results in greater reduced morbidity and mortality than the observed harms from screening and treatment? The indirect evidence pathway is evaluated by assessing the linkage coherence between KQ2 through KQ8. KQ2: Who are the persons at risk for the condition? KQ3: What are the performance characteristics of the screening test for the condition? KQ4: Does treatment of the screen detected condition result in improved intermediate outcomes? KQ5: Does treatment of the screen detected condition result in reduced morbidity and mortality? KQ6: Does an improvement in intermediate outcomes lead to improved health outcomes? KQ 7: What are the harms associated with screening and diagnostic testing for the condition? KQ8: What are the harms associated with treatment of the condition?

a rigorous, objective, and transparent approach that is well documented and continually updated in an online procedure manual and multiple publications.[1–8]

For each preventive service, the USPSTF seeks to determine the magnitude of net benefit (i.e., benefits minus harms) for implementing a service in a primary care population without symptoms. The first step in this process is to create an analytic framework that illustrates the key questions, populations, interventions, and outcomes to be included (Figure 1); a set of inclusion and exclusion criteria are developed that provide further detail, including information on comparisons, time periods, settings, and study designs. The analytic framework is then used to guide a series of systematic reviews and later the USPSTF's assessment of the evidence. The magnitude of net benefit could be directly determined by large, well-designed RCTs in a representative asymptomatic population with appropriately long follow-up of all relevant benefits and harms for a group randomized to receive the preventive service and a group randomized to not receive the preventive service (or practices could be randomized to implement or not implement the service). This is depicted in the analytic framework as key question 1, and the USPSTF defines this as the direct evidence pathway (Figure 1). A complete direct evidence pathway is often not available for USPSTF analyses as single studies may not have large enough sample sizes or duration to show long-term effects.

For many preventive services, the USPSTF looks for additional evidence because direct evidence may be absent or the direct evidence pathway may be incomplete (e.g., rare or long-term harms are less often observed in RCTs of screening but may be reported in cohort studies). Accordingly, the USPSTF also considers the indirect evidence pathway for many preventive services. This is defined as the indirect chain of evidence that connects the target population (far left side of the analytic framework) to the ultimate health outcome (far right side of the analytic framework) through a series of linked key questions (Figure 1).[5] In addition to the evidence for benefits, evidence for harms is necessary to determine the magnitude of net benefit. Recommendations are based not only on the adequacy of the evidence for each key question, but also the overall certainty of linking the key questions to create an unbroken chain of evidence.

Frequently, there are evidence gaps in the indirect pathway that must be considered. Studies may not follow participants long enough to fully assess long-term health outcomes, which can sometimes take decades to occur. Studies may not have adequate power to identify important beneficial outcomes or identify rare harms. Studies may not be done in primary care settings or on asymptomatic average risk populations. Specific populations of interest may not have been adequately included in studies and many elements needed to make a recommendation useful to clinicians may not have been studied (e.g., age to start, age to stop, and service delivery interval).

For each USPSTF recommendation, there is first an assessment of the certainty of the evidence, followed by an estimate of the magnitude of net benefit (defined as all the benefits minus all the harms).[1,5] The combination of certainty and magnitude of net benefit determines the letter recommendation (A recommendation if there is high certainty and substantial benefit, B recommendation if there is at least moderate certainty of moderate benefit, C recommendation if there is at least moderate certainty of small benefit, and D recommendation if there is at least moderate certainty of zero or negative benefit). The USPSTF uses the concepts of applicability and coherence as part of the judgment regarding the certainty of net benefit. Similarly, it uses the concepts of extrapolation and conceptual bounding to aid in judging the magnitude of benefit. This manuscript details the USPSTF's use of these methods using examples to highlight concepts.

APPLYING SIX CRITICAL APPRAISAL QUESTIONS TO EACH KEY QUESTION

The USPSTF's first step in assessing the overall certainty of the evidence is to answer six critical appraisal questions for

Table 1. Critical Appraisal Questions: Factors Considered for Evaluating Adequacy of Evidence for Each Key Question and Across Key Questions

Critical Appraisal Questions
1. Do the studies have the appropriate research design to answer the key question(s)?
2. To what extent are the existing studies of sufficient quality? (i.e., what is the internal validity?)
3. To what extent are the results of the studies generalizable to the general U.S. primary care population of interest to the intervention and situation? (i.e., what is the applicability?)
4. How many and how large are the studies that address the key question(s)? Are the results precise?
5. How consistent are the results of the studies?
6. Are there additional factors that assist us in drawing conclusions (e.g., fit within a biologic model)?

each key question in the analytic framework. The critical appraisal questions consider whether studies have the appropriate research design; risk of bias (internal validity); applicability; precision; consistency; and plausibility (Table 1).[1] This approach builds on the work of others and is a commonly used framework for evaluating evidence.

