Update on the Methods of the U.S. Preventive Services Task Force: Linking Intermediate Outcomes and Health Outcomes in Prevention

Tracy A. Wolff, MD, MPH,1 Alex H. Krist, MD, MPH,2 Michael LeFevre, MD, MSPH,3 Daniel E. Jonas, MD, MPH,4,5 Russell P. Harris, MD, MPH,4,5 Albert Siu, MD, MSPH,5 Douglas K. Owens, MD, MS,7,8 Matthew W. Gillman, MD, SM,9,10 Mark H. Ebell, MD, MS,11 Jessica Herzstein, MD, MPH,12 Roger Chou, MD,12 Evelyn Whitlock, MD, MPH,14 Kirsten Bibbins-Domingo, PhD, MD, MAS15

The U.S. Preventive Services Task Force (USPSTF) is an independent body of experts who make evidence-based recommendations about clinical preventive services using a transparent and objective process. Developing recommendations on a clinical preventive service requires evidence of its effect on health outcomes. Health outcomes are symptoms, functional levels, and conditions that affect a patient’s quantity or quality of life and are measured by assessments of physical or psychologic well-being. Intermediate outcomes are pathologic, physiologic, psychologic, social, or behavioral measures related to a preventive service. Given the frequent lack of evidence on health outcomes, the USPSTF uses evidence on intermediate outcomes when appropriate. The ultimate goal is to determine precisely a consistent relationship between the direction and magnitude of change in an intermediate outcome with a predictable resultant direction and magnitude of change in the health outcomes. The USPSTF reviewed its historical use of intermediate outcomes, reviewed methods of other evidence-based guideline-making bodies, consulted with other experts, and reviewed scientific literature. Most important were the established criteria for causation, tenets of evidence-based medicine, and consistency with its current standards. Studies that follow participants over time following early treatment, stratify patients according to treatment response, and adjust for important confounders can provide useful information about the association between intermediate and health outcomes. However, such studies remain susceptible to residual confounding. The USPSTF will exercise great caution when making a recommendation that depends on the evidence linking intermediate and health outcomes because of inherent evidence limitations.

Am J Prev Med 2018;54(1S1):S4–S10. Published by Elsevier Inc. on behalf of American Journal of Preventive Medicine. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
INTRODUCTION

The U.S. Preventive Services Task Force (USPSTF or Task Force), an independent body of experts who make evidence-based recommendations about clinical preventive services, strives continuously to advance its methods of making recommendations using a clear, transparent, unbiased process. To make a recommendation on a clinical preventive service, evidence must show an effect on health outcomes (HOs, outcomes that patients can experience or feel and affect how long a patient lives or the quality of life or both). Most studies on clinical preventive services are not designed to show an effect on HOs but more often show improvements in intermediate outcomes (IOs). Studies reporting HOs are not often available because of the long time frames and sample sizes required to show an effect. The methods to use evidence on IOs are not well established so the USPSTF commissioned a report on the use of IOs in evidence-based recommendations and recently updated its methods on how it will consider evidence on IOs.

The purpose of this paper is to describe a transparent method of assessing the evidence linking IOs to HOs to determine precisely a consistent relationship between the direction and magnitude of change in an IO with a predictable resultant direction and magnitude of change in the HOs.

BACKGROUND ON THE U.S. PREVENTIVE SERVICES TASK FORCE ANALYTIC FRAMEWORK

The USPSTF uses an analytic framework (Figure 1) to guide reviews of the evidence on a specific topic and to clarify the linkages in the evaluation of the benefits and harms of the proposed preventive service. Key questions (KQs) establish the necessary steps in the clinical logic that must be demonstrated to evaluate the benefits and harms of a clinical preventive service in primary care. KQs articulate the key aspects of the relevant populations, interventions, and outcome. For more information on the analytic framework see the procedure manual on the USPSTF website and more information on the use of indirect evidence throughout the analytic framework is available in Krist and colleagues. The focus of this manuscript is assessing the evidence for the KQ linking IOs to HOs (KQ6 in Figure 1).

