Behavioral counseling interventions to promote healthy behaviors can significantly reduce leading causes of disease and death. Recommendations for delivery of these interventions in primary care have been and continue to be an important part of the U.S. Preventive Services Task Force’s portfolio of clinical preventive services recommendations. However, primary and secondary research on the effectiveness of behavioral counseling interventions can be more complex than recommendations for screening or use of preventive medications. The nature of behavior change and interventions to promote it can lead to unique challenges. This paper summarizes and expands upon an extensive discussion held at the U.S. Preventive Services Task Force’s Expert Forum on behavioral counseling interventions held in November 2013. The paper describes the foundational challenges for using behavioral outcomes as evidence to support a Task Force recommendation. The paper discusses research design and reporting characteristics needed by behavioral counseling intervention researchers in order for their research to contribute to the evidentiary basis of a Task Force recommendation. Finally, the paper identifies critical issues that need to be considered by the Task Force and other stakeholders to maintain confidence and credibility in the standards of evidence for behavioral counseling recommendations.

(Am J Prev Med 2015;49(3S2):S150–S157) © 2015 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/).
describes the primary challenges for accepting behavioral outcomes as sufficient evidence to make a recommendation, discusses the ideal characteristics of research design and implementation to support evidence-based behavioral counseling recommendations, and offers suggestions for future consideration.

**Issues of Concern**

**Linking Behavior Change and Health Outcomes**

The USPSTF’s recommendations are based on scientific evidence that people who receive the preventive service experience better health outcomes than those who do not. In addition, the benefits must be large enough to outweigh the harms. The USPSTF looks for evidence to support specific linkages in the analytic framework (Figure 1), but overall a complete causal chain from intervention to outcome must be supported by evidence. Outcomes that patients can feel or care about (e.g., pain, function, survival) are weighted more heavily than intermediate or surrogate outcomes. Behavioral counseling intervention recommendations are generally expected to meet the same standards of evidence for linking interventions and health outcomes, but the nature of behavior change and interventions to promote it leads to some unique challenges compared with evidence for screening or use of preventive medications.

Time to outcome is one important challenge. The time frame for identifying a link between sustained behavior change and a health outcome can be as long as 10–20 years, and even longer in interventions targeting children. For example, eating a healthful diet and engaging in physical activity may lower the risk of obesity and hypertension, but it may take decades for the cardiovascular morbidity and mortality benefits to emerge. Likewise, stopping smoking can have immediate health benefits, including reductions in cardiovascular disease risks, but the long-term benefits of reduction in lung cancer, chronic obstructive pulmonary disease, and cardiovascular disease will not appear for many years. Most RCTs do not have long enough follow-up periods to demonstrate these important linkages between the behavioral outcome (e.g., stopping smoking) and the health outcome (e.g., lung cancer).

The target population for USPSTF behavioral counseling recommendations presents another challenge for linking behavioral outcomes and health outcomes. The USPSTF recommendations apply only to asymptomatic individuals or to those with unrecognized signs or symptoms of the target condition for which the preventive service is intended. These individuals may also be at lower risk for the condition being prevented. Effective behavioral counseling interventions can disrupt the natural history of the behavior–disease relationship and lead to continued low rates of morbidity and mortality. However, continued low outcome event rates in low-risk people make it difficult to demonstrate benefits. For example, although behavioral counseling interventions for improving diet and exercise in average-risk individuals showed changes in healthy behaviors (e.g., increased fruit and vegetable intake, increase minutes of exercise) and improvements in intermediate outcomes (e.g.,

---

**Figure 1.** Analytic framework for behavioral counseling interventions.

*Note: Key questions: 1. Do changes in patients’ health behavior improve health or reduce risk factors? 2. What is the relationship between duration of health behavior change and health improvement (i.e., minimum duration, minimum level of change, and change–response relationship)? 3. What are the adverse effects of health behavior change? 4. Does health behavior change produce other positive outcomes (e.g., patient satisfaction, changes in other health care behaviors, improved function, and decreased use of health care resources)? 5. Is risk factor reduction or measured health improvement associated with reduced morbidity or mortality? 6. Is sustained health behavior change related directly to reduced morbidity or mortality? 7. Are behavioral counseling interventions in clinical care related directly to improved health or risk factor reduction? 8. Are behavioral counseling interventions in clinical care related directly to reduced morbidity or mortality?*

settings is important as well. Because the USPSTF applicable to wider practice, the applicability of the trial participants in these trials are not widely representative of behaviors. Ideally, studies are conducted in populations and settings to accept behavioral outcomes as sufficient evidence of effectiveness, what are the ideal characteristics of the research needed to make an evidence-based determination around counseling interventions? The next section of the paper describes the key considerations and challenges of the research quality and evaluation of the evidence itself. Although many of these considerations are not unique to the behavioral counseling literature, explicitly highlighting the differences could help guide researchers as they design and implement their studies; funders as they prioritize investment in new research; and payers, guideline developers, and other decision makers as they create policies to disseminate best practices.