Given its mission of making clinical preventive service recommendations for primary care, the USPSTF

Table 2. Key USPSTF Considerations about Applicability, Coherence, Extrapolation, and Conceptual Bounding as They Relate to Certainty and Magnitude of Net Benefit

	Applicability	Coherence	Extrapolation	Conceptual bounding
Certainty	• Is the evidence applicable to an asymptomatic general population? • Is the evidence applicable to U.S. primary care clinicians? • Is the intervention acceptable, feasible, and available to the U.S. primary care setting?	• Can a coherent evidence chain link the population on the left side of the framework to the beneficial or harmful health outcomes on the right side of the framework? • Does a coherent evidence chain apply to early treatment when a condition is asymptomatic? • Does linkage coherence account for naturally occurring heterogeneity of conditions? A risk threshold necessary to identify the target population? Assessments of clinically important health outcomes? Harms including the full range of harms, severity of harms, patient perspective, and timing.	• Certainty may be lower when using extrapolation, and overall certainty must remain at least moderate to make a graded recommendation.	• Certainty may be lower when using conceptual bounding, and overall certainty must remain at least moderate to make a graded recommendation.
Magnitude	N/A	N/A	• Is there a gap in the evidence chain that can be overcome with the use of extrapolation to make an assessment on magnitude of net benefit? • Is it valid and clinically relevant to use extrapolation?	• Is there a gap in the evidence about benefits or harms that could be estimated at a minimum or maximum bound to allow an assessment on magnitude of net benefit? • Is it valid and clinically relevant to use conceptual bounding?

N/A, not applicable; USPSTF, U.S. Preventive Services Task Force.

faces a unique set of issues when considering the impact of applicability on certainty (Table 2). Applicability is like generalizability or external validity, but for the USPSTF, it has a very specific meaning. The USPSTF defines applicability as the extent to which results of studies are generalizable to the general asymptomatic U.S. primary care population or specific primary care populations of interest. Frequently, studies evaluate different populations, providers, or settings than the asymptomatic primary care population. For example, the evidence about screening high-risk individuals for syphilis mainly included studies in sexually transmitted infection clinics. The USPSTF judged this evidence as applicable to a similarly high-risk patient seen in a primary care setting as the performance of the test, treatment, and follow-up would be similar in both settings.[9] However, the USPSTF did not think the evidence for the benefit of treating autism, which was primarily from trials of children referred to highly specialized treatment centers, was applicable to a screen-detected population. The concern was that the screen-detected group would likely be much younger and less severely affected than the group seen in the treatment studies and therefore may not derive the same benefit.[10] Similarly, the USPSTF did not think that the evidence about reduced daytime sleepiness resulting from treating sleep apnea in patients referred to sleep centers for symptoms would be applicable to a screen-detected population with less severe symptoms.[11,12]

A common consideration about applicability is whether studies conducted outside the U.S. are applicable to the U.S. When developing the analytic framework, the USPSTF explicitly defines whether evidence from other countries will be included. Once studies are identified, the Evidence-based Practice Centers and USPSTF again consider the applicability of each study. For many preventive services, the USPSTF considers evidence from countries categorized as "very high" on the Human Development Index to be applicable to the U.S.[13] For example, the USPSTF decided that the evidence about folic acid supplementation from developed, but not underdeveloped, countries applied to the U.S., as the underlying nutritional status and risks and factors contributing to neural tube defects would be similar.[14,15] However, some evidence from other countries may not be applicable to the U.S. population. Treatments may not be available or disease prevalence may be very different. For example, the evidence regarding celiac disease from Northern European countries may not be applicable to the U.S. because of the much higher disease prevalence in the latter.[16] Likewise evidence from developing countries were excluded from the evaluation of vitamin supplementation to prevent cancer and cardiovascular disease because of the potential for confounding related to higher rates of malnutrition in those nations.[17]

