Evaluating comprehensive tobacco control interventions: challenges and recommendations for future action

Report of a workshop convened by the Institute for Global Tobacco Control, Johns Hopkins Bloomberg School of Public Health

This report summarises a workshop on evaluation of tobacco control interventions convened in Santa Fe, New Mexico in June 2001 by the Institute for Global Tobacco Control at the Johns Hopkins Bloomberg School of Public Health. The evaluation of such interventions is filled with complexities which intensify as the scope of tobacco control programmes increase. Evaluators are charged with the task of determining the effect of interventions in terms of magnitude of change, the relative contribution of programme components, and the relative impact for different populations. The report explores the theoretical foundations of tobacco control evaluation and provides a conceptual framework for capturing elements necessary for evaluating interventions. It then provides two case studies of challenges encountered when evaluating large scale tobacco control initiatives. The report summarises the discussions and recommendations of the workshop’s three working groups. Participants were certain and unanimous that the current state of evaluation research must be improved to evaluate accurately the dynamic nature of comprehensive tobacco control programmes. Hierarchical or multilevel modelling approaches were seen as promising for further research. Coordinated evaluation will provide a better understanding of local, state, and national tobacco control efforts.

This report describes the findings and recommendations of a workshop on evaluation of tobacco control interventions convened on 2 June 2001 in Santa Fe, New Mexico, USA by the Institute for Global Tobacco Control of the Johns Hopkins Bloomberg School of Public Health. Experts from around the world in epidemiology, biostatistics, economics, evaluation, health data systems, health communication, health behaviour and policy, clinical treatment, and computer simulation, were in attendance (see appendix). Attendees were representing a broad range of key stakeholder groups including academia, government, non-profit organisations, and private industry.

As the scientific foundation for tobacco control has advanced and funding for tobacco control programmes has increased, tobacco control interventions have become more comprehensive by including prevention, cessation and policy directed objectives at the community, state, and national levels. The expanded focus of these tobacco control programmes includes strategies aimed at changing individual behaviour, as well as strategies aimed at altering the social and political environments. As the complexity of comprehensive tobacco control programmes increases so does the difficulty associated with their evaluation. These programmes do not fit into the tightly defined, controlled, and presumably reproducible research model that is more suitable to epidemiologic testing. At present, unsuitable measures and methods are often used to assess interventions, including insufficient units of analysis and duration of follow up; errors in sampling; and gaps between the intervention units and the analysis methods. Appropriate measures and methods for evaluating community, state and national programmes are urgently needed to understand what works, to demonstrate success, to justify the costs of tobacco control interventions, and to assure continued funding.

A FRAMEWORK FOR APPROACHING EVALUATION CHALLENGES

A conceptual framework was developed based on workshop group discussions. This framework (fig 1) places the individual within a population context and acknowledges the dynamic relationship between individual characteristics, tobacco industry tactics, tobacco control interventions, and outcomes associated with changing policy, behaviour and disease outcomes. While simplistic, the framework captures elements necessary for evaluating tobacco control interventions, including the multiple factors that influence initiation and maintenance of smoking by individuals; various levels at which interventions occur (individual, local, state and national); and the competing influences of tobacco control interventions versus tobacco industry activities.

Individual factors

In fig 1, individuals are considered as having key characteristics (sex, age, race, socioeconomic status, and geographic location) in addition to their smoking status. Since time alters an individual’s key characteristics (that is, age increases, people move, smoking status changes) the framework is dynamic. Individuals have different motivations

Abbreviations: ASSIST, American Stop Smoking Intervention Study; IOC, initial outcomes index, NIC, National Cancer Institute; STOTC, strength of tobacco control; TCP, California Tobacco Control Prevention and Education Program
to change their behaviour; an individual’s motivations also may vary depending on such psychological factors as social support, readiness to change, and on how they balance the relative benefits and costs of quitting smoking. Consequently, the effectiveness of tobacco control initiatives, as well as any actions to counter the influence of tobacco industry, will depend on these factors.

**Population factors**
Specific characteristics of a state or community can influence tobacco control efforts, including regional, cultural, economic, and political factors. For example, greater resistance to state tobacco control policies is often found in more politically conservative communities where tobacco growing and production is a major activity and where the tobacco industry is considered an important component to the community’s economy. In addition, geographic location can greatly affect tobacco control outcomes; for example, urban communities often have different smoking rates and patterns than rural communities. The implementation and the evaluation of tobacco control interventions need to be informed by such community factors.

