A Graphical, Computer-Based Decision-Support Tool to Help Decision Makers Evaluate Policy Options Relating to Physical Activity

Antronette K. Yancey, MD, MPH, Brian L. Cole, DrPH, William J. McCarthy, PhD

Abstract: This pilot study builds on efforts to develop evaluation methods to compare and contrast potential strategies designed to increase population physical activity generally, and to reduce disparities in activity levels more specifically. The study presents a user-friendly, semi-quantitative decision-support tool of intermediate complexity that may better enable quick, flexible first-pass "ballpark" decision making by state and local health agencies instead of traditional evidence-based scientific reviews. The tool produces a summary score from ratings on 18 criteria, adjusted by fixed or variable weights to incorporate salient community contextual factors. Stair use, workplace activity breaks, and school construction siting are presented as samples. This first iteration of the decision-support tool is intended to be refined empirically by the experiences and policy outcomes of agencies adopting the innovation. This decision-support tool may expand the capacity of public health practitioners to conduct first-pass assessments of policy options for physical activity promotion in underserved communities.

(Am J Prev Med 2010;39(3):273–279)

Introduction

There is broad recognition of the need to bridge the gaps among research, policy, and practice.[1–6] The NIH have made "translational research" a major priority.[7] Large investments are being made in systematic reviews to synthesize research to identify "best practices and inform policy," such as *The Community Guide*,[8] Cancer Control P.L.A.N.E.T.[9] and the emerging field of health impact assessment (HIA).[10]

Merely distilling research findings and repackaging them in a form that can be understood by lay audiences, however, is usually not sufficient to bridge this gap,[2,3,11,12] in part because the questions facing policy makers differ from those addressed in systematic reviews. Although research seeks to perfect knowledge, policymaking aims to advance actions that appear most likely to produce desired near-term results. For the latter, "more research is needed" is not an acceptable justification for inaction in the face of pressing problems.

Even when there is general agreement among researchers about the causes of a particular health problem, ambiguity in research outcomes, questions about the validity of extrapolating to different populations, trade-offs and value judgments about the allocation of scarce public resources, and, of course, politics make it impossible to identify a single best solution.[2] Researchers in organizational behavior have described this kind of decision making with insufficient time and resources for exhaustive scrutiny, considerable uncertainty, and less than ideal options as "muddling through."[13] Adjustments along the way are to be expected.

Assessing and melding evidence from a variety of sources, along with gaps in the evidence base, present substantial challenges. Scientifically generated information can be evaluated by appropriate scientific standards, including precision, replicability, explanatory capacity, and predictive accuracy. The dramatic growth in evidence-based reviews in medicine,[14] public health,[15,16] and social policy[17] has yielded a number of criteria for grading scientific evidence. Differences between the aims of the scientific process and policymaking, however, create an imperfect fit between standards for judging scientific research and those used to make informed policy

From the Department of Health Services (Yancey, Cole, McCarthy), UCLA Kaiser Permanente Center for Health Equity (Yancey, McCarthy), Center for Health Policy Research (Yancey, Cole), Division of Cancer Prevention and Control Research (Yancey, McCarthy), School of Public Health, and the Psychology Department (McCarthy), University of California Los Angeles, Los Angeles, California

Address correspondence to: Antronette K. Yancey, MD, MPH, UCLA School of Public Health, 31-235 CHS, 650 Charles Young Drive South, Los Angeles CA 90095. E-mail: ayancey@ucla.edu.

0749-3797/$17.00
doi: 10.1016/j.amepre.2010.05.013

options. Definitions of evidence, purposes served, and rules for judging its acceptability and value are highly contextual and vary among fields.

This paper proposes a computer-based, graphically oriented decision-support tool (DST) to focus and guide community-level discussions involving complex health policy decision making. The tool aims to facilitate deliberative processes, not necessarily to specify the best policy option. It was first piloted on policies affecting physical activity because its development was supported by a state health department effort to assist local health departments in this relatively new arena of public health practice. Throughout the development process, however, efforts were made to make the application of the tool generalizable to other domains of public health.