COHERENCE TO ASSESS THE CERTAINTY OF INDIRECT EVIDENCE

The next step in assessing certainty is to determine the coherence of the evidence for the indirect evidence chain. The USPSTF defines coherence as the extent to which the key questions in the indirect evidence chain of the analytic framework can be linked to make an overall assessment of a preventive service's net benefit. If direct evidence (key question 1) is available, a judgment regarding coherence is likely not needed. For coherence, the USPSTF assesses adequacy of evidence for each key question and whether the key questions can be connected to link the population on the left side of the framework to the beneficial or harmful health outcomes on the right side of the framework. Coherence is not a simple summation of the adequacy of evidence for the key questions, but is determined by answering the critical appraisal questions across the entire analytic framework.

Using screening as an example, when considering the overall linkage coherence for indirect evidence, the USPSTF assesses whether the evidence about the screening test and the evidence about early treatment can be applied to the same population and linked to approximate direct evidence. In cases where the condition is asymptomatic (e.g., for hypertension and low bone density) and where the treatment trials have been conducted on populations similar to asymptomatic patients detected by screening, the body of evidence is coherent.[18,19] By contrast, the USPSTF did not consider the evidence to be coherent for cognitive impairment screening, because evidence of net benefit came from treatment trials that recruited patients from dementia clinics rather than on asymptomatic persons.[20] Of importance, there must be coherence in the evidence both for benefits and for harms to allow for the assessment of magnitude of net benefit.

Based on answers to the critical appraisal questions for the individual key questions and the overall linkage coherence, the USPSTF makes an overall assessment of the certainty of the evidence, which is categorized as high, moderate, or low (Table 3).[1]

USE OF EXTRAPOLATION TO ESTIMATE MAGNITUDE OF NET BENEFIT

When studies do not address all key questions needed to make a recommendation, the USPSTF can next consider whether extrapolation can overcome evidence gaps. A general definition of extrapolation as applied to USPSTF

Table 3. USPSTF Levels of Certainty Regarding Net Benefit.

Level of certainty	Description
High	The available evidence usually includes consistent results from a multitude of well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on the desired health outcomes. Because of the precision of findings, this conclusion is therefore unlikely to be strongly affected by the results of future studies. These recommendations are often based on direct evidence from clinical trials of screening or behavioral interventions. High-quality trials designed as "pragmatic" or "effectiveness" trials are often of greater value in understanding external validity.
Moderate	The available evidence is sufficient to determine the effects of the preventive service on targeted health outcomes, but confidence in the estimate is constrained by factors such as: • The number, size, or quality of individual studies in the evidence pool • Some heterogeneity of outcome findings or intervention models across the body of studies • Mild to moderate limitations in the generalizability of findings to routine primary care practice. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion
Low	The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: • The very limited number or size of studies • Inconsistency of direction or magnitude of findings across the body of evidence • Critical gaps in the chain of evidence • Findings are not generalizable to routine primary care practice • A lack of information on prespecified health outcomes • Lack of coherence across the linkages in the chain of evidence More information may allow an estimation of effects on health outcomes.

USPSTF, U.S. Preventive Services Task Force.

methods is an estimation of the magnitude of benefit or harm based on extending beyond what was originally studied in screening or treatment trials. Generally, when extrapolation is used, overall certainty is lower than when no extrapolation is used, and this is often reflected in the letter grades for recommendations. However, per USPSTF methods, to make a recommendation using extrapolation, the overall certainty about the net benefit must still be at least moderate.

Examples of how the USPSTF uses extrapolation to estimate net benefit include the following:

1. extrapolate evidence about intermediate outcomes to health-related outcomes[21,22] (e.g., the USPSTF extrapolated that a reduction in falls in older adults would result in improved health outcomes)[23,24];
2. infer long-term health outcomes based on shorter-term outcomes (the USPSTF was able to infer that short-term benefits with diet and exercise counseling for patients at cardiovascular risk—improved blood pressure, BMI, and lipid profiles—would result in long-term health benefits)[25];
3. extrapolate from primary trial data to infer the intervals or duration to deliver a preventive service (e.g., the USPSTF used modeling to extrapolate from the primary trial data to recommend intervals for colorectal cancer screening and the recommended ages for and duration of aspirin therapy)[26–29];
4. extrapolate the evidence from one preventive service to a newer one (although this is uncommon and highly dependent on how closely related the services are and the context of the rest of the recommendation and the evidence); and
5. extrapolate findings from one population studied to other relevant primary care populations.