HOW DOES THE TASK FORCE DEFINE INTERMEDIATE OUTCOMES AND HEALTH OUTCOMES?

IO is the term the Task Force uses to describe outcomes that may be influenced by a preventive service, but are not HOs in and of themselves. They are pathologic, physiologic, psychologic, social, or behavioral measures related to a preventive intervention. Examples include blood pressure, serum cholesterol, vitamin levels, viral levels, and physical activity measures. IOs are differentiated from HOs; to make a screening recommendation, the Task Force requires the evidence to demonstrate an effect on HOs, not just IOs. HOs, as used by the Task Force, are generally things that affect how long a patient lives or the quality of life and are often described as conditions that a patient can feel or experience. Alterations in physical or psychologic well-being (e.g., symptomatic disease) may affect the quality of life, and are usually considered HOs by the Task Force. Examples of HOs, as defined by the Task Force, include pain or dyspnea, functional status, disease-related quality of life, and child development.

DEVELOPING AN APPROACH FOR USING INTERMEDIATE OUTCOMES IN U.S. PREVENTIVE SERVICES TASK FORCE EVIDENCE-BASED RECOMMENDATIONS

In refining its approach to IOs, the Task Force went through a careful process that included reviewing its historical use of
IOs, reviewing methods of other evidence-based guideline-making bodies, consulting with other experts in evidence-based guideline methods, and reviewing scientific literature. The Task Force also considered the principles of establishing causation and of evidence-based medicine. The Task Force convened a workgroup of members who considered all of this information and had focused discussions about how to refine its methods.

The review of its portfolio revealed that the Task Force was generally consistent in considering the link between IOs and HOs, although the approach to gathering evidence on the link varied. This variation was primarily based on the characteristics of the IO and how well established the evidence base was supporting the IO. When the IO was more distant pathophysiologically from the HO, the Task Force generally used a full systematic review of the evidence linking the IO to HOs (e.g., the link between viremia and HOs in HIV Screening). When the IO was either very well established or very proximal pathophysiologically to the HO, the Task Force used selected reviews or considered the association established (often based on previous systematic reviews; e.g., the link between tobacco cessation and improvement in HOs).

The Task Force found little about assessing the adequacy of evidence for the link between IOs and HOs in the published literature, and few guideline-making groups discuss it in their methodology papers. More details are provided by Jonas and colleagues.

### APPRAOCH TO CONSIDERING THE LINK BETWEEN INTERMEDIATE AND HEALTH OUTCOMES

#### Types of Evidence

After following this careful process, the Task Force determined that the ultimate goal for assessing the evidence linking IOs and HOs is to determine precisely whether a consistent relationship exists between the direction and magnitude of change in an IO and a predictable resultant direction and magnitude of change in the HOs. As with other Task Force KQs, evidence should include the most rigorous studies available that evaluate primary care and prevention relevant interventions, populations, and settings. The major challenge is the lack of trials evaluating HOs and, therefore, the need for longitudinal studies that capture longitudinal outcomes that are more susceptible to bias; these studies are less likely to be performed because of the need for long follow-up times and large sample size. Evidence to evaluate the IO and HO link is likely to come from longitudinal studies that follow participants who received early treatment, stratify participants according to their response to treatment as measured by IOs, report HOs in each group, and determine the magnitude of association between IOs and HOs, with findings adjusted for important confounders. For example, a randomized trial may be followed by an analysis, in which patients are analyzed according to how they responded to an intervention, rather than the treatment to which they were allocated. An inherent limitation of such studies is that it is no longer randomized, and is therefore susceptible to confounding and other limitations of observational analyses. The usefulness of this type of evidence in evaluating the link between IOs and HOs will be, in part, dependent on the level and appropriateness of adjustment for known confounders, though even studies that adjust well are susceptible to residual confounding. Cohort studies that report changes in IOs and HOs following treatment may provide similar information. The Task Force may consider other types of observational evidence that provide epidemiologic support for causation, such as studies that assess the association between an IO and HO but do not evaluate changes in IOs and HOs as the result of a specified intervention.