It can be difficult to isolate the effects of a particular tobacco control programme from others that may be reaching the same target. Some programmes may be deliberately interrelated, while others may have unexpected points of synergy between programmes. There is often limited coordination between programmes, and currently surveillance mechanisms are not in place for tracking those that have been already implemented. In the future, knowing how, why and when specific targeted interventions changed their target community will be useful for designing more coordinated efforts.

**Tobacco industry**
There is a broad spectrum of covert and overt tobacco industry actions aimed at countering tobacco control efforts to reduce their sales. Some actions can be tracked through accessible data sources (advertising expenditures, point-of-sale advertising), while others are more difficult to access consistently (number of lobbyists or contributions to political campaigns). The tobacco industry documents contain budgetary data and other information on industry actions, which may be useful. While the issue of tobacco industry behaviour and its impact on programme outcome has been raised in past tobacco control literature, to our knowledge, there have been no systematic attempts to date to incorporate tobacco industry actions into evaluation methods for tobacco control programmes. Since tobacco control efforts are related to the countervailing force exerted by the tobacco industry (more successful tobacco control can trigger more response from the tobacco industry), it is difficult to enter both of these factors into typical regression models. Newer methods for analysing longitudinal data can handle these types of interrelated and multilevel data that are correlated over time. It is also possible to model the dynamic interrelationships and assess changes in policy, behavioural and disease outcomes.

**CASE STUDIES**
Two case studies were presented as “real world” examples of evaluation challenges. The first was the National Cancer Institute’s American Stop Smoking Intervention Study (ASSIST), which represents the largest national intervention study to date. The second case study was California’s Comprehensive Tobacco Control Program (TCP), one of the first comprehensive statewide programmes.

**American Stop Smoking Intervention Study (ASSIST)**
In October 1991, 17 state health departments were awarded contracts by the National Cancer Institute (NCI) to develop and implement the ASSIST project, with the overall goal of demonstrating that the application of statewide tobacco prevention and control programmes and policies would reduce cigarette consumption and smoking prevalence. The principal focus of ASSIST was altering the environmental and social influences affecting the population’s use of tobacco. NCI formed a partnership with the American Cancer Society to accomplish this goal. NCI directed states to focus on the following interventions: (1) developing media advocacy skills to increase pro-tobacco control media coverage; (2) strengthening support for (a) local and state clean indoor air laws, (b)
The primary objective of ASSIST was to demonstrate declines in levels of tobacco consumption and smoking prevalence in ASSIST as compared to the non-ASSIST states. The ASSIST evaluation developed a conceptual framework based on an ecological systems model that provided an opportunity to investigate the complex relationships between the social context, public health activity at the state level, tobacco use, and individual behaviour change. For the ASSIST evaluation, the unit of selection and analysis was the state and many of the factors included in the evaluation were measured only at the state level. However, tobacco control research at the state level provides a maximum of only 50 units (50 states plus the District of Columbia), limiting the number of factors that can be considered in modelling outcomes. Smoking prevalence is analysed using a two stage regression model. With a maximum of only 51 observations, even a modest degree of random variation severely limits the power of the analysis to detect an effect, especially for outcomes such as smoking prevalence. The evaluation relied on finding parsimonious models to represent complex constructs. New exposure measures were developed to assess state level factors. Outcomes were divided into three categories: initial—policy; intermediate—attitudes and behaviour; and final outcomes—prevalence and consumption.

Evaluation of ASSIST provided the opportunity to test a model for reducing tobacco usage at the state level. This was the first attempt to evaluate systematically the effectiveness of state level tobacco control across all states. Because there are only 51 units of analysis, the number of variables that can be included in any regression analysis is limited. Consequently, the ASSIST evaluation relied heavily on developing summary indices. ASSIST programme objectives focused on policy changes, media advocacy, and community mobilisation. Evaluation methodologies attempted to account for variability in individual and state conditions as well as tobacco industry efforts. Evaluators encountered numerous data issues (for example, limited baseline data, no measure of intervention dose or exposure, lack of measure(s) or data for pro-tobacco industry efforts) and design challenges (for example, ASSIST was not a randomised trial; there was diffusion of intervention strategies to non-ASSIST states; and a comprehensive evaluation plan was not put in place at the inception of the project). To deal with these challenges, ASSIST researchers developed a conceptual framework and identified pathways to link inputs to outcomes, and developed two new exposure measures to assess state level factors. Those two measures were the strength of tobacco control (SOTC), and initial outcomes index (IOI). Components of the SOTC measure included state resources allocated to tobacco control efforts, the capacity to implement those tobacco control activities, whether a state had a comprehensive tobacco control plan, and the extent to which its focus included policy change and media advocacy activities. The IOI served as a measure of early indicators of programmatic effectiveness of the intervention(s) was complex since not all surveys used the same methodology, and the interventions differed among local tobacco control activities in terms of type, target, and quality. Evaluators were not able to link programme components to programme outcomes. Although it is difficult to establish relationships between large scale social interventions and a change in tobacco use, numerous changes in intermediate outcomes were noted, including the passage of clean indoor air policies in public places, worksites and bars, and voluntary policies to ban smoking in homes. Giving the programme's longevity, links have also been made to more distal outcomes, including declines in smoking prevalence, and even to reduced mortality from coronary heart disease.

