Identifying promising practices for evaluation: the National Breast and Cervical Cancer Early Detection Program

Amy DeGroff,
Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, 4770 Buford Hwy, MS K-57, Atlanta, GA 30314, USA

Karen Cheung,
Public Health and Survey Research, ICF International, Atlanta, GA, USA

Nicola Dawkins-Lyn,
Public Health and Survey Research, ICF International, Atlanta, GA, USA

Mary Ann Hall,
Public Health and Survey Research, ICF International, Atlanta, GA, USA

Stephanie Melillo, and
Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, 4770 Buford Hwy, MS K-57, Atlanta, GA 30314, USA

Rebecca Glover-Kudon
Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, 4770 Buford Hwy, MS K-57, Atlanta, GA 30314, USA

Abstract

Purpose—The Centers for Disease Control and Prevention conducted a systematic screening and assessment process to identify promising practices implemented by grantees of the National Breast and Cervical Cancer Early Detection Program and its partners that were appropriate for rigorous evaluation.

Methods—The systematic screening and assessment (SSA) process was conducted from September 2010 through March 2012 and included five steps: (1) nominations of promising practices; (2) a first rating by subject matter experts; (3) field-based evaluability assessments; (4) a second rating by experts; and (5) use of results. Nominations were sought in three program areas including health education and promotion, quality assurance and quality improvement, and case management/patient navigation.

Results—A total of 98 practices were nominated of which 54 % were eligible for the first review by the experts. Fifteen practices were selected for evaluability assessment with ten forwarded for the second review by the experts. Three practices were ultimately recommended for rigorous evaluation, and one evaluation was conducted. Most nominated practices were based on evidence-based strategies rather than representing new, innovative activities. Issues were identified through the process including inconsistent implementation and lack of implementation fidelity.

adegroff@cdc.gov.

Conclusion—While the SSA was successful in identifying several programs for evaluation, the process also revealed important shortcomings in program implementation. Training and technical assistance could help address these issues and support improved programming.

Introduction

The Breast and Cervical Cancer Mortality Prevention Act of 1990 [1] established the National Breast and Cervical Cancer Early Detection Program (NBCCEDP) to be administered by the Centers for Disease Control and Prevention (CDC). Through the NBCCEDP, low-income, un-insured, and underinsured women receive important cancer screening services. To date, the NBCCEDP has served over 4.5 million women and diagnosed 62,121 breast cancers, 3,458 invasive cervical cancers, and 163,548 premalignant cervical lesions [2]. Between 2013 and 2014, the NBCCEDP provided nearly $154 million to all 50 states, the District of Columbia, five US territories, and 12 American Indian/Alaska Native tribal organizations. As the only nationally organized screening program in the USA, the NBCCEDP pays for screening and diagnostic services as well as partnership development, public education and recruitment, professional education, quality assurance and quality improvement, patient navigation, data management, and evaluation. The law currently mandates that at least 60 % of NBCCEDP funds be used to support direct clinical services, with up to 40 % allocated for the activities noted above.

The Affordable Care Act [3] is increasing the number of persons who have access to health insurance and, subsequently, to breast and cervical cancer screening. For most private insurance plans and newly eligible Medicaid expansion beneficiaries, the law requires coverage, with no cost-sharing, for clinical preventive services recommended by the United States Preventive Services Task Force (USPSTF), including screening for breast and cervical cancer [4]. By removing cost as a barrier, the Affordable Care Act offers new opportunities for the NBCCEDP to increase access to screening for a broader population than has traditionally been served by the program. In particular, the NBCCEDP could leverage its non-screening components to help integrate public health and primary care, reducing an existing community clinical services gap and facilitating health care access for a broader population than traditionally reached [5]. A possible model for the NBCCEDP is CDC's Colorectal Cancer Control Program (CRCCP), a program encouraging population-level impact through use of evidence-based strategies focused on health systems and organizational policy change [6].

In anticipation of healthcare reform and full implementation of the Affordable Care Act, CDC funded several projects to explore future directions for the NBCCEDP [7]. Recognizing the possibility for scaling up non-screening activities, one project sought to identify promising practices implemented by NBCCEDP grantees or partners that have potential to achieve population-level effects and that are well positioned for rigorous

evaluation. The project aligns with an important public health goal to identify and promote evidence-based practice [8].

Evidence-based approaches are often identified through efficacy studies involving experimental or quasi-experimental designs whereby key variables are tightly controlled in order to isolate intervention effects. Currently, CDC largely relies on the Guide to Community Preventive Services [9] to identify evidence-based strategies. However, sole reliance on interventions identified through efficacy studies is limiting. In particular, evidence-based interventions are expensive to implement and often overly complex for real-world replication by community-based public health practitioners facing many constraints [10-12]. Other methods are needed to gather *evidence* from public health programs, including the NBCCEDP. With over 20 years of implementation experience, NBCCEDP grantees and their partners offered an opportunity to identify innovative public health practices (e.g., interventions, strategies, policies, procedures, processes, and/or activities) ripe for evaluation.