The last use of extrapolation is inter-related with the concept of applicability. As an example, the USPSTF determined when considering the evidence for abdominal aortic aneurysm screening that the results demonstrating benefits in male smokers were not applicable to female smokers because of the lower prevalence and the mixed body of evidence among female smokers and that the USPSTF could not use other data or logic to extrapolate what the net benefit would be for female smokers based on the findings in male smokers.[2,30,31]

In all the above examples, the USPSTF extrapolates from existing evidence only when the case for doing so is strong, requiring a high bar of evidence to extrapolate and still deem the certainty of evidence sufficient. The USPSTF may use non-trial evidence (e.g., epidemiologic data, natural history) for extrapolation and when possible will use formal methodologic approaches (e.g., epidemiologic models, decision models) to structure the extrapolation. For example, decision models can help simulate the effects of different interventions or health outcomes;

epidemiologic models can project the course of disease or allow inferences about etiology of disease.[3]

When there is inadequate evidence to make a judgment of the magnitude of net benefit, the USPSTF will not substitute extrapolation for evidence. This principle is exemplified for prostate cancer screening. Epidemiologic data clearly show that African American men have a higher prevalence of prostate cancer and are more likely to suffer harms from prostate cancer, yet the primary U.S. screening trial only included 4% African American men and the European trial did not report race and ethnicity.[32–34] The USPSTF considered the findings from the U.S. and European trial data to be applicable to African American men—meaning African American men would be expected to derive at least the same net benefit and harm as other men included in the studies. However, knowing that African American men have a higher prevalence and mortality from prostate cancer does not provide sufficient certainty to extrapolate greater net benefit than observed in the primary trials.[35]

When extrapolation is used to reach a letter grade recommendation, the scientific rationale for the recommendations and the methods used to review and judge the evidence are explicitly stated in the recommendation statement. If the USPSTF commissions a model, a manuscript detailing the methods and results of the modeling is provided; if existing models are reviewed, they are included in the recommendation statement or in a separate manuscript.

USE OF CONCEPTUAL BOUNDING TO ESTIMATE MAGNITUDE OF NET BENEFIT

Similarly, the USPSTF can next consider whether conceptual bounding can overcome evidence gaps when determining the net benefit. The USPSTF defines conceptual bounding as identifying a theoretical upper or lower bound for benefits or harms. The magnitude of this conceptual bound is used to set a threshold for which the existing direct evidence about the opposing benefits or harms must overcome or be below to say that a service has a net benefit or a net harm. The USPSTF must have at least moderate certainty when setting a conceptual bound in order to have at least moderate certainty in estimating the overall magnitude of net benefit.

When considering conceptual bounding, the USPSTF estimates the likely upper and lower bounds of the benefits or harms in the absence of evidence. Various inputs inform the bounds such as baseline risk of study participants, natural history, epidemiology of health conditions, invasiveness or severity of diagnostic tests or treatment, and applicability of related evidence. Based on a systematic audit of its portfolio, it was determined that the USPSTF used some form of bounding in nearly one quarter of its recommendations.[36]

A common use of upper threshold bounding by the USPSTF is to bound the benefits or harms as "no greater than small." This occurs more often for harms than benefits and most often for counseling services—the harms of nearly half of counseling recommendations are bounded as no greater than small. This is exemplified in the tobacco smoking–cessation counseling recommendation in which studies did not report harms related to counseling but the USPSTF bounded these harms as no greater than small because counseling is non-invasive and unlikely to result in serious physical or psychologic harms.[37] For benefits, if the magnitude of benefit is estimated only from studies of an intervention conducted by highly trained clinicians using specialized equipment for persons at higher risk, the observed benefit might be considered the upper bound that might reasonably be anticipated in the general population, where the population is lower risk and the setting has less-experienced clinicians. When considering screening for carotid artery stenosis, trials selected higher risk patients with carotid artery stenosis and compared carotid endarterectomy performed by skilled surgeons to an outdated medical management (many trial participants were not on lipid-lowering medications). Using this as an upper bound for benefit, the USPSTF found that observed benefits would not outweigh harms and thus there is no net benefit and likely harm with screening asymptomatic patients in primary care.[38]