**Populations/Settings**

Ideally, studies are conducted in populations and settings that are closely representative of the target population, that is, for whom they are intended in actual practice. The USPSTF focuses on well-conducted trial-level evidence of benefit for counseling interventions in primary care. Often, participants in these trials are not widely representative of the community’s primary care population. Although this issue of efficacy versus effectiveness is not unique to behavioral counseling trials, it can be exaggerated in this evidence base, in that trials often have more-restrictive inclusion criteria for entry into the study than the general primary care population of interest. Participants may be mostly or partially volunteers and may have high attrition of eligible compared with enrolled participants. Problems extrapolating from volunteer samples may be particularly exaggerated in studies of behavior change, where patients may view enrollment as a means to achieve their own personal goal, desiring the added support and structure a study could provide. Unfortunately, important information about who is or is not enrolled, indicators or descriptors of readiness to change, and other factors such as adherence and degree of efficacy are not commonly reported.

Just as certain trial populations’ findings might be less applicable to wider practice, the applicability of the trial settings is important as well. Because the USPSTF’s target audience is primary care clinicians, their recommendations center on behavioral counseling interventions that can be readily delivered in or referred from primary care. Interventions conducted in community settings like schools, workplaces, and churches may not be applicable to primary care practice. Community-based counseling interventions may not be available for referral, and some aspects of the intervention may simply be impossible to replicate in primary care settings. For example, counseling interventions in settings in which there is an existing social network may perform differently with groups of strangers, or worksite settings may be able to provide onsite activities such as lunchtime group walks, which would not be useful in most primary care settings. Other important setting details, like the years in which the study took place and the geographic location may affect if and how the evidence is considered, as behaviors and cultural norms change over time and from place to place.

The primary population as well as any important subpopulations of interest should be defined a priori, that is, before the start of the trial. Populations studied in behavioral counseling trials can include a broad range of participants, or can focus on narrow, targeted populations (e.g., subpopulations at higher risk for bad health outcomes or with a disproportionate need for counseling). Well-conducted trials that include a broad range of participants are helpful in understanding the overall efficacy or effectiveness of a counseling intervention. However, these trials may be more costly to conduct and may have underwhelming effects as a result of larger findings in the whole population obscuring smaller subpopulation effects.

More often, a body of counseling literature consists of a range of trials that target a variety of specific populations. For example, the USPSTF recommendation for counseling to prevent sexually transmitted infections (STIs) excluded primary trials conducted exclusively in HIV-positive people. Although there are HIV-positive individuals seen in primary care who need STI prevention counseling, the behaviors and motivations in this population were deemed unique to this subgroup; therefore, it is unreasonable to extrapolate findings from this subgroup to a broader primary care population. Multiple studies conducted in narrow populations will not necessarily provide meaningful evidence for broader population recommendations.

**Interventions**

One of the distinguishing features of behavioral counseling interventions is that they are usually more complex than other preventive services such as screening and use of preventive medications. Clear reporting of the intervention components is important, as is the intensity,
frequency, and duration of the intervention; who delivered the intervention; and the fidelity of its delivery and adherence of the participant to various components of the intervention. This information is not routinely or consistently reported. Without knowing this level of detail, it is difficult to assess the reproducibility of findings. Many counseling interventions may appear to be “one-offs,” such that there is limited evidence for similar interventions. The more complex an intervention is, the less likelihood there is for reproducibility.

The issue of aggregating “like” interventions is not unique to counseling interventions. It is analogous to understanding the preventive benefit of a medication. For example, would it be reasonable to group all medications to lower cholesterol versus a specific class of medications (e.g., statins); versus a specific drug (e.g., simvastatin); versus a specific dose (e.g., simvastatin 40 mg); versus a specific dose, frequency, duration (e.g., simvastatin 40 mg daily for 12 months), when evaluating effectiveness of chemoprevention? The challenges are magnified with the complexity of counseling interventions. For example, within the healthy lifestyle counseling literature, even those interventions restricted to similar populations (e.g., people with cardiovascular risk factors) and with a similar primary aim (e.g., reduce cardiovascular disease), the interventions varied widely in the messages (e.g., different dietary and physical activity targets); components (e.g., diaries, goal setting, social support, problem solving, motivational interviewing); adjuts (e.g., low-cost exercise opportunities, cooking demonstrations); intensity (e.g., 30 minutes to greater than 20 hours); and frequency and duration (e.g., single visit to weekly sessions for more than 1 year). This level of detailed reporting is necessary to aggregate and then compare “like” counseling interventions and to be explicit about what works. Additionally, details on the fidelity of the delivery of the intervention, who is delivering the counseling, as well as the adherence of participants to the intervention allow for judgment of the efficacy versus effectiveness of any particular intervention, and equally important, awareness and consideration around implementation issues with counseling in real-world settings.