The systematic screening and assessment method (SSA), developed by the Robert Wood Johnson Foundation in collaboration with CDC, sequences expert peer review with evaluability assessment as a means to identify promising public health practices worthy of rigorous evaluation. The SSA includes criteria such as reach to the target population, generalizability, data availability, and potential impact [13]. Both expert review [14] and evaluability assessment [15] are recognized as systematic evaluation approaches. By examining program design, implementation, and data availability, evaluability assessments —which can be considered pre-evaluations—help ascertain whether further evaluation is justified, feasible, and likely to provide useful information. Developers initially applied the SSA to obesity prevention [16], and it has since been used in food policy, active transportation, food service guidelines, and joint use agreements. Given limited evaluation resources, the SSA is a cost-effective approach to identify the most promising public health practices best positioned for evaluation [13]. At the same time, valuable feedback and technical assistance can be delivered even to those programs that would not benefit from further evaluation.

To identify promising public health practices implemented by NBCCEDP grantees or partners, CDC initiated an SSA in 2010. In doing so, CDC hoped to identify non-screening practices with potential for population impact, improve its understanding of current public health practices among NBCCEDP grantees, and bridge the divide between promising and evidence-based practice. The purpose of this paper is to present results from the SSA. An evaluation team led by staff from ICF International with CDC participation carried out the SSA.

Materials and methods

The SSA was conducted from September 2010 through March 2012. Figure 1 represents the five-step SSA process [13] that reduces a large number of candidate public health practices to a priority listing of those most promising and appropriate for rigorous evaluation. Each step in the process is described below.

Step 1—Nomination of practices: The evaluation team selected three priority areas for assessment including (1) health education and promotion (HEP), (2) quality assurance and quality improvement (QA/QI), and (3) case management or patient navigation (CM/PN) that are defined in Table 1. These areas were selected for their potential to achieve greater population-level effects in the new healthcare environment. On December 1, 2010, CDC held a webinar with grantees of the NBCCEDP and the National Comprehensive Cancer Control Program (NCCCP) as well as other partners to launch the SSA project. An electronic nomination form was circulated broadly to grantees and partners through various listservs (e.g., NBCCEDP, NCCCP, American Cancer Society, Susan G. Komen Foundation). The nomination form collected voluntary information on the practice name, type (e.g., HEP), and description (priority population, activities, outcomes of interest, data collection efforts). For purposes of the project, "practice" was defined as interventions, strategies, policies, procedures, processes, and/or activities implemented in the field with the intention to affect a change or achieve a particular impact. Examples of impact may be to decrease morbidity, mortality, or disability, or to increase quality of life or effectiveness and efficiency of public health. Nominations were collected through 21 January 2011.

The evaluation team reviewed all nominated practices and contacted programs for additional information. Inclusion and exclusion criteria were applied in order to select practices for review by experts. For inclusion, nominated practices had to address one of the three priority areas, have a plausible goal related to improving screening prevalence and adherence to diagnostic follow-up, have been currently implemented for at least 6 months, have a measurable outcome, and be suitable for implementation in other contexts. Practices undergoing a rigorous evaluation, already evaluated, or a strict replication of an existing evidence-based model (without major adaptations) were excluded using a standard template; the evaluation team prepared a summary of each practice including a project description and information addressing several areas (e.g., intended outcomes, reach, implementation, sustainability).

Step 2—First informal meeting of experts: Twenty-three researchers and practitioners representing academia, federal and state agencies, and nonprofit organizations with expertise in at least one of the three priority areas were recruited for the project. Experts were sent eight to nine practice summaries consistent with their area of expertise (e.g., HEP) to review using a rating assessment tool accessed online. To reduce potential bias and ensure candid discussion, pseudonyms were used for all nominated practices. Table 1 summarizes the nine criteria assessed for each practice. The assessment tool included up to four Likert scale questions for each criterion. A value of one was associated with low scores, such as "low proportion of population reached," and a value of four represented favorable ratings, such as "highly acceptable to stakeholders." Experts also had an option of "unable to assess" if they believed inadequate information was available to rate the criterion. A final summative item using the four-point Likert scale was included, "How strongly do you recommend this practice for evaluability assessment?" with four reflecting "strongly recommend."

Three to four experts reviewed and rated each practice. Using SPSS version 16.0, a mean score for each of the nine criteria was calculated on individual practices. In addition, all criterion scores were averaged into an overall composite mean score for the practice. The

overall composite mean score refers to the mean of the mean scores. Criteria rated as "unable to assess" were treated as missing data and not included in means calculations. Criteria were treated equally with no weighting used. Graphs were produced for each criterion to provide a visual aid for determining which practices might have the highest potential impact, the highest feasibility for adoption, and so on.

In April 2011, seventeen experts convened in Atlanta, Georgia and six experts joined by telephone to discuss the ratings and make recommendations for evaluability assessments. Experts were placed in groups representing the three priority areas to review the rating results and discuss the candidate practices. The full group then reconvened, and each priority group recommended practices for the evaluability assessments along with suggested issues for site visitors to explore during the assessment process. The evaluation team reviewed recommendations from the experts and selected practices for evaluability assessments.

Step 3—Evaluability assessments: To gather detailed information on each practice, evaluation team members reviewed background documents provided by staff. Teams of two to three then conducted two and a half day site visits between July and October 2011. The evaluation team spoke with nine to 12 stakeholders at each site for approximately 1–2 hours each to better understand the practice and its implementation—including its history, goals, and data sources. During the site visit, the evaluation team facilitated a logic model-building exercise with staff to explicate the relationships between practice inputs, activities, outputs, and outcomes [17].