The DST was conceptualized as a practical tool to complement evidence-based systematic reviews, such as the *Community Guide*,[8] that could integrate generalizable evidence from the scientific literature with locally specific, contextualized information. It comes out of the recognition that the systematic scientific reviews were not addressing the specific, somewhat unique, local issues confronting policymakers where the potential health effects of the options being weighed rarely, if ever, have the level of certainty required of studies included in systematic reviews. The DST was created in response to a specific request from the California Department of Health Services to provide guidance on assessing policies to promote physical activity, particularly among populations that are more sedentary and for which current promotion activities have been of limited effectiveness. Also shaping these efforts to develop the DST were the concurrent efforts of project team members to train local health departments and community groups in HIA methods as a part of a foundation-supported community-wide obesity intervention.[18,19] Because of short time lines, competing priorities, and limited budgets, conducting a comprehensive HIA of policy options is rarely an option. For those instances in which some knowledge of potential health effects is available, but a full HIA is not feasible, other tools are needed to quickly synthesize, evaluate, and communicate options.

See related Commentary by Brennan in this issue.

Methods

The DST evolved through numerous iterations. Building on the experience of the *Community Guide* and other systematic reviews of public health interventions, all of these iterations used a range of criteria to rate multiple interventions. As project staff applied these iterations to different scenarios and considered the needs of potential users, additional capabilities were added, including allowing users to set their own weights for different rating criteria and to specify additional criteria. Efforts were also made to make both the rating process and output graphically based to increase the DST's user-friendliness and intuitiveness.

Selection of Criteria for the Decision-Support Tool

Based on a review of the criteria others have used for comparing physical activity, nutrition, and related interventions and the public health practice experiences of the investigators,[20–31] policy assessment factors were identified and then condensed into a list of 18 specific criteria (Figure 1).

The first version of the DST was pilot tested by five investigators and staff members (see below for Rater procedures). Subsequently, the assessment criteria were rephrased for clarity and to enhance generalizability beyond the physical activity interventions being evaluated in the initial testing. The criteria were grouped into four categories: feasibility, quality of evidence, population health impact, and disparities reduction. This iteration of the DST also allows users to specify additional criteria of their own. Graphic displays for both inputs (i.e., ratings pages) replaced the numeric-only format in the first iteration for visual simplicity and ease of interpretation. The fundamental rating procedures and weighting, however, were not changed.

The aim of using these criteria is to augment and facilitate careful judgment, not to replace it with mechanistic algorithms for determining feasibility and effectiveness. The tool provides a first-pass analysis by effectively and efficiently communicating the state of existing community-level evidence. Judgments about decisions important to the community may be rendered more confidently and expeditiously. Although not a substitute for full-fledged policy analysis, the tool can be used to quickly assemble, organize, and communicate information when more in-depth policy analysis is not feasible.

Description of the Decision-Support Tool

The DST (Figures 1 and 2) is a graphical, spreadsheet-based tool designed to help decision makers systematically synthesize, weigh, and compare evidence specific to public health–relevant interventions and policies in a relatively short time frame. Informed, deliberative decision making is supported by presenting ratings of alternatives in multiple graphical displays that facilitate easy comparison. The DST also guides group discussion about the relative importance of different rating criteria for a given decision and indicates how to rate alternatives based on these criteria. The DST allows comparison of up to five alternatives rated on 18 preset rating criteria encompassing four domains: feasibility, evidence of effectiveness, population impact, and disparities reduction. Up to four additional user-specified criteria can also be included. Given inherent uncertainties about the potential effects of alternatives, differing opinions, and inconsistent evidence about the relative merits of alternatives in different domains, the DST does not aim to select an overall best alternative or to replace judgment about trade-offs and uncertainty, but rather it aims to support reasoned and transparent decision