The USPSTF is primarily concerned with making recommendations about whether behavioral counseling works, but not necessarily which counseling interventions work better (i.e., comparative effectiveness). This means much of the behavioral counseling evidence in any given topic may not be considered. The same principles of needing adequate detail to understand the intervention and compare “like” interventions are needed. Although the USPSTF can provide some guidance on parameters that increase the likelihood that counseling will be effective (e.g., must be intensive rather than brief), comparative effectiveness trials are often better for determining necessary and sufficient components of effective interventions. For implementation of counseling interventions, researchers, payers, and other decision makers are often interested in the minimal effective counseling “dose” for any given behavior. In the USPSTF’s recent recommendation for healthful diet and physical activity to prevent cardiovascular disease, they were able to identify a mean range of the number of contacts reported but the evidence base did not allow them to identify the minimal number of contacts in order to be effective. Similarly, in a recent review of STI counseling, it was clear that the interventions with the most hours of contact were most consistently effective; however, some lower-intensity interventions in well-conducted trials were also effective, making it very difficult to identify a minimum threshold.

**Comparison Groups**

Commonly, details about the control group or comparator intervention are not reported in behavioral counseling trials. Ideally, control group descriptions will parallel intervention descriptions, including details around counseling messages/components, intensity, frequency, and duration, if applicable. The distinctions between “no care” or “usual care” versus “minimal” intervention versus an “active” intervention need to be clearly defined. Control arms are often described as “usual care,” which, depending on the years, location, and setting performed, can be anything from “no care” to an “active” intervention. As with other research characteristics, this issue of adequate reporting around the control group is not a unique challenge to the behavioral counseling literature. However, unlike being referred for a screening colonoscopy or receiving an immunization, patients are not reliant on the recommendation or guidance of a primary care clinician to initiate healthy behavior changes, such that the usual care group frequently engages in behavior change(s) as well.

Understanding what are “like” interventions allows for more intelligent synthesis and comparison of results. For example, trials in which usual care actually entails an active intervention may show less benefit (smaller effect sizes) than trials using no or minimal intervention. In a review of healthy lifestyle interventions for the USPSTF, trials done more recently (in 2010s) showed a smaller effect than older trials (from the 1990s). Some of this reduced effect may be due to evolution in the standards of care over time (i.e., some counseling versus no counseling). In addition to advances in standards of usual care, the evolution of public and community awareness, incentives, and education around healthy...
behaviors over time means that “usual care” is highly dependent on temporal and geographic setting. Robust reporting around control groups is essential particularly across different healthcare systems. For example, in trials conducted outside the U.S., the healthcare systems are often substantially different, and “usual care” may differ markedly from any U.S. standard.

Outcomes

Ultimately, the desired benefit of behavior change is improvement in health outcomes, that is, those health outcomes that can be experienced by an individual such as better quality of life, less disease, or greater functional ability. For the USPSTF, recommendations are issued based on the magnitude and certainty of net benefit on health outcomes, as opposed to surrogate or intermediate outcomes.13 For example, increased physical activity (a behavioral outcome) may lead to decreased blood pressure (an intermediate outcome), which ultimately would be expected to result in decreased heart attacks or stroke or death (health outcomes). However, the web of behavioral, intermediate, and health outcomes and the relationship among them can be very complex and time-dependent (Figure 2).

Ideally, primary trials need to define primary and secondary outcomes of interest a priori. Too often, trials report multiple outcomes, and therefore may be prone to selective reporting bias. Even within a specific body of literature, similar outcomes can be defined and measured using a variety of methods, which may not be comparable. If the outcomes being measured and reported are not health outcomes, it is important to understand the evidence supporting the causal links of behavioral and intermediate outcomes to health outcomes. For example, identifying the evidence base to support an increase in physical activity or a decrease in blood pressure change leading to decreased cardiovascular morbidity and mortality is critical to establishing linkages. The necessity in establishing the “link” between intermediate and health outcomes is not unique to behavioral counseling studies. However, it can be compounded, as counseling trials can be costly and resource intensive if powered to detect benefits in health outcomes, especially if intended as primary or secondary prevention (versus treatment).

