Health behaviors are an important determinant of many chronic diseases. Current knowledge suggests that behavioral patterns contribute more to premature death than genetic predisposition, social circumstances, environmental exposures, and health care errors (1, 2). Behavioral counseling interventions are preventive services designed to help persons engage in healthy behaviors and limit unhealthy ones. Integration of behavioral counseling interventions with primary care delivery increases the reach of effective prevention strategies (3), which is particularly important with recent health care reform efforts. Advances in implementation science, which focuses on adoption and implementation of evidence-based interventions, further enhance our ability to move along the continuum from research to practice recommendations to practice changes (4).

Despite advances in the development of evidence-based behavioral counseling interventions, challenges remain in aggregating the research into recommendations for practice. Many thoughtful and constructive analyses focus on the importance of several factors in the design and conduct of studies (5–7). These factors relate to study populations, practice settings, and intervention protocols and components (8–10). These recommendations emphasize the importance of the external validity of the research base, best summarized by the adage, “If you want to get research into practice, you must get practice into research” (8, 10).

The U.S. Preventive Services Task Force (USPSTF) acknowledged the potential of modifying health behaviors for prevention as early as 1996 (11) and has previously described its methods for evaluating behavioral counseling interventions in primary care and for estimating the certainty and magnitude of the net benefit of these preventive services (12, 13). Briefly, evaluation of these interventions focuses on 2 primary questions: Do interventions in the clinical setting influence persons to change their behavior, and does changing health behavior improve health outcomes with minimal harms? (Figure 1) The systematic evidence reviews done for the USPSTF seek intervention studies that can answer whether interventions lead to changes in health outcomes and whether they can be implemented in primary care to effect behavior change.

As the USPSTF has developed recommendations for behavioral interventions in the past several years, the systematic reviews underpinning those recommendations have reinforced important challenges to making more robust recommendations. The goal of this commentary is to encourage behavioral scientists, primary care clinicians, and funders to conduct research that optimizes the ability to aggregate evidence on behavioral counseling interventions for primary care–based practice recommendations. Examples draw from 2 recent USPSTF evidence reviews, “Behavioral Counseling to Promote Physical Activity and a Healthful Diet to Prevent Cardiovascular Disease in Adults” (14) and “Screening, Behavioral Counseling, and Referral in Primary Care to Reduce Alcohol Misuse” (15).
risk status are often not clearly described in behavioral counseling studies. For example, the USPSTF alcohol misuse evidence review focused on interventions for non-alcohol-dependent patients with risky or harmful patterns of alcohol use. One challenge of identifying appropriate studies was that some constructed inclusion and exclusion criteria to limit the number of potential participants with alcohol dependence, whereas others did not collect or report data on whether any participants met criteria for alcohol dependence (15). The evidence review also noted that behavioral counseling trials used varying terminology to describe the included populations, further complicating considerations of population comparability (15).

Failure to report the proportions of study participants with and without elevated risk factors or relevant diagnoses creates additional challenges. For example, the USPSTF evidence review on behavioral counseling to promote physical activity and a healthful diet to prevent cardiovascular disease included studies in which trial participants were unselected adults (not screened for cardiovascular disease risk factors) or were screened and did not meet criteria for hypertension, hyperlipidemia, diabetes, or cardiovascular disease. Studies with mixed populations were included if no more than 50% of trial participants had relevant diagnoses. As a result, many studies were included that enrolled and analyzed mixed groups of patients. Because researchers tend to test more intensive interventions in high-risk patients, disentangling the effects of intervention intensity from the populations’ risk status was difficult (14).

### Intervention Protocols

#### Feasibility of Practice or Referral

The USPSTF behavioral counseling recommendations focus on interventions that are feasible for primary care clinicians or office staff to deliver or that are available by referral from primary care and delivered in other clinical or community settings (12).

Interventions that afford flexibility in delivery and require minimal specialized training are more likely to be adaptable for delivery by primary care teams. These behavioral interventions typically comprise brief motivational messages with supportive materials offered during routine health care visits. Referred interventions are more intensive and are delivered outside of the primary care setting by persons with more extensive training and expertise. Key to judging feasibility of referral is the ease of enrollment in the intervention, perhaps through such mechanisms as direct enrollment from the provider’s office by telephone or e-mail (16, 17), and lack of enrollment barriers (such as high cost or limited geographic availability). Interventions provided by professionals in settings commonly available in most clinical or community settings are also referable. The effectiveness of some community interventions, including complementary public health approaches through regulation or community-wide campaigns, is evaluated by the Community Preventive Services Task Force (18).