Often less evidence about potential harms for screening exists, compared with evidence about potential benefits. In such situations, the USPSTF may draw general conclusions from evidence on expected yield of screening in terms of false-positive test results. If the prevalence of the condition is low and the specificity of the test is $<100\%$, the positive predictive value may be low and false-positive test results will be expected. If the diagnostic workup is invasive or otherwise carries clinically important potential for harm, as exemplified by screening for ovarian cancer,[39] the USPSTF can infer that at least some harms will result from screening, because some persons with false-positive screening tests will undergo an invasive diagnostic protocol for no possible benefit. Similarly, if overdiagnosis (and therefore labeling and overtreatment) is common, and if the treatment has some adverse effects, the USPSTF may infer that implementation of routine screening will cause at least some incremental harms, even in the absence of studies that characterize harms. This approach establishes a floor for harms and recognizes that the harms must at least be greater than what is known about the number of false-positives, the invasiveness of the

diagnostic workup, inappropriate labeling, and the expected amount of overtreatment.

Whenever the USPSTF uses conceptual bounding, it specifically calls attention to the use in the recommendation statement. For example, language from the recommendation about primary care interventions to support breastfeeding explicitly say, "There is adequate evidence to bound the potential harms of interventions to support breastfeeding as no greater than small, based on the nature of the intervention, the low likelihood of serious harms, and the available information from studies reporting few harms."[40]

FUTURE DIRECTIONS

The USPSTF continually works to advance the methods it uses to make recommendations about preventive services. The USPSTF strives to ensure that when it makes a recommendation with an A, B, C, or D grade that there is at a minimum moderate certainty of positive or zero/negative net magnitude of benefit. At the same time, the USPSTF is careful to not solely wait until there is incontrovertible evidence to judge the magnitude of net benefit, as it could run the risk of leading to patient harm by either overuse of an ineffective service or underuse of an effective service.[41] The use of applicability and coherence are key to understanding the certainty of the evidence and extrapolation and conceptual bounding in select circumstances can be used to estimate the magnitude of net benefit.

The USPSTF will continue to revisit and advance these concepts, both through evaluation of its past use of these strategies and observation of other bodies working to make evidence-based recommendations. Specific areas that deserve further attention include the role of modeling, further considerations to define when these methods are inappropriate, communicating use of these methods, the role of new data sources, and future methods advancements.

ACKNOWLEDGMENTS

Publication of this article was supported by the Agency for Healthcare Research and Quality (AHRQ) through contract #HHSA290201600006C. The U.S. Preventive Services Task Force (USPSTF) is an independent, voluntary body. The U.S. Congress mandates that AHRQ support the operations of the USPSTF. The findings and conclusions in this document are those of the authors, who are responsible for its content, and do not necessarily represent the views of AHRQ. No statement in this report should be construed as an official position of AHRQ or the U.S. Department of Health and Human Services. No financial disclosures were reported by the authors of this paper.

The authors gratefully acknowledge the helpful review and input from Drs. Kirstin Bibbins-Domingo and Iris Mabry-Hernandez; and the efforts of Colleen Barclay and Rachael Palmieri Weber to conduct the audit of U.S. Preventive Services Task Force (USPSTF) recommendations.

Daniel E. Jonas and Russell P. Harris have received funding from the Agency for Healthcare Research and Quality (AHRQ) to conduct systematic reviews for the USPSTF.