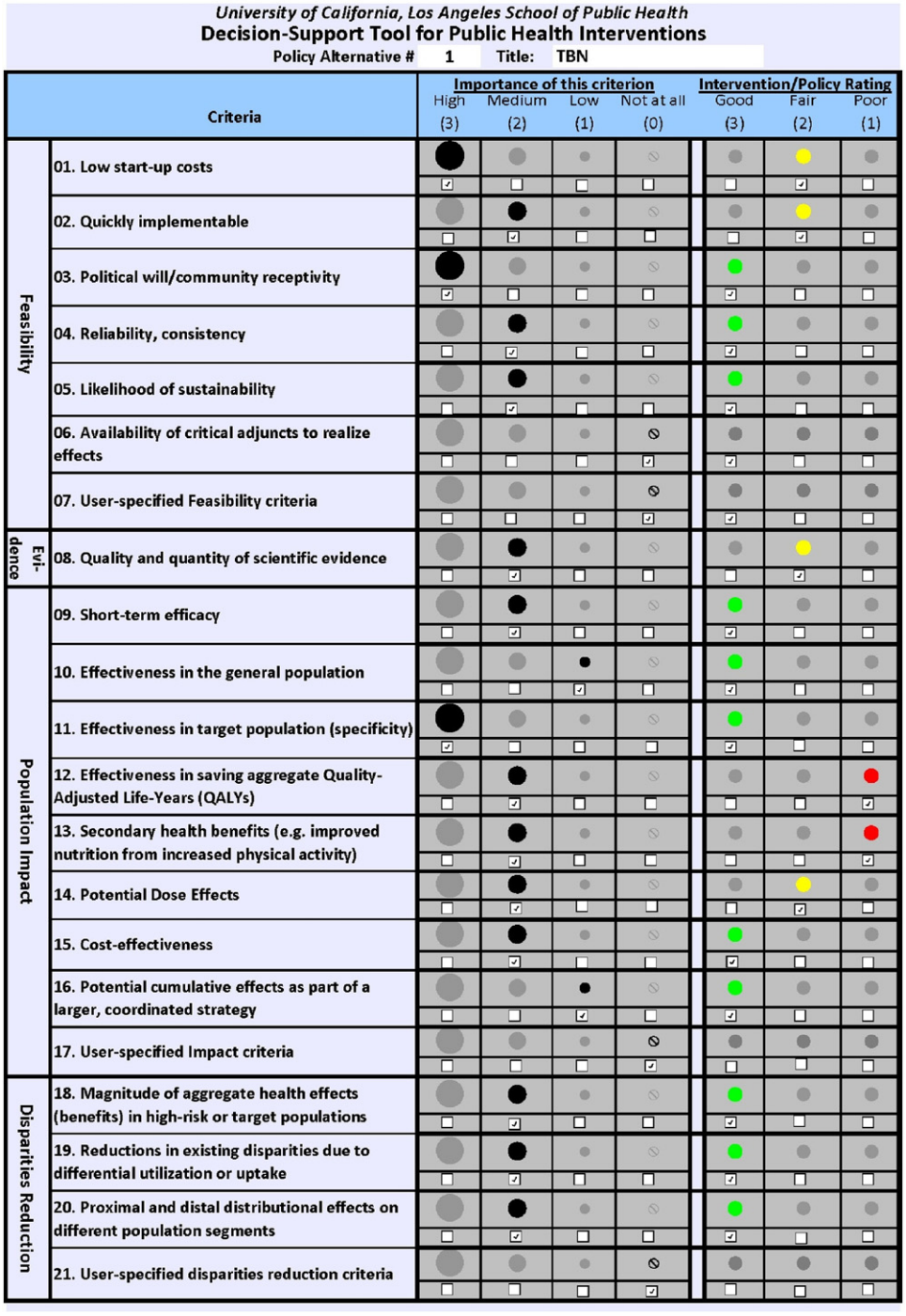