#### Components and Intensity of Interventions

The USPSTF evidence reviews aim to specify the components and intensity required to achieve reductions in behavioral risk factors and associated conditions. Because most behavioral interventions include multiple components, it is often not feasible to estimate the effects of individual components. Studies that evaluate multicomponent interventions often do not adequately describe the main components or use differing terminology for similar approaches, making it difficult to assess the comparability of interventions. For example, the evidence review on diet and physical activity did not address the effectiveness of specific intervention components, which would probably be helped by more thorough and consistent reporting of counseling intervention elements (14).

Intervention intensity is typically classified according to the duration and number of contacts, but there is no single consistent standard. Definitions of intensity are usually derived from the available literature and defined in an ad hoc manner for each evidence review, which limits comparability across recommendations. For example, interventions to reduce high-risk alcohol use were classified as very brief (≤5 minutes), brief (>5 to ≤15 minutes), extended (>15 minutes), brief multicontact (each contact ≤15 minutes), and extended multicontact (some contacts >15 minutes). Diet and physical activity interventions were categorized as low (≤30 minutes), medium (>30 minutes to <6 hours), or high intensity (>6 hours).
Outcomes and Assessment
Assessing Potential Adverse Effects

Adverse effects from behavioral interventions could derive from labeling; discrimination associated with screening; or the effects of the interventions themselves, such as increases in other unhealthy behaviors (for example, from alcohol misuse to illicit substance use or from weight control to disordered eating) or injuries (such as those from increased physical activity). Sufficient information about adverse effects is rarely reported for all types of preventive interventions, and the USPSTF generally considers the potential adverse effects of behavioral counseling interventions to be no greater than small (19, 20). A recent meta-analysis of physical activity trials reported an increase in risk factors (such as blood pressure or cholesterol level) for as much as 10% of study participants (21). The availability of consensus-derived measures of potential harms that can be routinely included in data collection protocols would greatly facilitate the uniform reporting of adverse effects of behavioral counseling interventions. As noted in the evidence review on alcohol misuse, potential adverse effects of screening and behavioral counseling interventions for alcohol misuse have received little attention in published studies. Investigators found no studies reporting on illegal substance use, stigma, labeling, discrimination, or interference with the physician–patient relationship (15).

Behavioral Outcome Measures

Many behavioral outcomes are commonly uncovered in a systematic review of studies that focus on the same risk behavior. Although there have been efforts to standardize assessment measures, particularly in areas with substantial behavioral intervention research, such as tobacco use and cessation, more typically several valid measures are used for key behavioral outcomes. To complicate matters further, different assessment methods can be used to compute the same outcome measure. For example, with diet one can calculate the percentage of energy consumed from fats by using food-frequency questionnaires, brief dietary assessments, or 24-hour recalls (22). As evidence emerges about associations between dietary components and disease, different components are emphasized for different diseases, creating further complexity. For example, studies associating diet with incident cardiovascular disease often focus on saturated fat intake (23), whereas studies of diet and cancer focus on the percentage of calories from fat; grams of fiber; and, more recently, fruit and vegetable intake (24).

Equally complex are measures of physical activity, which include self-reported surveys; job classification; motion sensors; and physiologic markers, such as doubly labeled water (25). Self-reported measures using methods as varied as detailed diaries, logs, checklists of activities, or global self-reported measures of physical activity are most commonly used (25). Efforts to pool results of behavioral counseling interventions can be hampered by noncomparability of interventions and outcomes and by varying validity of measured and reported outcomes.

The lack of standardization of behavioral measures across studies often precludes meta-analysis, making it more challenging to precisely determine net benefit, assess heterogeneity of effects by meta-regression, and provide specific recommendations for behavior change targets. For example, the evidence review on diet and physical activity noted a large variation in the type of dietary outcome reported, and it was not clear whether the included outcomes were primary or secondary outcomes or whether trials se-

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? (Reproduced with permission of Elsevier from the American Journal of Preventive Medicine. 2002;22:267-84.)
that serve as examples for this article. The evidence review also reported an association between the type of outcome measure and the effect size. For example, effect sizes for physical activity measured as minutes per week were larger than those in studies using other measures (14).