Following each site visit, the evaluation team drafted reports describing each practice and including programmatic recommendations. Summary reports were provided to all participating practices. Staff members were invited to participate in a follow-up meeting by telephone to review the report, refine the logic model, and discuss recommendations. Based on results of the evaluability assessments that revealed significant implementation problems for several practices, the evaluation team selected 10 of the 15 practices for final review by the experts.

Step 4—Second informal meeting of experts: Each expert was provided detailed reports and logic models developed from the evaluability assessment and asked to review and rate the practices using the nine established criteria (Table 1), again through the web-based rating assessment tool. Ten to thirteen experts reviewed and rated each practice. Consistent with the first informal meeting, the evaluation team calculated a mean score for each of the nine criteria on each practice as well as a mean, composite score. As before, no weighting was used and criteria rated as "unable to assess" were treated as missing data. The experts were convened by telephone in January 2012 to discuss each practice and review assessment ratings. Individual experts presented a summary of each practice to the group. At the end of the discussion, experts were asked to vote for the top three practices they would recommend for rigorous evaluation.

Step 5—Use of SSA results: Following the second informal meeting, the evaluation team met to discuss results and prioritize a practice for evaluation to CDC leadership. Considerations influencing selection included the potential costs in terms of both time and

money to carry out the evaluation. The evaluation team synthesized results from the SSA and reflected upon the experience in order to extract lessons learned across similar practices.

Results

Nomination of practices: The nominations process yielded a total of 98 practices including 32 practices in HEP, 39 in QA/QI, and 27 in CM/PN. Sixteen of the nominated practices addressed more than one category. Most often, CM/PN practices included a health education and promotion component. Nominations were evenly distributed across the US: 24 nominations were received from the Northeast, 25 from the Southeast, 22 from the Midwest, 24 from the West, and three from US territories (Fig. 2).

Based on inclusion and exclusion criteria applied by the evaluation team, 45 practices were excluded. The remaining 53included20HEP, 21 QA/QI, and 12CM/PN. Practices were excluded for issues related to design (e.g., no plausible logic to increase screening rates), implementation (e.g., no longer implemented, partial implementation of a practice), evaluation (e.g., data collection not possible, already evaluated), and capacity (e.g., no response to project inquiries).

First informal meeting of experts: Based on the experts' assessment ratings, composite mean scores ranged from 2.3 to 3.6 for the 53 practices with a median of 3.19. The lowest mean scores recorded were for feasibility of adoption (2.78) and intervention sustainability (2.64); stakeholder acceptance (3.35) and feasibility of implementation (3.34) received the highest mean scores.

Following extensive discussion, the experts recommended 12 practices for evaluability assessment and identified several others than might warrant additional consideration. Experts also shared the following observations:

- Nominated practices predominately reflected adaptations of Community Guide-recommended strategies rather than new, innovative approaches,

- Selecting between untested models and those based on evidence-based strategies was difficult, and

- Strong partnerships seemed critical to practice implementation, especially for highly ranked ones in the HEP category.

Following the informal meeting, the evaluation team met to discuss the recommendations from the experts. Based on the experts' advice and CDC's interest in practices with potential for population-based effects or those addressing the screening needs of vulnerable populations, the team considered additional practices for evaluability assessment. Aside from examining ratings from the experts, the team engaged in the same assessment process using the web-based tool (results not presented). Based on that process, three other practices were added for a total of 15 identified for evaluability assessment. The practices represented the following categories: six HEP, three QA/QI, and six practices spanning multiple categories (four HEP–CM/PN, one HEP–QA/QI, one reflecting all three categories).

Evaluability assessments: Fifteen evaluability assessments were conducted, and reports summarizing assessment results for each site visit were drafted. Five practices were removed from further review given significant implementation problems that included the lack of implementation fidelity with the program design and inconsistent, non-systematic implementation. Consequently, ten practices were forwarded for final review and rating by the experts including three HEP, three QA/QI, three HEP–CM/PN, and one HEP–QA/QI.

Reflecting on the evaluability assessment process, evaluation team members identified several commonly observed issues:

- Staff were not implementing enough of a given activity to achieve the desired outcome,

- Staff were implementing too many different activities to achieve the desired outcome,

- Staff were implementing activities inconsistently or without fidelity,

- Staff did not adequately document implementation efforts, and

- Staff often requested assistance in collecting, interpreting, and presenting evaluation data.

Second informal meeting of experts: Table 2 summarizes the experts' ratings by criteria for each of the ten practices, numbered 1–10. Consistent with the first informal meeting, mean scores for individual criteria were lowest for feasibility of adoption (2.35) and intervention sustainability (2.59) and highest for feasibility of implementation (3.55) and stakeholder acceptance (3.53).

Figure 3 provides the composite and overall recommendation scores for each practice. Composite scores ranged from 2.86 (practice #4) to 3.42 (practice #1); overall recommendation scores ranged from 2.50 (practice #4) to 3.70 (practice #5). Results of voting conducted during the informal meeting of experts for the top three practices were nearly consistent with composite scores; the three practices receiving the most votes were included in the four practices receiving the highest composite scores. Of interest, three practices received no votes from experts in this final exercise.