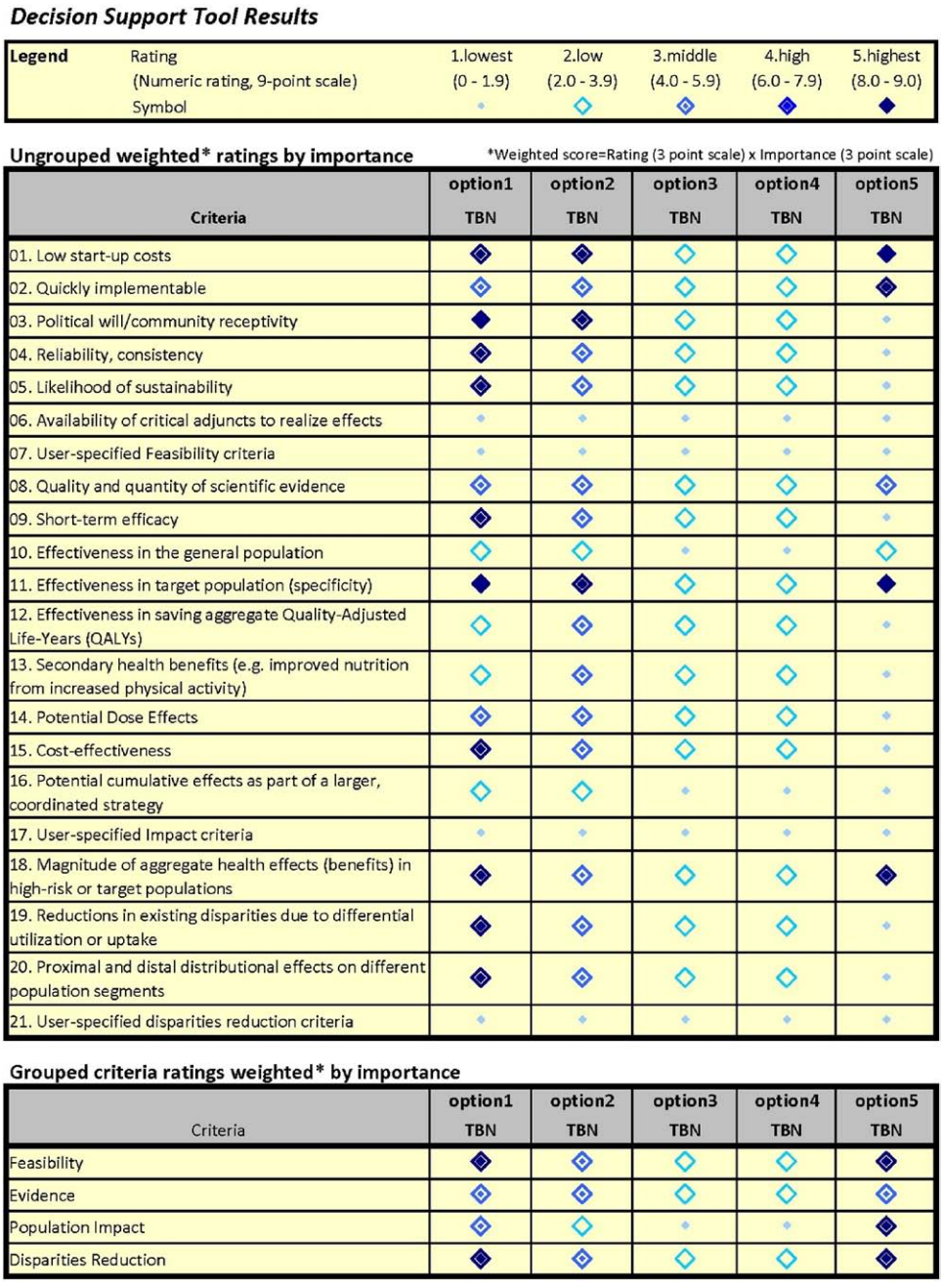