**Linking Behavior Change to Clinical Biometric Markers and Health Outcomes**

The USPSTF recommendations supporting certain behavioral counseling interventions rely on evidence that changing health behavior improves health outcomes with minimal harms. Few behavioral counseling studies are designed to measure effects on health outcomes, such as death; disability; quality of life; or acute events, such as a stroke. Even the assessment of intermediate biometric risk factors, such as lipid level, blood pressure, and blood glucose level, is uncommon. In the absence of direct evidence for improvements in health outcomes, alternative indications through an indirect chain of evidence to epidemiologic and other types of studies can show that the target behavior improves health outcomes (12). These associations are often represented in the analytic framework by dotted lines between changes in health behavior and intermediate health improvements or risk factor reduction and between intermediate health improvements and reductions in morbidity or mortality (Figure 1). One of the largest challenges for the USPSTF in developing positive recommendations for behavioral counseling interventions is identifying evidence that associates behavioral outcomes (such as physical activity) with health outcomes (such as cardiovascular mortality).

This challenge is apparent in the 2 evidence reviews that serve as examples for this article. The alcohol misuse evidence review states, “The assumption underlying brief behavioral counseling interventions in primary care is that reducing overall alcohol consumption or adopting safer drinking patterns will reduce the risk for medical, social, and psychological problems” (15). The review further notes that many of the available data are from cross-sectional or cohort studies that do not provide experimental evidence. The science was better for diet and physical activity to prevent cardiovascular disease outcomes, with 16 of the 25 trials of healthful diet counseling interventions reporting 1 or more intermediate health outcomes; however, it was more limited for specific disease outcomes, with only 3 of 39 trials reporting health outcomes related to cardiovascular disease (13).

**Recommendations**

Quality behavioral counseling intervention studies exist. The 2 reviews discussed here included more than 80 studies that met inclusion criteria based on the PICOTS (population, intervention, comparator, outcomes, timing of outcomes measurement, and setting) approach (26). The challenges we outline relate to aggregating evidence to develop practice guidelines. Opportunities to optimize behavioral counseling intervention research exist for investigators as they design studies and disseminate their findings, scientists who can lead consensus efforts in key areas, and funders who can prioritize needed research. We highlight recommendations in 3 areas.

**Design for Dissemination**

The imperative for real-world relevance of behavioral counseling intervention research long predates this commentary. Several papers articulate key gaps between research and practice and offer thoughtful and constructive approaches to creating more robust practice-relevant research (5–7, 10). Green and Glasgow (8) aptly noted a relative neglect of external validity in the evidence-based practice literature. The consequences of this neglect underscore the challenges encountered in developing evidence-based practice guidelines. The consensus is that pragmatic studies designed and implemented with representative populations in actual practice settings have the most value for rapid translation and may best inform practice guidelines (10).

No researcher begins a study with the desire to minimize its contribution, yet decisions made during the design and implementation of a study have important implications for translation of its findings. Models and resources are available for researchers who are committed to optimizing the translation potential of their research. Glasgow provides a cogent summary of 3 models: the RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance), EIT (Evidence Integration Triangle), and PRECIS (Pragmatic-Explanatory Continuum Indicator Summary) models (10).
In this commentary, we focus on the PRECIS model because it is particularly relevant to behavioral interventions that are integrated into the primary care setting (27). This model recognizes several aspects of study design that relate to challenges encountered by the USPSTF about study populations, intervention components and intensity, and feasibility or referability of intervention protocols.