Step 5—Use of SSA results: The evaluation team recommended the top three practices identified by the experts to decision makers in the CDC's Division of Cancer Prevention and Control and solicited their input to identify the first practice for evaluation. At the time, resources were available to support one evaluation. The Colorado NBCCEDP's Women's Wellness Connection bundled payment system, a QA/QI practice, was selected given several factors, including that the practice was innovative and appeared to be highly replicable for other NBCCEDP and cancer programs, staff were willing to provide clinical and financial data for analysis, and the evaluation could be completed within existing time and cost constraints. The team designed and conducted an evaluation study of the Colorado NBCCEDP's Women's Wellness Connection bundled payment system in 2012–2013 [18].

The team also reflected on the SSA results identifying the following lessons learned:

- Most practices were based on evidence-based strategies rather than new, innovative activities, and many encompassed a spectrum of activities from educating women (HEP), addressing barriers (CM/PN), and supporting adherence and timely care (QA/QI).

- Many nominated practices were not well planned or implemented; only three of the 98 nominated practices (3 %) were ultimately recommended for rigorous evaluation.

- Documentation of program implementation was limited, and practice staff expressed a need for assistance in collecting, analyzing, and reporting data.

Discussion

We found the SSA an effective process for identifying promising cancer screening-related practices implemented in real-world settings ready for rigorous evaluation. By applying criteria such as acceptability to stakeholders and feasibility of adoption, areas typically ignored in highly controlled intervention research, the SSA allows evaluators to assess field-based practices and contribute toward building practice-based evidence from the "bottom up," a current gap. Recently, CDC authors, drawing on the SSA as well as the reach, effectiveness, adoption, implementation, maintenance (RE-AIM) model [19] and the integrative validity model [12], proposed a new framework that recognizes a continuum of evidence-based practice, from emerging to best [20]. The framework is meant to complement traditional, systematic review approaches (e.g., Community Guide) and, like the SSA, contribute to building practice-based evidence. The framework integrates two distinct concepts, public health impact and quality of evidence, in order to assess practices. Five components comprise public health impact: effectiveness, reach, feasibility, sustainability, and transferability. Strength of evidence is assessed via a range: weak, moderate, strong, and rigorous.

While the SSA proved an effective method, the process yielded far fewer practices ready for evaluation than anticipated. Despite a seemingly successful nomination effort generating 98 practices for review, only a few were ultimately found appropriate for further evaluation. Other SSAs have had similar findings [16]. Our results may reflect the NBCCEDP's traditional emphasis on screening provision; as noted, grantees are required by law to expend at least 60 % of funds on clinical service delivery. Grantees may simply not have adequate resources to implement strong practices in the three foci for this SSA.

However, results may also reflect limited capacity in program planning, implementation, and evaluation for the non-clinical intervention activities we assessed. A 2013 survey of NBCCEDP grantees identified the need for training in implementation of specific strategies such as systems change, quality improvement, and provider assessment and feedback, as well as in program monitoring and evaluation [21]. However, through the SSA, we identified other types of shortcomings including partial and inconsistent implementation, lack of implementation fidelity, and inadequate "dose" of the practice because implementers were either not doing enough of a given activity or trying to do too many different things. This is significant given that strong empirical evidence has demonstrated that the level of

implementation affects outcomes [22]. Further, one of the lowest scored criteria for both informal meetings was feasibility of adoption, suggesting the practice would be difficult to adopt in its current form. We also found that documentation of implementation efforts was limited, and programs faced other challenges to program monitoring and evaluation.

Based on these findings, CDC program consultants and staff assigned to provide regular technical assistance to grantees, as well as the grantees, received extensive training in program logic modeling in 2013 to support improved planning, implementation, and evaluation. By developing a logic model, program planners and implementers can help ensure a plausible relationship between the proposed activities and the intended outcomes and also identify outputs and outcomes for program monitoring and evaluation. In addition, logic models benefit program planning by improving clarity about the program among program staff and stakeholders, identifying needed resources, specifying the sequencing of activities that should be implemented, and serving as a basis for program evaluation [17]. Under the most recent NBCCEDP funding announcement, grantees were required to develop a program logic model to specify relationships between program activities and intended outcomes. To help grantees meet this requirement, CDC developed related guidance and program consultants and evaluators are providing technical assistance. Other training to address evaluation planning is forthcoming, an area for which SSA participants explicitly requested assistance. As a screening program, the NBCCEDP is data-driven and supports a strong culture of data use [23, 24]; therefore, this capacity should be lever-aged as programs potentially expand non-screening efforts.

Through the SSA, we also found that grantees often implement multiple strategies to achieve individual screening objectives, as demonstrated by practices falling into more than one category (e.g., HEP and CM/PN). Such multiple component approaches, if implemented appropriately, may increase impact as synergies across activities are realized. Innovation may not rest in the practice alone, but also in how evidence-based strategies are combined to enhance impact. Effective organized screening programs involve integrating strategies such as patient outreach and recruitment, client and provider reminders, patient navigation, and provider assessment and feedback.