California Tobacco Control Prevention and Education Program

The California Tobacco Control Prevention and Education Program (TCP) was one of the most aggressive anti-tobacco campaigns ever launched, nationally or internationally, and illustrates the complexity of evaluating the dynamic interplay between the individual, intervention(s), and the tobacco industry at the state level. In November 1988, California voters approved the California Tobacco Tax and Health Promotion Act of 1988 (Proposition 99), which increased the state surtax on cigarettes by 25 cents per pack (and an equivalent amount on other tobacco products). Revenues from the new tobacco tax were earmarked for the TCP and directed to tobacco related disease research, health education against tobacco, and health care for medically indigent families. The primary goal of the TCP was to reduce smoking among California adults and adolescents, reduce exposure to secondhand smoke, reduce youth access to tobacco products, and reveal and counter tobacco industry influence. The focus of the TCP was to “denormalise” or reduce the acceptability of tobacco use in California communities. Programme progress was determined by examining trends in per capita cigarette consumption and smoking prevalence.

The TCP was comprehensive and consisted of multiple interventions that addressed individual, social, and environmental factors that contributed to tobacco use. The primary TCP components included community programmes, a statewide mass media campaign, and a school based tobacco use prevention education programme. The “interventions” were not delivered in a uniform manner, but each community decided for itself what was feasible in the local situation.

The evaluation of the TCP included large triennial surveys and smaller ongoing surveys. Determining the extent and effectiveness of the intervention(s) was complex since not all surveys used the same methodology, and the interventions differed among local tobacco control activities in terms of type, target, and quality. Evaluators were not able to link programme components to programme outcomes. Although it is difficult to establish relationships between large scale social interventions and a change in tobacco use, numerous changes in intermediate outcomes were noted, including the passage of clean indoor air policies in public places, worksites and bars, and voluntary policies to ban smoking in homes. Given the programme's longevity, links have also been made to more distal outcomes, including declines in smoking prevalence, and even to reduced mortality from coronary heart disease.

WORKING GROUP DELIBERATIONS AND RECOMMENDATIONS

The participants were divided into three working groups. The first group addressed exposures, or factors influencing smoking initiation and cessation. The second group focused on outcomes, of the dependent variables that can be measured to index the consequences of interventions. The distinction between exposures and outcomes may be arbitrary at times, reflecting the assumed model of causal relationships; a factor that is an exposure in one evaluation might be an intermediate outcome in another (for example, smoke-free workplaces might be considered an “exposure” since their existence can influence smoking behaviour, but also achieving smoke-free workplaces is often considered an intermediate outcome of a tobacco control intervention). The third group discussed methodological approaches to relate exposures to outcomes. All three groups discussed the models and tools currently being used to evaluate interventions at the national, state, and local levels. They explored the relationship between exposures and short, intermediate, and long term outcomes and focused on mechanisms when working on interventions and outcomes with incomplete data and variable data systems.

Group 1: exposures

In the context of evaluating comprehensive tobacco control programmes, “exposures” are factors that are associated with tobacco use reduction. Exposures include media, policies (for example, clean indoor air, taxation, youth access restrictions,
product, norms, and health care system cessation efforts. Also critical to understanding how and why a tobacco control programme works are factors that are not themselves considered “exposures” but that may affect exposures, such as resources expended on tobacco control, community mobilisation, media and policy advocacy and lobbying, community capacity to conduct and sustain tobacco control efforts, community readiness to accept policies, state and local cultural and political factors, and demographics.

Challenges
The group identified the following challenges: heterogeneity of exposure information (for example, types and sources of information, quantitative and qualitative data, assessment of dose delivered and dose received); lack of clarity on how much data are needed and for how long; and the lack of measures for tracking the tobacco industry in a dynamic fashion.