Figure 1. Sample rating page for a policy option with user weights and ratings shown

making and to prompt questions that should be addressed before choosing an alternative.

Rater Procedures

Although the DST may be used by an individual, it is ideally suited for use by a small group that collectively evaluates and rates each specific alternative. Step-by-step instructions guide users on optimal use of the DST. The first step is to define up to five alternative policies or interventions for comparison. Next, users specify the importance (i.e., weight) of each of the 18 preselected rating criteria on a 4-point scale from 0 (*not at all important*) to 3 (*highly important*). Users may also add up to four of their own rating criteria that are not included in the preset list. After reviewing available evidence (e.g., research studies, reports, expert testimony, and stakeholder comments), they score each of the rating criteria for a given alternative from 1 (*poor*) to 3 (*good*), then repeat this process for the remaining alternatives. The DST then displays the weighted scores for the alter-



Decision Support Tool Results

Legend	Rating	1.lowest	2.low	3.middle	4.high	5.highest
	(Numeric rating, 9-point scale)	(0 - 1.9)	(2.0 - 3.9)	(4.0 - 5.9)	(6.0 - 7.9)	(8.0 - 9.0)
	Symbol					

Ungrouped weighted* ratings by importance

*Weighted score=Rating (3 point scale) x Importance (3 point scale)

Criteria	option1 TBN	option2 TBN	option3 TBN	option4 TBN	option5 TBN
01. Low start-up costs					
02. Quickly implementable					
03. Political will/community receptivity					
04. Reliability, consistency					
05. Likelihood of sustainability					
06. Availability of critical adjuncts to realize effects					
07. User-specified Feasibility criteria					
08. Quality and quantity of scientific evidence					
09. Short-term efficacy					
10. Effectiveness in the general population					
11. Effectiveness in target population (specificity)					
12. Effectiveness in saving aggregate Quality-Adjusted Life-Years (QALYs)					
13. Secondary health benefits (e.g. improved nutrition from increased physical activity)					
14. Potential Dose Effects					
15. Cost-effectiveness					
16. Potential cumulative effects as part of a larger, coordinated strategy					
17. User-specified Impact criteria					
18. Magnitude of aggregate health effects (benefits) in high-risk or target populations					
19. Reductions in existing disparities due to differential utilization or uptake					
20. Proximal and distal distributional effects on different population segments					
21. User-specified disparities reduction criteria					

Grouped criteria ratings weighted* by importance

Criteria	option1 TBN	option2 TBN	option3 TBN	option4 TBN	option5 TBN
Feasibility					
Evidence					
Population Impact					
Disparities Reduction					

Figure 2. Results displayed in iconic form (results may also be viewed in numeric form, bar graphs, and as a scatterplot)

natives in three different formats—table of numbers; table of icons (similar to *Consumer Reports*®; see Figure 2); and bar graphs. Ratings can be viewed in full detail, with the ratings of each criterion shown or aggregated into the four rating domains. Seeing the visual comparison of alternatives, users can quickly identify the relative strengths and weaknesses of each alternative, identify a preferred alternative, and discuss options for improving on the selected alternative.

Pilot Testing

Pilot testing of the original DST (numeric rating without graphics) was conducted by three members of the investigative team who reviewed the literature on three different environmental approaches for increasing physical activity: (1) stair-use prompts in workplaces; (2) school siting/design; and (3) physical activity breaks in workplaces. These examples were chosen to illustrate a wide variety of considerations and their handling by the DST, to allow comparison of similar policies (i.e., examples 1 and 3) and to test the robustness of the DST when confronted with a more complex policy (i.e., example 2). The school siting policy was more difficult to assess with the DST because there is a less-developed body of literature and also because it is not primarily a public health intervention.

Definitions and the basis of ratings for the workplace activity break example are shown in Appendixes A and B (available online

Table 1. Concordance of ratings for three sample policies assessed using the decision-support tool in pretesting

Policy	M % agreement	Kappa inter-rater agreement	SE
Stair-use promotion in workplaces	60.8	0.5785	0.0782
School siting (in relation to effects on physical activity)	50.0	0.1091	0.1299
Workplace activity breaks	46.9	0.1563	0.0778

Note: $n=3$ raters

at www.ajpm-online.net). A similar table for the workplace stair-use prompts is available online at www.ph.ucla.edu/hs/health-impact/methodology.htm. To facilitate comparisons and discussions of the alternatives, the weighted ratings are presented in multiple numeric and graphical formats.