The PRECIS tool outlines 10 domains of study design and identifies criteria to determine where each design feature lies on a continuum from pragmatic to explanatory. These domains relate to participant eligibility, intervention flexibility, comparison interventions, practitioner expertise, follow-up intensity, primary trial outcomes, participant and practitioner adherence to study protocols, and analysis of primary outcome. Research teams can map the characteristics of planned study features by using the wheel graphic illustrated in Figure 2. The PRECIS tool employs the terms “pragmatic” and “explanatory” to describe a continuum of study design choices on each spoke of the wheel. In Figure 2, pragmatic design features are toward the rim of the wheel and explanatory features are toward the hub (hence the “E” in the center of the wheel). For example, for practitioner adherence to the intervention protocol, pragmatic trials allow practitioners to customize their approach on the basis of their practice setting, whereas explanatory trials would more closely monitor practitioner adherence and take corrective steps with those who do not adhere to the protocol. Using the PRECIS tool during the study design phase ensures that each detail of the study design receives careful attention from the outset and can facilitate later reporting of study features. The PRECIS tool has been used to evaluate behavioral counseling interventions for smoking cessation (28) and weight loss (29), among others.

Further opportunities occur in the reporting of study findings. Researchers’ consistent use of reporting guidelines, including the CONSORT (Consolidated Standards of Reporting Trials) statement and its extensions and the TREND (Transparent Reporting of Evaluations with Nonrandomized Designs) statement checklist, and enforcement of such reporting by journal editors can alleviate challenges related to the lack of clarity about the risk status of study populations and understanding the components and intensity of interventions (30-32). Green and Glasgow (8) propose several quality rating criteria to assess external validity that could be applied during peer review of proposed research. These criteria address reach and representativeness, program or policy implementation and adaptation, outcomes for decision making, and maintenance and institutionalization and could be used to augment guidelines, such as CONSORT and TREND.

**Table.** Possible Definitions of Intensity for Primary Care–Feasible and Primary Care–Referable Interventions

<table>
<thead>
<tr>
<th>Intensity</th>
<th>Primary Care–Feasible</th>
<th>Primary Care–Referable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low</td>
<td>0–15 min (approximately 1 very brief contact, with little time for anything else from education)</td>
<td>≤30 min (approximately 1 brief session)</td>
</tr>
<tr>
<td>Low</td>
<td>16–30 min (approximately 1 brief contact, with time for some discussion of factors, such as motivation, behavioral skills, barriers, or self-efficacy)</td>
<td>&gt;30 min–2 h (approximately one to four 30-min sessions)</td>
</tr>
<tr>
<td>Medium</td>
<td>&gt;30 min–2 h (approximately one 30-min session plus one to six 15-min follow-up contacts)</td>
<td>3–25 contacts (approximately weekly for &lt;6 mo)</td>
</tr>
<tr>
<td>High</td>
<td>&gt;2–6 h (approximately two 30-min sessions plus five to twenty 15-min follow-up contacts)</td>
<td>26–52 contacts (approximately weekly for 6–12 mo)</td>
</tr>
<tr>
<td>Very high</td>
<td>&gt;6 h</td>
<td>&gt;53 contacts (approximately &gt;1 h/wk for 12 mo)</td>
</tr>
</tbody>
</table>

**Develop Consensus Terminology and Measures**

A more consistent terminology for the intensity of behavioral counseling interventions can facilitate cross-study comparisons by meta-analysis and can also allow for more detailed assessments of dose–response effects. Such a terminology could also inform the extent to which a counseling intervention is feasible for primary care delivery or referral. The Table suggests consistent, descriptive language for intervention intensity that is scaled to reflect different categories of interventions. Further refinement of these categories could facilitate more general conclusions about the overall effectiveness of behavioral interventions of different intensity. Use of these metrics, along with better reporting of data on the number, frequency, and duration of intervention contacts and the time invested per contact, would further assist reviewers of behavioral counseling interventions.

The development, dissemination, and use of consensus measures are challenging. The National Institutes of Health and professional organizations can bring together leading researchers in the field to evaluate and recommend robust measures, and there are examples of such efforts. In 1993, the National Center for Health Statistics sponsored a Consensus Workshop on Dietary Assessment that recommended methods for assessing dietary intake at the population level, as well as specific individual measures of intake, including alcohol use (33). In the mid-2000s, a workgroup convened by the Society for Research on Nicotine and Tobacco published 5 recommendations for consensus measures of abstinence from tobacco in clinical studies (34). Some federal funding organizations provide guidance on several constructs, including behavioral mea-

---

**Figure 2.** The PRECIS tool employs the terms “pragmatic” and “explanatory” to describe a continuum of study design choices on each spoke of the wheel.
sures and measures of health status and other health outcomes. For example, the Cancer Control and Population Sciences division of the National Cancer Institute provides research measures and databases as part of a Web-based resource for behavioral research (35, 36). Although these resources can be helpful, consensus panels do not necessarily evaluate possible common measures from the perspective of aggregating clinical trial data for evidence-based practice recommendations. For example, the National Center for Health Statistics recommends the use of multiple 24-hour recalls for dietary measures, which is an intensive measurement that may not be practical for primary care–based intervention studies. Moreover, there is often little incentive for investigators to include common measures in their research.