Limitations

While the SSA offers a systematic means to identify promising practices and identify those most appropriate for rigorous evaluation, it does have limitations. Results suggest the SSA process may be slightly improved by specifying whether replications of evidence-based strategies or whether truly innovative practices are preferred as nominations. The validity and reliability of the assessment tool used by the experts have not been examined. Additionally, the number of persons reviewing each practice wasrelatively small. The experts were comprised of highly respected researchers and practitioners in the field, and the discussions that followed the objective rating may prove equally important in determining which practices the experts select to move forward for evaluability assessment or for evaluation. In addition, experts may not have had as complete and comprehensive information about each practice as would be ideal. However, the evaluation team compiled

extensive materials and conducted thorough evaluability assessments in order to describe the practices as fully as possible.

Conclusion

Although the SSA proved an effective method, fewer practices ready for rigorous evaluation were identified than anticipated. An unexpected and important consequence of the SSA process was the identification of needed improvements in program planning, implementation, and evaluation. Training and technical assistance could address these deficiencies, supporting improved programming and better preparing programs for evaluation. In particular, program logic model training could support improvements in both program planning and process and outcome evaluation. Similarly, training addressing factors including fidelity, dose, and reach could improve program implementation and, subsequently, the likelihood of achieving related outcomes.

Acknowledgments

References

1. Breast and Cervical Cancer Mortality Prevention Act of 1990, Pub. L. No. 101-354. 1990.

2. National Breast and Cervical Cancer Early Detection Program. [Accessed 8 May 2014] 2014. http://www.cdc.gov/cancer/nbccedp/about.htm

3. Patient Protection and Affordable Care Act, Pub. L. No. 111-148. 2010.

4. USPSTF A and B Recommendations: U.S. Preventive Services Task Force. [Accessed 8 May 2014] https://www.healthcare.gov/what-are-my-preventive-care-benefits/

5. Institute of Medicine. Primary care and public health: exploring integration to improve population health. The National Academies Press; Washington: 2012.

6. Joseph D, DeGroff A, Hayes N, Wong F, Plescia M. The Colorectal Cancer Control Program: partnering to increase population level screening. Gastrointest Endosc. 2011; 73(3):429–434. [PubMed: 21353839]

7. Plescia M, Wong F, Pieters J, Joseph D. The National Breast and Cervical Cancer Early Detection Program in the era of health reform: a vision forward. Cancer. 2014; 120(16):2620–2624. [PubMed: 25099907]

8. Frieden TR. A framework for public health action: the health impact pyramid. Am J Public Health. 2010; 100:590–595. [PubMed: 20167880]

9. The Guide to Community Preventive Services. [Accessed 8 May 2014] 2014. http://www.thecommunityguide.org/index.html

10. Green LW. Public health asks of systems science: to advance our evidence-based practice, can you help us get more practice-based evidence? Am J Public Health. 2006; 96:406–409. [PubMed: 16449580]

11. Glasgow RE, Green LW, Klesges LM, Abrams DB, Fisher EB, Goldstein MG, et al. External validity: we need to do more. Ann Behav Med. 2006; 31:105–108. [PubMed: 16542124]

12. Chen, HT.; Garbe, P. Assessing program outcomes from the bottom-up approach: an innovative perspective to outcome evaluation. In: Chen, HT.; Donaldson, SI., editors. Advancing validity in outcome evaluation: theory and practice. New directions for evaluation. Vol. 130. Sage; Thousand Oaks, CA: 2011. p. 93-106.

13. Leviton, LC.; Gutman, MA. Overview and rationale for the systematic screening and assessment method. In: Leviton, LC.; Kettel Khan, L.; Dawkins, N., editors. The systematic screening and

assessment method: finding innovations worth evaluating. New directions for evaluation. Vol. 125. Sage; Thousand Oaks, CA: 2010. p. 7-31.

14. Rossi, PH.; Lipsey, MW.; Freeman, HE. Evaluation: a systematic approach. Sage; Thousand Oaks: 2003.

15. Wholey, JS. Assessing the feasibility and likely usefulness of evaluation. In: Wholey, JS.; Hatry, HP.; Newcomer, KE., editors. Handbook of practical program evaluation. Jossey Bass; San Francisco: 2004. p. 15-39.

16. Dawkins, N.; Wethington, H.; Khan, LK.; Grunbaum, JA.; Robin, L.; Barnes, SP., et al. Applying the systematic screening and assessment method to childhood obesity prevention. In: Leviton, LC.; Kettel Khan, L.; Dawkins, N., editors. The systematic screening and assessment method: finding innovations worth evaluating. New directions for evaluation. Vol. 125. Sage; Thousand Oaks, CA: 2010. p. 33-49.

17. Fretchling, JA. Logic modeling methods in program evaluation. Wiley; San Francisco; 2007.

18. Cheung, K.; Boyce, SP.; Heisler, H.; Cash, S.; Lara, C.; Kinsella, E., et al. Evaluation of a bundled payment system in the National Breast and Cervical Cancer Early Detection Program: Women's Wellness Connection in Colorado; American Public Health Association annual meeting; Boston, MA. 2-6 Nov; 2013.

19. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. Am J Public Health. 2003; 93:1261–1267. [PubMed: 12893608]

20. Spencer LM, Schooley MW, Anderson LA, Kochtitzky CS, DeGroff AS, Devlin HM, et al. Seeking best practices: a conceptual framework for planning and improving evidence-based practices. Prev Chronic Dis. 2013 doi:10.5888/pcd10.130186.

21. DeGroff, A.; Royalty, J.; Gressard, L.; Myles, Z.; Melillo, S.; Rice, K.; Joseph, K.; Wong, F.; Miller, J.; Kammerer, W.; Helsel, W.; Howe, W.; Rockwell, W. Implementation of the National Breast and Cervical Cancer Early Detection Program; Prevent Cancer Foundation's Dialogue for Action conference; Baltimore, MD. 21 March 2014; 2014.

22. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol. 2008; 41:327–350. [PubMed: 18322790]

23. Yancy B, DeGroff A, Royalty J, Marroulis S, Mattingly C, Bernard V. Using data to effectively manage a national screening program. Cancer. 2014; 120(16):2575–2583. [PubMed: 25099900]

24. DeGroff A, Royalty J, Howe W, Buckman D, Gardner J, Poister T. When performance management works: a study of the National Breast and Cervical Cancer Early Detection Program. Cancer. 2014; 120(16):2566–2574. [PubMed: 25099899]
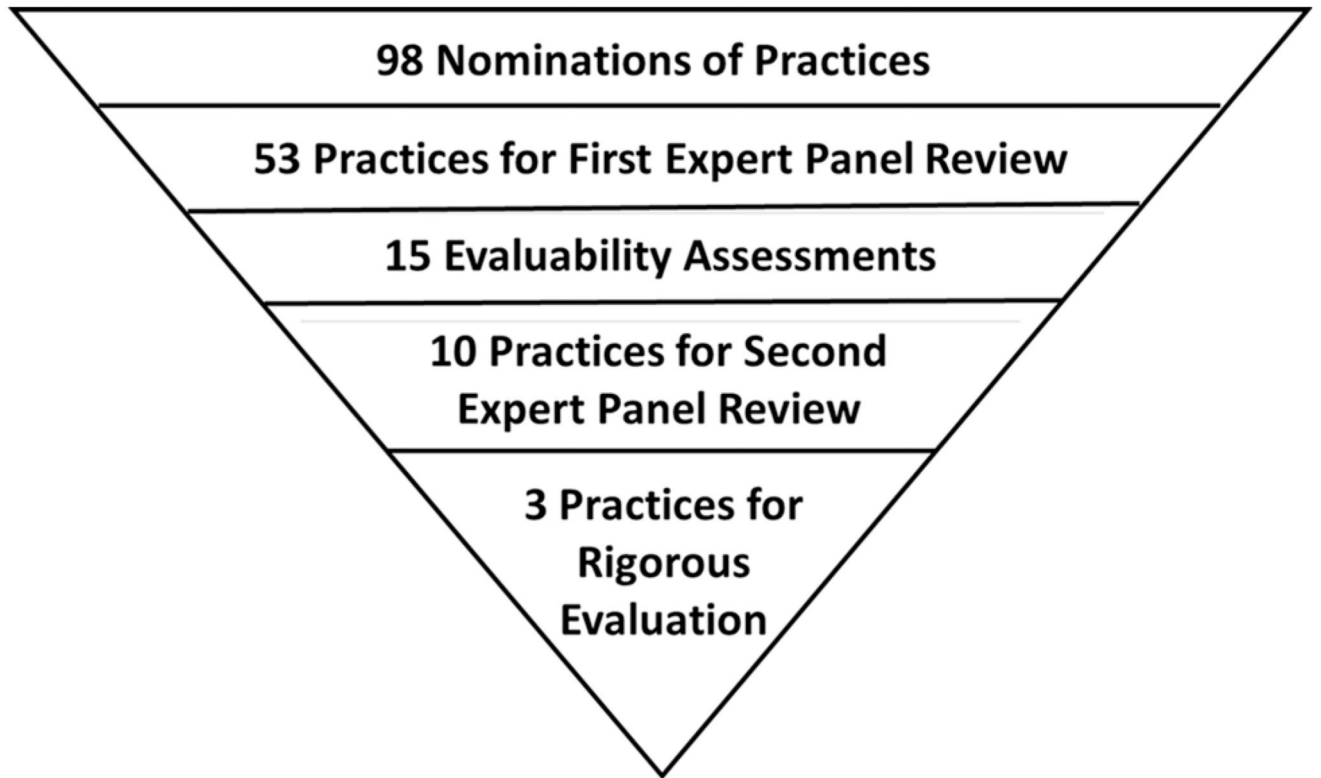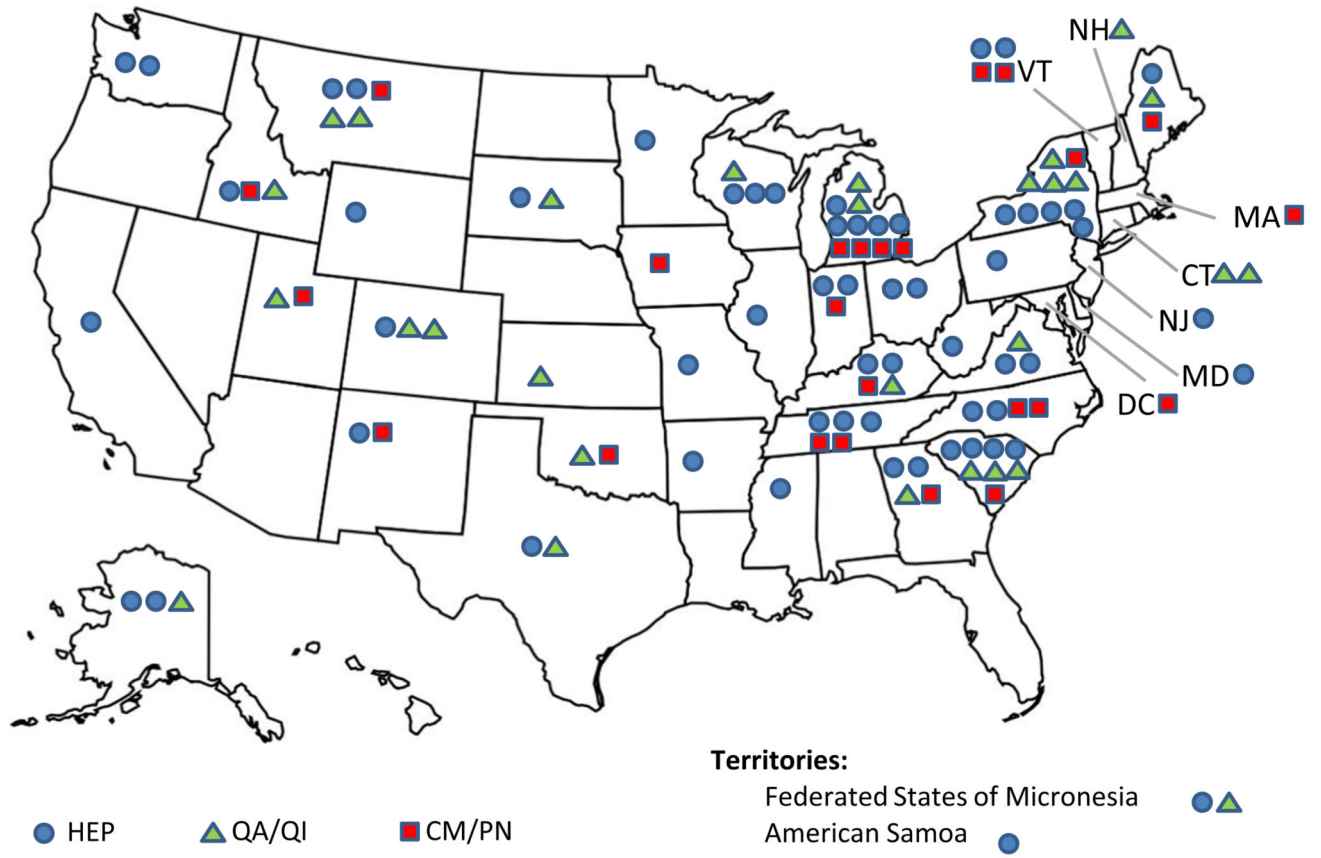
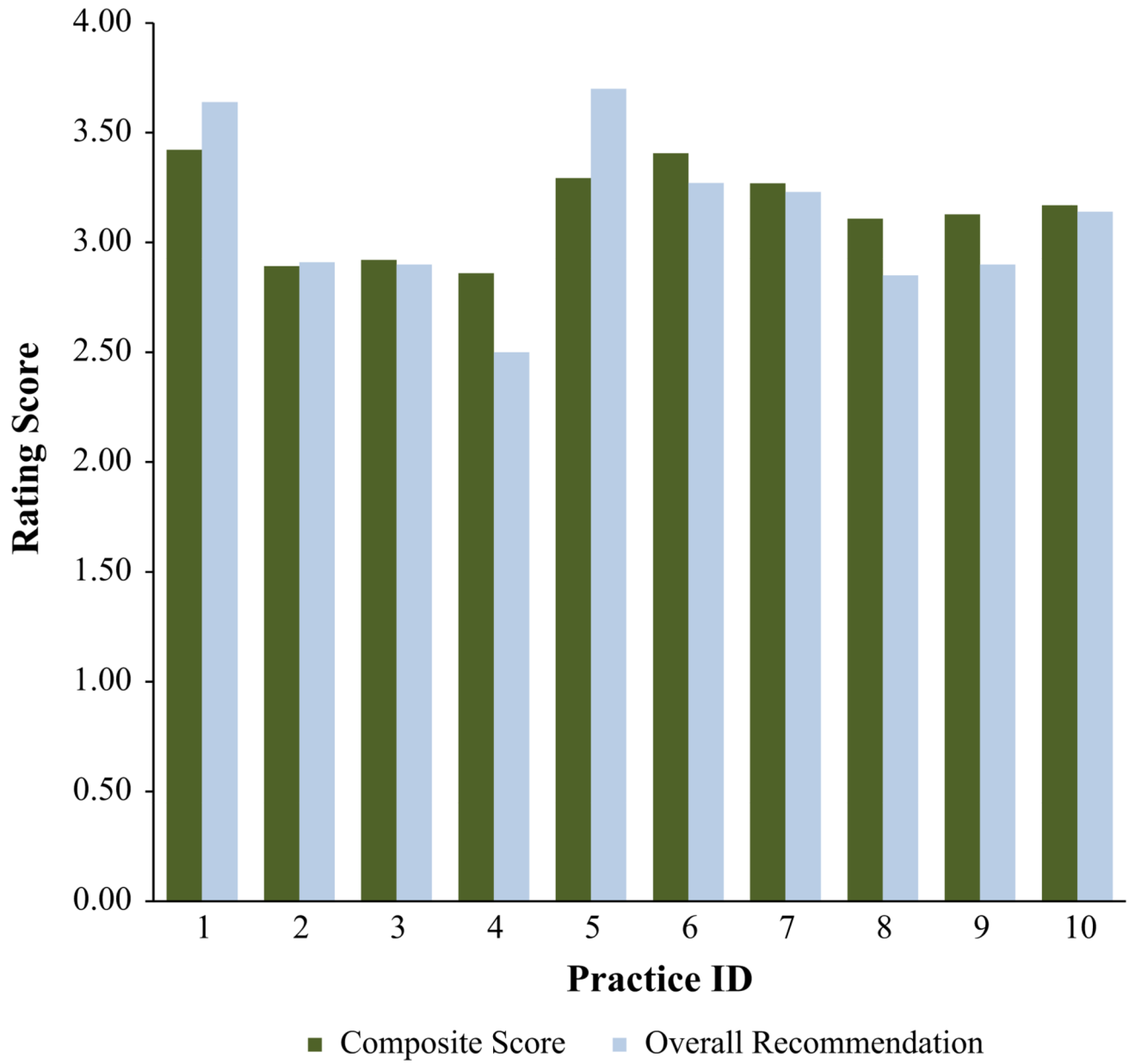**Fig. 2.**
Distribution of nominated practices