REFERENCES

1. U.S. Preventive Services Task Force. Procedure manual. https://www.uspreventiveservicestaskforce.org/Page/Name/procedure-manual. Published 2017. Accessed January 2017.
2. Bibbins-Domingo K, Whitlock E, Wolff T, et al. Developing recommendations for evidence-based clinical preventive services for diverse populations: methods of the U.S. Preventive Services Task Force. *Ann Intern Med*. 2017;166(8):565–571. https://doi.org/10.7326/M16-2656.
3. Owens DK, Whitlock EP, Henderson J, et al. Use of decision models in the development of evidence-based clinical preventive services recommendations: methods of the U.S. Preventive Services Task Force. *Ann Intern Med*. 2016;165(7):501–508. https://doi.org/10.7326/M15-2531.
4. Leipzig RM, Whitlock EP, Wolff TA, et al. Reconsidering the approach to prevention recommendations for older adults. *Ann Intern Med*. 2010;153(12):809–814. https://doi.org/10.7326/0003-4819-153-12-201012210-00007.
5. Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med*. 2007;147(12):871–875. https://doi.org/10.7326/0003-4819-147-12-200712180-00007.
6. Guirguis-Blake J, Calonge N, Miller T, Siu A, Teutsch S, Whitlock E. Current processes of the U.S. Preventive Services Task Force: refining evidence-based recommendation development. *Ann Intern Med*. 2007;147(2):117–122. https://doi.org/10.7326/0003-4819-147-2-200707170-00170.
7. Barton MB, Miller T, Wolff T, et al. How to read the new recommendation statement: methods update from the U.S. Preventive Services Task Force. *Ann Intern Med*. 2007;147(2):123–127. https://doi.org/10.7326/0003-4819-147-2-200707170-00171.
8. Petitti DB, Teutsch SM, Barton MB, Sawaya GF, Ockene JK, DeWitt T. Update on the methods of the U.S. Preventive Services Task Force: insufficient evidence. *Ann Intern Med*. 2009;150(3):199–205. https://doi.org/10.7326/0003-4819-150-3-200902030-00010.
9. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Screening for syphilis infection in nonpregnant adults and adolescents: U.S. Preventive Services Task Force Recommendation Statement. *JAMA*. 2016;315(21):2321–2327. https://doi.org/10.1001/jama.2016.5824.
10. Siu AL, Bibbins-Domingo K, Grossman DC, et al. Screening for autism spectrum disorder in young children: U.S. Preventive Services Task Force Recommendation Statement. *JAMA*. 2016;315(7):691–696. https://doi.org/10.1001/jama.2016.0018.
11. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Screening for obstructive sleep apnea in adults: U.S. Preventive Services Task Force Recommendation Statement. *JAMA*. 2017;317(4):407–414. https://doi.org/10.1001/jama.2016.20325.
12. Jonas DE, Amick HR, Feltner C, et al. Screening for obstructive sleep apnea in adults: evidence report and systematic review for the U.S. Preventive Services Task Force. *JAMA*. 2017;317(4):415–433. https://doi.org/10.1001/jama.2016.19635.
13. UN Development Programme. Human development reports. http://hdr.undp.org/en/content/human-development-index-hdi. Published 2017. Accessed February 2017.
14. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Folic acid supplementation for the prevention of neural tube defects: U.S.