Recommendations
• Standardisation of exposure assessment is critical; standard measures, or “key exposures”, should be identified and collected at the local, state, and national levels.
• Data systems should be in place to track the identified key exposures, making it possible for evaluations to control for other interventions besides that of the index programme.
• National agencies (NCI, Centers for Disease Control and Prevention, American Legacy Foundation, Robert Wood Johnson Foundation, and others) should take a leading role in defining core exposures and approaches for measuring them. A meeting of stakeholders (that is, administrators, programme implementers, health promotion activists, participants, and scientific evaluators) is needed to draft a national surveillance plan which should specify core exposures and measurement strategies.
• NCI’s newspaper database should be used as a resource to assess media advocacy exposures and as a basis for model building. This also could be used to assess the relationship of newspaper coverage and policy outcomes.

Group 2: outcomes
To document the outcomes of tobacco control initiatives, public health practitioners and researchers need milestones and proximal markers of success; therefore, evaluation data need to be set up to capture sufficient information to document the short-term outcomes (for example, policies), intermediate outcomes (for example, behaviours and attitudes), and long-term outcomes (for example, health outcomes).

Challenges
Participants grappled with defining the critical short, intermediate, and long-term outcomes. For example, are there “good” data linking clean indoor air policy, smoking bans at home, confirmed awareness of media campaigns, or insurance coverage of nicotine replacement therapy to “key” outcomes? Additionally, participants discussed the lack of information documenting the implementation or “process outcomes” of programmes—for example, recording which activities take place, where they take place, who conducts the activities, as well as the quality, fidelity, and rate of implementation.

Recommendations
The participants offered the following recommendations:
• Convene a meeting of stakeholders (that is, administrators, programme implementers, health promotion activists, participants, and scientific evaluators) to draft a national surveillance plan and decide what are the “key” outcome categories; define what to measure and how to measure it; link key intermediate indicators to final outcomes.
• As part of a nationwide surveillance plan, states should support qualitative studies to capture the readiness of communities and what is happening “on the front lines” in terms of social, cultural, and political environments so that policymakers can compare statewide and nationwide why certain interventions are working in certain states and communities. These qualitative studies can help account for variability in implementation of programmes as well as capturing data on the process of programme implementation.
• Once surveillance is standardised, develop a state report card or ranking system to compare state tobacco control activities.
• Define standard morbidity outcomes (for example, health outcomes, toxic exposure measures, hospital records) to monitor changes better over time. To accomplish this, a new surveillance system is needed to evaluate biomarkers. Specifically, the Centers for Disease Control and Prevention should develop a state-specific health and nutrition examination survey. Components should include tests to measure tobacco-specific carcinogen exposures (for example, spot urine and blood samples), an exposure questionnaire, and a scanning device to capture the actual brand smoked and bar code on cigarette packs and chewing tobacco tins.

Group 3: methods
In order for programme administrators at the local, state, and national levels to know what is the optimal mix of interventions and funding to reduce the burden of tobacco use, we need systematic data collection across states and local communities and appropriate methods to analyse these data. With limited number of observations, assessing small effects of a particular programme may be difficult in the presence of far stronger secular trends.

Challenges
Participants grappled with numerous issues related to determining the appropriate statistical methodologies and models to use for hierarchical and time series data. The lack of cooperation to collect data systematically was seen as a crucial obstacle to providing a macro-level analysis of tobacco control interventions. It is understood that current quantitative methods are limited and there is a need to integrate better qualitative and quantitative ideologies.

The group discussed how research designs often fail to match the needs of the real world. Participants recognised the need for models that are useful to practitioners (that is, state, county, community, city evaluators). Studies, such as those conducted at the University of Michigan, which are linking academia with the community through participatory research, build community capacity and gather information that is relevant to researchers, practitioners, and recipients of the intervention(s) were discussed as possible models for future tobacco control interventions.

The methods group debated different methods for analysing the smoking prevalence and the time series consumption data, similar to the type of data used in the ASSIST evaluation. There was considerable disagreement about how to model the state and time effects, and such details as whether to treat the state effects as fixed or random. This issue of random versus fixed effects has implications for the generalisability of the estimated intervention effects and the power of the statistical inferences. Overall, these modelling issues were viewed as needing more attention and could not be resolved at this workshop.