Pilot testing was done by individuals working independently to verify the clarity of items and reliability. After the testers had become thoroughly familiar with the research literature on the three topics, they independently rated each policy option (school siting, workplace activity breaks, and prompts of stair usage) by the 18 decision criteria. The project data manager then evaluated the degree of concordance among raters for each item on each policy option using the kappa procedure in SAS version 9.2.

Results

Pilot Study Findings

Table 1 shows the mean kappa statistic measure of inter-rater agreement for the 18 DST items as applied to three examples affecting physical activity levels: promotion of stair use in workplaces, school siting, and workplace activity breaks (e.g., 10-minute organized exercise breaks). Overall, mean inter-rater agreement for the three policies was 46.9%–60.8%, with kappa ratings of 0.58, 0.10, and 0.16, based on the ratings of three observers.

Discussion and Implications

The DST guides users through an evaluation of what is known about a set of policy and environmental change interventions, prompting them to consider feasibility, the strength of research evidence, population impact, and influence on disparities. It allows them to weigh the relative importance of rating criteria. Multiple options are provided for viewing concise, easy-to-understand text and graphical summaries.

The interactive capacity of the tool prompts users to question and discuss rating criteria and rankings, and to incorporate their own judgments. Unlike evidence-based scientific reviews, the DST explicitly treats the research-to-policy interface as an iterative, bi-directional exchange of information between equals with disparate and often divergent aims and constraints. Measures of efficacy and effectiveness are used by researchers to assess a set of intended outcomes, while policymakers need to weigh diverse outcomes, affecting different constituencies. RCTs, the gold standard for researchers, are rarely available in areas being explored for policy action. Where such trials have been conducted, the findings must be integrated with and weighted against other types of information that may be equally relevant.

Both peer-reviewed and "gray" literature has informed the development of this decision-support tool. Particular emphasis has been placed on identifying and reducing health disparities among underserved ethnic groups, gender groups, and income groups.

Limitations

Ratings were more discordant for the exercise break and the school siting policies than for stair use, despite the fact that the investigators were exposed to the same literature and had worked together for many years. Several explanations for this discrepancy are possible: (1) The stair usage literature was more mature and extensive than was the literature on the utility of structural integration of short activity bouts or the physical activity–related consequences of locating new school construction, so that the raters' judgments of the relative feasibility and effectiveness of stair prompts were informed by consensus in the field to a much greater extent than the corresponding judgments governing where new schools should be located or whether and in what settings exercise break should be implemented; (2) the raters' backgrounds and disciplines are quite disparate with respect to practice experience; and (3) the complexity of decisions regarding where to locate a new school was considerably greater than that regarding how best to motivate public use of the stairs or organizational uptake of exercise break policies and practices. It is somewhat like comparing the decision inputs into the penalty for running a red light and that accompanying a homicide conviction. Considerably more deliberation understandably occurs in the latter instance. Similarly, no one would argue that the choice of where to locate a school should depend exclusively or even primarily on how it affects students' physical activity, whereas the decision about stair prompts or activity

breaks depends almost entirely on its demonstrated effectiveness in increasing physical activity.

Factors limiting application of this tool include:

1. the paucity of intervention outcome data, especially for environmental interventions, and particularly in low-income communities of color[32–35];
2. missing information on rating criteria used in the DST applicable to specific communities and target populations;
3. disparate dimensions that may be conflated into single measures; and
4. the small evidence base for empirically validating the tool.