Preventive Services Task Force Recommendation Statement. *JAMA*. 2017;317(2):183–189. https://doi.org/10.1001/jama.2016.19438.
15. Viswanathan M, Treiman KA, Kish-Doto J, Middleton JC, Coker-Schwimmer EJ, Nicholson WK. Folic acid supplementation for the prevention of neural tube defects: an updated evidence report and systematic review for the U.S. Preventive Services Task Force. *JAMA*. 2017;317(2):190–203. https://doi.org/10.1001/jama.2016.19193.
16. U.S. Preventive Services Task Force. Celiac disease: screening. www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryDraft/celiac-disease-screening?ds=1&s=celiac. Published 2017. Accessed February 2017.
17. Fortmann SP, Burda BU, Senger CA, Lin JS, Whitlock EP. Vitamin and mineral supplements in the primary prevention of cardiovascular disease and cancer: an updated systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2013;159(12):824–834. https://doi.org/10.7326/0003-4819-159-12-201312170-00729.
18. Siu AL. Screening for high blood pressure in adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2015;163(10):778–786. https://doi.org/10.7326/M15-2223.
19. U.S. Preventive Services Task Force. Screening for osteoporosis: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2011;154(5):356–364. https://doi.org/10.7326/0003-4819-154-5-201103010-00307.
20. Moyer VA. Screening for cognitive impairment in older adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2014;160(11):791–797. https://doi.org/10.7326/M14-0496.
21. Jonas DE, Ferrari RM, Wines RC, Vuong KT, Cotter A, Harris RP. Evaluating evidence on intermediate outcomes: considerations for groups making healthcare recommendations. *Am J Prev Med*. 2018;54(1S1):S38–S52.
22. Wolff TA, Krist AH, LeFevre M, et al. Update on the methods of the U.S. Preventive Services Task Force: linking intermediate outcomes and health outcomes in prevention. *Am J Prev Med*. 2018;54(1S1):S4–S10.
23. Moyer VA. Prevention of falls in community-dwelling older adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2012;157(3):197–204. https://doi.org/10.7326/0003-4819-157-3-201208070-00462.
24. Michael YL, Whitlock EP, Lin JS, Fu R, O'Connor EA, Gold R. Primary care-relevant interventions to prevent falling in older adults: a systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2010;153(12):815–825. https://doi.org/10.7326/0003-4819-153-12-201012210-00008.
25. LeFevre ML. Behavioral counseling to promote a healthful diet and physical activity for cardiovascular disease prevention in adults with cardiovascular risk factors: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2014;161(8):587–593. https://doi.org/10.7326/M14-1796.
26. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *JAMA*. 2016;315(23):2564–2575. https://doi.org/10.1001/jama.2016.5989.
27. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of benefits, burden, and harms of colorectal cancer screening strategies: modeling study for the U.S. Preventive Services Task Force. *JAMA*. 2016;315(23):2595–2609. https://doi.org/10.1001/jama.2016.6828.
28. Bibbins-Domingo K. Aspirin use for the primary prevention of cardiovascular disease and colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2016;164(12):836–845. https://doi.org/10.7326/M16-0577.
29. Dehmer SP, Maciosek MV, Flottemesch TJ, LaFrance AB, Whitlock EP. Aspirin for the primary prevention of cardiovascular disease and colorectal cancer: a decision analysis for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2016;164(12):777–786. https://doi.org/10.7326/M15-2129.
30. LeFevre ML. Screening for abdominal aortic aneurysm: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2014;161(4):281–290. https://doi.org/10.7326/M14-1204.
31. Guirguis-Blake JM, Beil TL, Senger CA, Whitlock EP. Ultrasonography screening for abdominal aortic aneurysms: a systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2014;160(5):321–329. https://doi.org/10.7326/M13-1844.
32. Andriole GL, Crawford ED, Grubb RL 3rd, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst*. 2012;104(2):125–132. https://doi.org/10.1093/jnci/djr500.
33. Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360(13):1320–1328. https://doi.org/10.1056/NEJMoa0810084.
34. Schroder FH, Hugosson J, Carlsson S, et al. Screening for prostate cancer decreases the risk of developing metastatic disease: findings from the European Randomized Study of Screening for Prostate Cancer (ERSPC). *Eur Urol*. 2012;62(5):745–752. https://doi.org/10.1016/j.eururo.2012.05.068.
35. U.S. Preventive Services Task Force. U.S. Preventive Services Task Force: prostate cancer screening draft recommendations. www.screeningforprostatecancer.org. Published 2017. Accessed April 2017.
36. Jonas DE, Barclay C, Palmieri Weber R, Harris RP. When and how has the U.S. Preventive Services Task Force used conceptual bounding? An audit of the current portfolio. (Prepared by RTI International-University of North Carolina Evidence-based Practice Center [EPC] under Contract No. HHSA290-2012-0001-5i.) Rockville, MD: Agency for Healthcare Research and Quality. www.uspreventiveservicestaskforce.org/Page/Name/methods-and-processes. Published 2017. Accessed August 2017.
37. Siu AL. Behavioral and pharmacotherapy interventions for tobacco smoking cessation in adults, including pregnant women: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2015;163(8):622–634. https://doi.org/10.7326/M15-2023.
38. LeFevre ML. Screening for asymptomatic carotid artery stenosis: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2014;161(5):356–362. https://doi.org/10.7326/M14-1333.
39. Moyer VA. Screening for ovarian cancer: U.S. Preventive Services Task Force reaffirmation recommendation statement. *Ann Intern Med*. 2012;157(12):900–904. https://doi.org/10.7326/0003-4819-157-11-201212040-00539.
40. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Primary care interventions to support breastfeeding: U.S. Preventive Services Task Force recommendation statement. *JAMA*. 2016;316(16):1688–1693. https://doi.org/10.1001/jama.2016.14697.
41. Harris RP, Helfand M, Woolf SH, et al. Current methods of the U.S. Preventive Services Task Force: a review of the process. (suppl). *Am J Prev Med*. 2001;20(3):21–35. https://doi.org/10.1016/S0749-3797(01)00261-6.