### Assessing the Adequacy of the Evidence for the Intermediate to Health Outcome Link

In updating its approach to IOs, the Task Force considered the established criteria for causation (i.e., Bradford Hill Criteria); tenets of evidence-based medicine; and consistency with its current standards (Table 1). Currently, the Task Force asks a number of critical appraisal questions to determine the adequacy of evidence for each linkage (KQ) in its analytic framework; the Task Force has three categories for adequacy based on the answers to these critical appraisal questions: convincing, adequate, and inadequate. When considering the IO-HO linkage, the USPSTF applies the critical appraisal questions to the body of literature and the appropriateness of the study designs to answer the KQs (individual study quality is addressed at an earlier step in the process); the risk of bias in the body of literature; the generalizability; the number and size of available studies; the precision of the pooled results; and the consistency of results across studies.

The Task Force also considers whether there are additional factors that assist in drawing conclusions (critical appraisal question 6 in Table 1). Additional factors that may be considered for the IOs to HOs link include biological plausibility; proximity; adjustment for confounding; benefits for other outcomes; duration of studies; evidence from other interventions/exposures or populations; magnitude of association; dose–response; relationship to treatment; and positive (but not statistically significant) direction of effects on an HO.
Table 1. Factors Considered for Evaluating Adequacy of Evidence for Key Questions (Critical Appraisal Questions and Definitions)

<table>
<thead>
<tr>
<th>Critical appraisal questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do the studies have the appropriate research design to answer the key question(s)?</td>
</tr>
<tr>
<td>2. To what extent are the existing studies of sufficient quality (i.e., what is the internal validity)?</td>
</tr>
<tr>
<td>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)?</td>
</tr>
<tr>
<td>4. How many and how large are the studies that address the key question(s)? Are the results precise?</td>
</tr>
<tr>
<td>5. How consistent are the results of the studies?</td>
</tr>
<tr>
<td>6. Are there additional factors that assist us in drawing conclusions (e.g., fit within a biological model)?</td>
</tr>
</tbody>
</table>

**Definitions for levels of adequacy for key questions**

- **Convincing:** After evaluating the evidence using the critical appraisal questions, the Task Force concludes that there are sufficient well-conducted studies of appropriate design that demonstrate consistent and precise results focused on outcomes and generalizable to the intended U.S. primary care population and setting. The consistency of a large number of individual studies and lack of heterogeneity in pooled results strengthens the case for the evidence to be deemed convincing.

- **Adequate:** After evaluating the evidence using the critical appraisal questions, the Task Force concludes that the evidence is sufficient to answer a key question, but it is less convincing because of one or more significant limitations in factors, such as the appropriateness of study design, quality of studies, applicability of results, overall precision, and/or heterogeneity of evidence.

- **Inadequate:** After evaluating the evidence using the critical appraisal questions, the Task Force concludes that the evidence is insufficient to answer a key question because of a complete lack of evidence or a fatal flaw in one or more of the following factors: consistency of results, precision, applicability, and/or study quality and design. Inadequate evidence (for either benefits or harms) may create a critical gap in the evidence chain.

HO, health outcome; IO, intermediate outcome; Task Force, U.S. Preventive Services Task Force.

In general, if the answers to all of the critical appraisal questions (Table 1) are positive, the Task Force considers the evidence convincing; that is, there are sufficient well-conducted studies of appropriate design that demonstrate consistent and precise results for linking an IO to an HO in the intended U.S. primary care population and setting. Evidence for a KQ is deemed as adequate when there is enough evidence to answer a KQ, but there are one or more significant limitations in factors, such as the appropriateness of study design, quality of studies, applicability of results, overall precision, and heterogeneity of evidence. Evidence for a KQ is inadequate when there is not enough evidence to answer a KQ because of a complete lack of evidence or a fatal flaw in one or more of the following factors: consistency of results, precision, generalizability, and internal validity and study design. Adequacy of the evidence at the KQ level is the first step in the USPSTF deliberation on the evidence; adequacy at the KQ level contributes to the Task Force’s assessment of overall certainty across the entire analytic framework that, along with the magnitude of net benefit, determines the grade of the overall recommendation. For more detail on the Task Force’s methods and the grades, the reader is referred to the USPSTF Procedure Manual.