Dissemination, Training, and Further Development

The DST is publically available on the authors' university-based website (www.ph.ucla.edu/hs/health-impact/methodology.htm). Along with written instructions on how to use the DST, the website includes a short video showing users how to complete the rating procedures and how to interpret results. Even with these user supports, the tool will undoubtedly be daunting to some potential users. Visitors to the website, whether or not they use the DST, are encouraged to provide feedback on the tool, its support, and potential applications. Users are also asked to submit their results to the website for posting in a gallery of DST results. Specific steps needed to advance the DST include:

1. refinement of graphical interfaces to allow users with minimal training to use the tool;
2. expanded database of proven applications; and
3. multiple, embedded self-training modules to show users how to use the tool and to guide them toward its appropriate use.

The current version of the DST requires some knowledge of health interventions in order to assign relative weights and understand technical terms. Definitions of terms can be embedded in future versions to make the tool more accessible. Providing users with the broader, conceptual knowledge of health and intervention methodology is probably beyond the self-training capabilities of the tool. Input from pretesting the DST with legislative aides and health department analysts whose expertise encompasses both policymaking and public health is recommended for refining the tool for use by a wider audience of decision makers and community stakeholders. Tool limitations and strategies for their mitigation should be highlighted in accompanying user-support documents and services.

Decisions affecting the fitness and well-being of communities are made regularly by legislators and their staff, local health department senior managers, school officials, and advocacy groups without the benefit of this structured approach. Use of the DST would ensure a more systematic and evidence-based assessment. Clearly, there is demand for such an approach, as this work has already been utilized in Robert Wood Johnson Foundation, IOM,[36] and National Physical Activity Plan (www.physicalactivityplan.org) deliberations and shares some similarities with a recent CDC effort to prioritize environmental obesity control interventions.[37] Evidence-based refinement of the tool is a critical next step, as is adapting the tool for web-based applications. The ultimate indicator of decision-support tool utility, however, is the effectiveness and sustainability of interventions adopted in the "real world" through its use.

This study was supported by the USDA-funded Nutrition Network of the Cancer Prevention and Nutrition Section, California Department of Public Health, CDC-funded Cancer Prevention and Control Research Network, and core funding for the CEHD.

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

References

1. Anderson LM, Brownson RC, Fullilove MT, et al. Evidence-based public health policy and practice: promises and limits. Am J Prev Med 2005;28(5S):226–30.
2. Brownson RC, Royer C, Ewing R, McBride TD. Researchers and policymakers: travelers in parallel universes. Am J Prev Med 2006; 30(2):164–72.
3. Lang T, Rayner G. Overcoming policy cacophony on obesity: an ecological public health framework for policymakers. Obes Rev 2007; 8(1S):165–81.
4. Sorian R, Baugh T. Power of information: closing the gap between research and policy. Health Aff (Millwood) 2002;21(2):264–73.
5. WHO. World Report on knowledge for better health: strengthening health systems. Geneva: WHO, 2004.
6. WHO, Canada. Public Health Agency of Canada. Preventing chronic diseases: a vital investment. Ottawa: WHO, Public Health Agency of Canada, 2005.
7. Woolf SH. The meaning of translational research and why it matters. JAMA 2008;299(2):211–3.
8. Preventive Services Taskforce: Guide to Community Preventive Services. www.thecommunityguide.org/index.html.
9. Consortium of cancer control organizations. Cancer Control P.L.A.N.E.T. (Plan, Link, Act, Network with Evidence-based Tools). ccplanet.cancer.gov/sponsors.html.
10. Cole BL, Fielding JE. Health impact assessment: a tool to help policy makers understand health beyond health care. Annu Rev Public Health 2007;28:393–412.
11. Davis P, Howden-Chapman P. Translating research findings into health policy. Soc Sci Med 1996;43(5):865–72.