Measuring Behaviors

One of the greatest challenges in understanding the clinical relevance of behavioral outcomes is the variability in how behaviors are defined and measured. For example, physical activity can be defined as minutes per week, METs per week, or moderate or vigorous activity minutes per week and can be reported as a change in number of minutes per week (continuous) or change in number of people meeting recommended number of minutes per week (dichotomous). Interpreting the magnitude of findings for continuous measures, especially more-modest changes (e.g., increase in 20 minutes per week of physical activity or 2-mmHg reduction in systolic blood pressure), and their relationship to health outcomes is very difficult, and not established for many behavioral or physiologic measures. Dichotomous outcomes, although more clinically intuitive, are not commonly or consistently reported (e.g., reporting how many participants met physical activity goals or had a blood pressure less than 140/90 mmHg). Meaningful dichotomous outcome thresholds are often not established or agreed upon, leading to the inconsistency of how and if these outcomes are reported. For example, abstinence has been the most important treatment goal for alcohol, tobacco, or illicit drug use, but there are also health benefits to reduction in use, not just abstinence, which

---

Figure 2. An example of the complex relationships between behavioral counseling interventions and outcomes.  
Note: This figure demonstrates the complex relationships between behavioral outcomes, intermediate outcomes, and health outcomes using the example of diet and physical activity interventions.
should be considered. Whether continuous or categorical, standardization both of operational definitions and types of reported outcomes is critical for comparing results across studies.

From the measurement perspective, behavioral outcomes can be self-reported (subjective) or objectively measured. One is not inherently better than the other, but each must be viewed in context. Although subjective measures of behaviors (e.g., exercise log) cannot be “blinded” and are subject to both recall and social desirability bias, objective measures of behaviors (e.g., pedometer, VO₂ max) may also be limited and may not be more clearly associated with health outcomes. Both subjective and objective measures of behavior changes should be validated in the population being studied. Often, a combination of complementary outcomes measures is helpful; for example, concordant changes (in magnitude and direction of effect) in multiple measures of a single behavior or measures of multiple behaviors is a more robust finding in favor of a counseling intervention than discordant changes across different outcomes. Consistent findings of beneficial behavioral outcomes and positive change in intermediate outcomes (i.e., physiologic measures), regardless of how the outcome is measured, lead to a more robust evidence base.

Assessing Harms

Harms are an underappreciated and under-reported outcome; again, the phenomenon is not limited to behavioral counseling literature. When determining “net” benefit, the USPSTF assesses the magnitude of the health benefit in relation to the harms of the intervention. Although the harms of other preventive services have been well delineated (e.g., overdiagnosis, complications from testing or subsequent diagnostic testing, adverse drug events), the actual and potential harms of behavioral counseling interventions have been less well described. Counseling interventions have few hypothesized harms, among them anxiety and depression; injuries (e.g., from physical activity); economic costs; and the opportunity cost of delivering the intervention. Nonetheless, explicitly assessing for actual harms and clearly describing potential harms (e.g., stigma, paradoxical changes in behavioral or intermediate outcomes) are important to accurately evaluating the evidence. This clarity will help assess if the paucity of harms reported for behavioral counseling interventions is due to the lack of harms or lack of evidence on harms.

Timing

The duration of the intervention and its components (e.g., active intervention, maintenance or booster sessions) and the timing of the measurement of outcomes are necessary to understand the duration of benefit and whether the benefit is limited to the period around the counseling or is sustained beyond the end of the intervention. Often, behavioral counseling trials have a limited duration of follow-up such that little is known about the sustained benefit over time. Understanding the duration of benefit also affects the understanding of linkages between behavioral and health outcomes, such that if changes in behaviors are transient (and not sustained) there may not be downstream health benefits. Likewise, depending on the timing and duration of the studies, health outcomes may be limited to immediate outcomes (e.g., ultraviolet radiation [UV] protection counseling resulting in less sunburn) versus delayed outcomes (i.e., UV protection counseling decreasing risk of skin cancer). Adherence to the intervention over time is also important to describe, as real-world adherence is likely to be lower, and generally decreases over time. Last, because many counseling interventions are typically longitudinal in nature (e.g., requiring multiple contacts over many weeks, months, or even years), and behavior changes (e.g., initiating positive changes or terminating harmful behaviors) may occur over a long time horizon, longer-term follow-up (with or without maintenance counseling) may be needed to observe meaningful health benefits.