The Task Force exercises caution when making a recommendation that depends in large part on the evidence linking IOs and HOs. The Task Force infrequently deems the evidence convincing to establish the linkage because of the inherent limitations of the evidence linking IOs and HOs. Strong associations between changes in IOs that are not linked to a specific intervention and HOs (e.g., association between uric acid and cardiovascular disease without evidence on reduction in cardiovascular disease from urate-lowering treatment) would be deemed less informative than studies in which responses are due to specific interventions, and cross-sectional studies on the association between IOs and HOs would likely not be considered to establish this link.

**EXAMPLES OF WHEN THE TASK FORCE USED THE EVIDENCE ON INTERMEDIATE OUTCOMES**

The Task Force’s recommendation on screening HCV is an example of when the Task Force used the evidence linking IOs to HOs to make a recommendation (the dotted line in Figure 2). In this case, sustained virologic
response (lack of detectable HCV) served as the IO. Although adequate evidence suggested that screening with anti-HCV antibody testing was accurate and that there were limited harms of screening and early treat-
ment,8,9 there was inadequate direct evidence that screening improves morbidity and mortality and inadequate evidence that treatment leads to improved HOs. There was adequate evidence that evidence that treatment with antiviral regimens resulted in sustained virologic response (an IO); and 19 longitudinal studies reported that a sustained virologic response was strongly linked with improved HOs (i.e., reductions in hepatic complications, hepatocellular carcinoma, and mortality). Using the evidence linking IOs and HOs, the Task Force made a recommendation to screen specific groups of adults.

The criteria elucidated by Jonas and colleagues2 (additional factors discussed in Assessing the Adequacy of the Evidence for the Intermediate to Health Outcome Link and in Table 1) highlight many of the considerations that the Task Force used to assess sustained virologic response as an IO and the adequacy of its link to HOs. HCV is the causative agent for hepatitis C–related morbidity and mortality, so elimination of detectable virus could reasonably be assumed to be related to HOs (biological plausibility). Furthermore, HCV induces hepatic fibrosis (another IO in this pathway), and evidence links sustained virologic response to lack of progression of hepatic fibrosis, which is more closely tied to HOs (proximity). There was a high degree of consistency in the longitudinal studies linking sustained virologic response to HOs, and pooling of underpowered studies linking sustained virologic response to hepatocellular carcinoma also showed this relationship (positive direction of effects on HO). Although the treatment trials were not large enough to determine the effect of treatment on HOs, the direction of the effect on mortality and hepatocellular carcinoma was also consistent.

Although there was adequate evidence to make a recommendation, the HCV recommendation highlights some of the challenges in using evidence on the link between IOs and HOs. Although the Task Force found consistent evidence that improvement in virologic IOs (sustained viral response) following antiviral therapy was associated with improvement in HOs, there was some imprecision around the estimates and uncertainty with regard to the degree to which residual confounding could have affected the estimates. This is important because the Task Force is not interested solely in whether an association exists, but in determining the strength of the association (which is necessary to estimate the magnitude of the benefit). Studies of HCV treatment did not consistently adjust for the same set of confounding factors, which is a common challenge when synthesizing observational studies. To address this challenge,
the Task Force required that studies it used to establish this link must have adjustment for confounding for a key set of variables—factors associated with worse clinical outcomes (e.g., age, sex, genotype, fibrosis stage). Most studies evaluated patients with more advanced fibrosis (evidence from other populations), such that applicability to screen-detected populations was less certain. Finally, because of the long duration of time between screening and effect on HOs, assessment of the primary evidence base posed additional challenges; in this case, a published simulation model provided additional evidence of the potential benefits of screening based on the published primary studies.10 Given the limitations, the evidence was determined to be adequate (but not convincing) and the Task Force was able to make a recommendation to screen for HCV.