12. Lavis JN. Research, public policymaking, and knowledge-translation processes: Canadian efforts to build bridges. J Contin Educ Health Prof 2006;26(1):37–45.
13. Lindblom CE. The science of "muddling through." Public Admin Rev 1959;19(2):79–88.
14. Haynes RB. What kind of evidence is it that evidence-based medicine advocates want health care providers and consumers to pay attention to? BMC Health Serv Res 2002;2(1):3.
15. Kohatsu ND, Robinson JG, Torner JC. Evidence-based public health: an evolving concept. Am J Prev Med 2004;27(5):417–21.
16. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. Am J Public Health 2004;94(3):400–5.
17. Davies P, Boruch R. The Campbell collaboration: does for public policy what Cochrane does for health. Br Med J 2001;323(7308):294–5.
18. Solomon L, Standish MB, Orleans CT. Creating physical activity-promoting community environments: time for a breakthrough. Prev Med 2009;49:334–5.
19. Samuels SE, Bullock SL, Woodward-Lopez G, et al. To what extent have high schools in California been able to implement state-mandated nutrition standards? J Adolesc Health 2009;45(3S):S38–44.
20. Blumberg LJ. Balancing efficiency and equity in the design of coverage expansions for children. Future Child 2003;13(1):205–11.
21. Finkelstein E, French S, Variyam JN, Haines PS. Pros and cons of proposed interventions to promote healthy eating. Am J Prev Med 2004;27(3S):163–71.
22. James C, Carrin G, Savedoff W, Hanvoravongchai P. Clarifying efficiency–equity tradeoffs through explicit criteria, with a focus on developing countries. Health Care Anal 2005;13(1):33–51.
23. Jha P, Nagelkerke JD, Ngugi EN, et al. Public health. Reducing HIV transmission in developing countries. Science 2001;292(5515):224–5.
24. Lightwood J. The economics of smoking and cardiovascular disease. Prog Cardiovasc Dis 2003;46(1):39–78.
25. Pavlovich WD, Waters H, Weller W, Bass EB. Systematic review of literature on the cost-effectiveness of nutrition services. J Am Diet Assoc 2004;104(2):226–32.
26. Schmid TL, Librett J, Neiman A, Pratt M, Salmon A. A framework for evaluating community-based physical activity promotion programmes in Latin America. Promot Educ 2006;13(2):112–8.
27. Summerbell CD, Waters E, Edmunds LD, Kelly S, Brown T, Campbell KJ. Interventions for preventing obesity in children. Cochrane Database Syst Rev 2005(3):CD001871.
28. Tengs TO. Cost-effectiveness versus cost-utility analysis of interventions for cancer: does adjusting for health-related quality of life really matter? Value Health 2004;7(1):70–8.
29. Tsai AG, Glick HA, Shera D, Stern L, Samaha FF. Cost-effectiveness of a low-carbohydrate diet and a standard diet in severe obesity. Obes Res 2005;13(10):1834–40.
30. Wang LY, Yang Q, Lowry R, Wechsler H. Economic analysis of a school-based obesity prevention program. Obes Res 2003;11(11):1313–24.
31. Weightman A, Ellis S, Cullum A, Sander L, Turley R. Grading evidence and recommendations for public health interventions: developing and piloting a framework. London: Health Development Agency, 2005.
32. Kumanyika SK. Minisymposium on obesity: overview and some strategic considerations. Annu Rev Public Health 2001;22:293–308.
33. Marcus BH, Williams DM, Dubbert PM, et al. Physical activity intervention studies. What we know and what we need to know. A scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity); Council on Cardiovascular Disease in the Young; and the Interdisciplinary Working Group on Quality of Care and Outcomes Research. Circulation 2006;24(114):2739–52.
34. Matson-Koffman DM, Brownstein JN, Neiner JA, Greaney ML. A site-specific literature review of policy and environmental interventions that promote physical activity and nutrition for cardiovascular health: what works? Am J Health Promot 2005;19(3):167–93.
35. Yancey AK, Kumanyika SK, Ponce NA, et al. Population-based interventions engaging communities of color in healthy eating and active living: a review. Prev Chronic Dis 2004;1(1):A09.
36. IOM (U.S.). Local government actions to prevent childhood obesity. Washington DC: National Academies of Sciences, 2009.
37. Khan LK, Sobush K, Keener D, et al. Recommended community strategies and measurements to prevent obesity in the U.S. MMWR Recomm Rep 2009;58(RR-7):1–26.

Appendix

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.amepre.2010.05.013.