In summary, synthesizing evidence across a body of literature for any given behavioral intervention or behavioral outcome is immensely challenging, largely because of the complexity and heterogeneity across multiple dimensions (i.e., studied populations, evaluated interventions and comparators, measured outcomes, the relationship of behavioral and intermediate outcomes to health outcomes, and timing and persistence of benefits). Understanding and communicating these challenges will allow for advances in both primary (trials and observational studies) and secondary (meta-analyses and systematic reviews) research for behavioral counseling fields. These challenges will require more attention and consistency on the part of researchers to clearly and accurately report details around the study design and findings; encourage researchers (and funders) to establish reproducibility of findings; and will require more-thoughtful syntheses across individual studies to understand multicomponent interventions with multiple outcomes.

Future Considerations

The primary question for participants at the USPSTF Expert Forum and the major focus for this paper is whether the standards of evidence for the net benefits of behavioral counseling intervention should be the same as those for other preventive services such as screening or the use of
preventive medications. Currently, the primary standard of evidence considered by the Task Force is evidentiary support for each causal link in an analytic framework either through direct or indirect evidence. From this perspective, behavioral outcomes are acceptable if there is sufficient evidence from either clinical trials or longitudinal observational studies that behavior change results in improved health outcomes. As discussed in the panel and summarized thus far in this paper, obtaining evidence for health outcomes is challenging from both methodologic and conceptual perspectives.

The path forward for addressing methodologic challenges is relatively clear. This paper is not the first to call for the development of a common set of behavioral outcome measures and measures of harm to be included in behavioral counseling trials, which would ease the use of current standards in evaluating the evidence base. Other known and applicable advances that the research community can take are clearer descriptions of interventions for easier aggregation of studies for meta-analysis, consideration of the representativeness of the primary care population of interest in the development of eligibility criteria, and inclusion of at least intermediate outcomes in studies to help connect behavioral outcomes to health indicators. Finally, the use of existing large-scale longitudinal cohort studies to more carefully examine linkages from behavior change to health outcomes over time can provide support for accepting behavioral outcomes as a reasonable standard of evidence.

With regard to conceptual challenges, the Task Force can evaluate current methodologic standards against three key issues. First, the Task Force is interested in primary prevention, that is, interventions that can be applied to asymptomatic individuals who do not have the condition targeted for prevention. If we are able to achieve behavioral change in these populations, it can bode well for maintaining low risk status over time and disrupt the behavior–disease relationship. Such an effect would be evidenced by continued low rates of morbidity and mortality. Conceptually, it is a challenge to demonstrate improvements in health or reductions in negative health outcomes when people are starting and staying healthy over time. Second, comparative effectiveness studies, which are not currently considered in evidence reviews, could have a useful place in the development of behavioral counseling recommendations to help define more specifically the active ingredients of effective interventions. Third, behavior change is complex, and behavior change interventions can occur at multiple levels in multiple contexts. Primary care interventions that lead to successful behavior change may be necessary, but not sufficient, to improve population health. Assuming some synergy among levels and types of interventions, it is important to consider the unique contributions of behavioral counseling interventions despite concerns about time to observed health effects.

Forum participants identified the potential advantages of convening a consensus body to help guide the development of evidence standards for behavioral counseling interventions: a process in which the USPSTF can play an important role. Such a consensus body (e.g., an IOM committee) could provide standards that include

1. common language for describing the components of behavioral interventions;
2. a parsimonious set of common behavioral outcome measures for inclusion in behavioral counseling trials;
3. a common set of potential harms to assess in behavioral counseling intervention trials; and
4. a catalogue of existing longitudinal cohort studies that could provide insights to linkages between behavior change and health outcomes over time.

In addition, a consensus body could articulate a framework for developing evidence-based behavioral counseling interventions that takes into consideration

1. the methodologic rigor of the current Task Force methodology;
2. the challenge of demonstrating health outcome benefits both with regard to the goal of primary prevention and the time frame of an RCT given the natural history of chronic diseases with behavioral risk factors;
3. the types and quality characteristics of comparative effectiveness studies that could be considered in the development of recommendations; and
4. the importance of primary care– or health care–based behavioral counseling interventions in the context of multilevel interventions.

Publication of this article was supported by the Agency for Healthcare Research and Quality (AHRQ).

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 or the USPSTF. No statement in this report should be construed as an official position of AHRQ or the U.S. Department of Health and Human Services.

Administrative and logistical support for this paper was provided by AHRQ through contract HHSA290-2010-00004i, TO 4.

No financial disclosures were reported by the authors of this paper.
References