Fig. 3.
Composite and overall recommendation scores for final ten practices

Table 1

Priority areas and assessment criteria

Practice categories	Definition
Priority areas	
Health education and promotion (HEP)	Practices that include public education to convey information on the need for, benefits of, and ways to overcome barriers to breast or cervical cancer screening with the goal of informing, encouraging, and motivating people to seek recommended screening
Quality improvement/quality assurance/ professional development (QA/QI)	Practices that increase health care providers' knowledge, awareness, and adherence to breast and cervical cancer screening and quality guidelines
Case management/patient navigation (CM/PN)	Practices that support women through the screening, diagnostic, or cancer care process to ensure adherence and timely care; often includes use of case managers or patient navigators to help women overcome barriers and obtain services

Assessment criteria	Definition
Expert panel rating tool assessment	
Potential impact	Intervention appears to have potential for impact on increasing screening rates for breast and cervical cancer
Reach to target population	Percentage of the target population that is reached or in some other way positively affected by the intervention
Acceptability to stakeholders	Potential or actual evidence that the intervention is acceptable and even attractive to pertinent collaborators, gatekeepers, and other necessary groups, such as community members, physicians, and patients
Feasibility of implementation	Likelihood that the intervention as designed can be or has been implemented fully, given the clarity of its goals, objectives, and strategies; complexity and leadership requirements; financial and other costs; and training and supervision requirements
Feasibility of adoption	Potential for other grantees, sites, or entities to adopt the intervention—particularly for multiple states or regions or racial/ethnic groups
Transportability or generalizability	Degree to which the intervention demonstrates or has potential to be adapted for other settings that differ in size, resources, and demographics
Intervention sustainability	Likelihood that the intervention can continue over time without special resources or extraordinary leadership
Staff and organizational capacity	Sponsoring organization and staff have the capacity to participate fully in a brief assessment, learn from it, and further develop the intervention
Sustainability of health effect	Intended health effect of the intervention will endure over time

Table 2

Ratings of final ten practices with evaluability assessments

Assessment criteria	Ratings by practice ID									
	1	2	3	4	5	6	7	8	9	10
Reach to target audiences	3.46	4.00	3.20	2.86	3.09	3.45	3.55	3.29	3.35	3.64
Acceptability to stakeholders	3.83	2.45	3.45	3.79	3.77	3.75	3.64	3.35	3.77	3.50
Feasibility of implementation	3.75	2.77	3.65	3.63	4.00	3.68	3.42	3.50	3.55	3.54
Feasibility of adoption	2.60	2.00	1.80	2.08	2.20	2.82	2.92	2.77	2.27	2.08
Generalizability	3.05	3.00	2.67	2.42	2.90	3.00	3.42	3.21	2.60	2.83
Sustainability of health effect	3.55	4.00	3.43	3.08	3.60	3.64	3.22	3.42	3.45	3.67
Practice sustainability	3.00	3.10	1.90	2.08	2.73	3.00	2.92	2.38	2.36	2.43
Potential impact	3.59	3.42	3.28	3.27	3.59	3.59	3.08	3.41	3.30	3.88
Staff capacity	3.83	3.40	3.00	2.58	3.64	3.73	3.54	2.77	3.50	3.07
Composite score	3.42	2.89	2.92	2.86	3.29	3.41	3.27	3.11	3.13	3.17
Overall recommendation	3.64	2.91	2.90	2.50	3.70	3.27	3.23	2.85	2.90	3.14