A collaborative effort among funding agencies, behavioral science professional organizations, practice-based research networks, and the USPSTF could lead to consensus recommendations for a parsimonious set of core measures related to behavioral outcomes and meaningful clinical indicators of health effects. Once such measures are identified, funding agencies could require their inclusion as a condition of funding. Collaboration with initiatives in electronic health records to include assessment of health behaviors could also contribute robust behavioral outcome data that could be linked with clinical outcomes.

**Mine Existing Data for Linkages Between Behavior Change and Clinical Outcomes**

The problem of publication bias, in which studies with negative results are not published (35), may have a corollary problem related to failure to analyze and report results for all of the data collected during a clinical trial. Behavioral counseling studies often include secondary measures that could help connect the dots between changes in health behavior and intermediate health improvements or risk factor reduction. For example, the evidence review on alcohol misuse searched for reduced incidence of medical, social, and psychological problems that could be assessed by patient follow-up surveys. Other relevant data that can be obtained from self-report include employment stability, health care utilization, and legal issues. Although these data were lacking in the published literature, they may be available in existing data sets. Pooling data across trials for these types of outcomes may provide data with adequate power to explore health outcome effects. Opportunities may also exist to conduct long-term follow-up for disease incidence by using electronic medical records data for large cohorts of participants in behavioral counseling studies. Further, national longitudinal studies that include biometric assessments, such as blood pressure, blood glucose level, and lipid level, may be an untapped resource for secondary analyses to better inform associations between behavior change and health outcomes. Support from funding agencies for these types of secondary analyses could greatly enhance the work of the USPSTF.

**CONCLUSION**

Behavioral counseling interventions are important primary and secondary preventive care strategies, and the USPSTF is committed to developing and disseminating recommendations to ensure that effective interventions achieve the broadest reach into health care delivery. Effective synthesis and incorporation of evidence for behavioral counseling interventions into USPSTF recommendations is challenged by gaps in the current evidence base. The behavioral science community can better align primary care–based intervention studies with the key questions that guide the development of evidence-based recommendations. This commentary joins other voices calling for careful attention to study populations and the pragmatic aspects of intervention protocols in the design and dissemination of research, for greater consistency in key behavioral measures, and for further research that links behavior change to health outcomes.

From University of Iowa, Iowa City, Iowa; Group Health Research Institute, Seattle, Washington; Kaiser Permanente Research Affiliates Evidence-based Practice Center, Portland, Oregon; and University of Texas Health Sciences Center at San Antonio, San Antonio, Texas.

**Financial Support:** The USPSTF is an independent, voluntary body. The U.S. Congress mandates that the Agency for Healthcare Research and Quality support the operations of the USPSTF.

**Potential Conflicts of Interest:** Disclosures can be viewed at [www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum =M13-2128](http://www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum =M13-2128).

**Requests for Single Reprints:** Susan J. Curry, PhD, College of Public Health, University of Iowa, 105 River Street, S153A CPHB, Iowa City, IA 52242; e-mail, Sue-curry@uiowa.edu.

Current author addresses and author contributions are available at [www.annals.org](http://www.annals.org).

**References**

Current Author Addresses: Dr. Curry: College of Public Health, University of Iowa, 105 River Street, S153A CPHB, Iowa City, IA 52242. Dr. Grossman: Group Health Cooperative, Group Health Research Institute, Metropolitan Park East Building, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101. Dr. Whitlock: Kaiser Permanente Research Affiliates Evidence-based Practice Center, Center for Health Research, Kaiser Permanente Northwest, 3800 North Interstate Avenue, Portland, OR 97227. Dr. Cantu: University of Texas Health Science Center at San Antonio, 7730 Floyd Curl Drive, San Antonio, TX 78228.