An example of when the Task Force was unable to make a recommendation because of inadequate evidence on the link between IOs and HOs is screening for iron deficiency anemia in young children (aged 6–24 months; Figure 3).11,12 In its review of evidence, the Task Force found convincing evidence on the accuracy of screening with hemoglobin measurement and limited evidence on treatment and iron-related levels. However, no studies (neither trials nor observational studies) were found in relevant populations and settings that assessed the association between change in iron status (the IO) as a result of intervention and improvement in child HOs (e.g., growth, neurodevelopmental outcomes). Therefore, the Task Force was unable to make a recommendation for or against screening for iron deficiency anemia in young children (I statement).

CONCLUSION AND NEXT STEPS
The Task Force may consider the evidence linking IOs and HOs when making recommendations and will judge the adequacy of that evidence based on an established, transparent approach with updated considerations throughout the Task Force’s processes (Table 2). Assessing IOs has a unique set of challenges for children and the Task Force presents these in another manuscript in this supplement by Kemper and colleagues.13 The Task Force’s methods of using the evidence linking IOs and HOs may be useful to other evidence-based guideline-making bodies when there is limited direct evidence.

The Task Force plans to focus future efforts to further advance its approach related to defining, identifying, and linking IOs and HOs. Other future work might include exploration of the unique issues associated with counseling for healthful behaviors, use of modeling to understand the relationship between IOs and HOs, and the appropriateness of establishing surrogate outcomes (“an outcome meant to substitute for a clinically meaningful endpoint and predicts change in health outcomes”)14 that could be used across multiple recommendations (e.g., magnitude of change in blood pressure as a surrogate for HOs across multiple related cardiovascular recommendations).

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, the U.S. Department of Health and Human Services, or the National Institutes of Health. No financial disclosures were reported by the authors of this paper.

Dan Jonas and Russell Harris were supported in their work by the Agency for Healthcare Research and Quality, U.S. DHHS Contract No.: HHSA29020100005i, Task Order #3. Roger

---

Table 2. Updated Approach to Considering the Link Between Intermediate and Health Outcomes

<table>
<thead>
<tr>
<th>1. At the research plan phase, the USPSTF will determine if the IO-HO link might be important in the USPSTF deliberation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. If IO-HO link likely to be important, insert dotted line in AF for the IO-HO link.</td>
</tr>
<tr>
<td>b. Determine level of review necessary to understand IO-HO link and clearly articulate justification and any key questions in the research plan.</td>
</tr>
<tr>
<td>2. Systematic review will synthesize evidence in a way that the USPSTF can use the critical appraisal questions with additional factors to assess the evidence on the IO-HO link.</td>
</tr>
<tr>
<td>3. Assessment by the USPSTF of all evidence using established processes with special attention to the adequacy of the evidence for the IO-HO link using the critical appraisal questions (and additional factors).</td>
</tr>
<tr>
<td>4. Given the inherent limitations of the evidence base for the IO-HO link, the USPSTF will exercise great caution when making a recommendation that depends in large part on the evidence linking intermediate and health outcomes.</td>
</tr>
<tr>
<td>5. The USPSTF will clearly articulate its assessment of the adequacy of the evidence and the rationale for using the evidence on intermediate outcomes in its recommendation statement.</td>
</tr>
<tr>
<td>6. The USPSTF will communicate in its recommendation statement the research gaps related to intermediate outcomes.</td>
</tr>
</tbody>
</table>

AF, analytic framework; IO-HO, intermediate outcomes to health outcomes; USPSTF, U.S. Preventive Services Task Force.
Chou and Dan Jonas have received funding from the AHRQ to conduct systematic reviews for the USPSTF.

REFERENCES