The methods group also discussed the challenge of simulation models and how to best use simulation to show policymakers the potential impact of particular policies in given certain social and political climates. “SimSmoke” was...
presented as an example of a dynamic simulation model. “SimSmoke” models the effects of different policy approaches to tobacco control, including the effect of price interventions (taxes), policies to reduce youth access to cigarettes, policies to encourage cessation treatment, mass media policies, and clean air laws on smoking rates and mortality across sex, age, and racial/ethnic groups. These simulations can show the effects of a particular dose of an intervention (for example, size of tax, expenditures on media, extent of clean air laws) at a given point in time. However, there was concern that simulation models, while useful, can be misleading if all assumptions inherent in their construction are not made explicit and if valid standard errors are not presented in conjunction with projections and estimates made from the simulation models. These standard errors need to incorporate the sampling variability of any external estimates used in the simulation and include variability for assumed parameters based on agreed upon prior distributions.

**Recommendations**

**Quantitative methods**

To identify the relative contribution of individual programme components by assessing the variation in intensity of tobacco control activity across time and geographic units (for example, county, state), the following approaches should be considered:

- hierarchical models that appropriately nest individuals within the catchment area of the intervention
- social network analysis to understand diffusion
- simulation models to synthesise evaluation research and provide direction for programme development
- randomised trials to complement comprehensive programme evaluations
- online reporting to assess individual contributions or programme components
- geographic identifiers to track exposures across time, populations, and media markets
- national surveys, cohorts, and databases to supplement surveillance; pilot studies of such a surveillance system should be conducted in states where considerable data are already available (for example, California and Florida).

**Qualitative methods**

To understand the timing and active ingredients of tobacco control interventions, the following complementary qualitative methods were proposed:

- case studies of programme implementation
- focus groups
- ethnographic studies
- consensus conferences of experts who use non-linear models, physicists, neural network experts, statisticians, and epidemiologists, and practitioners to develop new systems approaches to evaluation to capture the extent, effectiveness, and efficacy of individual level and population level interventions.

**CONCLUSIONS**

As one of the first workshops to address the overall challenge of evaluating tobacco control programmes, it is not surprising that the assembled participants made a number of basic recommendations, including further meetings. Across all three working groups, participants agreed that an ongoing process was needed to develop data resources and methods for more effective evaluation. This need is immediate, given the current scope of funding and intervention and the expectations for success. Leadership for furthering evaluation approaches needs to be assumed by appropriate agencies and foundations.

Workshop participants were certain and unanimous that the current state of evaluation research, including development of measures and methods, must be improved. There is a need to identify empirically based theoretical principles that can guide future interventions. Appropriate methods and outcome measures for evaluation of higher level public health initiatives, including tobacco control programmes, need to be developed. Rigorous experimental control and random assignment may not be possible. Hierarchical or multilevel modelling approaches are promising for further research. If evaluation approaches are not improved, the potential consequences include misleading and inaccurate estimation of programme impact, the promotion of less effective strategies, and loss of financial and political support. Investing in evaluation research can lead to more effective tobacco control programmes. Developing strong and effective tobacco control programmes will curb the huge loss of life and economic burden caused by continued use of tobacco.

The utility of the workshop's recommendations goes beyond tobacco control evaluation and surveillance in the USA and has application for global tobacco control. There is a critical need to be able to monitor and assess tobacco related diseases, as well as assess efforts being developed to fight the epidemic worldwide. The Tobacco Control Country Profiles and the NATIONS electronic data system have identified a variety of indicators to monitor tobacco control within nations: smoking prevalence and consumption, laws and regulations, morbidity and mortality, industrial organisations, tobacco economics, and programme tobacco control interventions. Data assessing countries’ resources and infrastructure to deliver interventions are also needed. Developing conceptual models of what works and determining a core set of indicators to evaluate country efforts is essential. These indicators should also highlight political and cultural factors that impede development and adoption of tobacco control policies and practices.

**ACKNOWLEDGEMENTS**

This article gives the proceedings of the workshop. There was no attempt to achieve group consensus on all issues; consequently, the article should not be necessarily construed as reflecting the views of all participants. GlaxoSmithKline provided funding for this workshop through a grant to the Institute for Global Tobacco Control at the Johns Hopkins Bloomberg School of Public Health. The views presented in this article are not necessarily those of the funding organisation. For more information please contact Dr Frances Stillman, Co-Director, Institute for Global Tobacco Control at the Johns Hopkins Bloomberg School of Public Health, tel. (001) 410 995 3435.

**APPENDIX**

**List of participants**

- Ronald Borland—VicHealth Centre for Tobacco Control
- David Burns—University of California at San Diego School of Medicine
- Marion Ceraso—Johns Hopkins Bloomberg School of Public Health
- Michael Cummings—Roswell Park Cancer Institute
- Francesca Dominici—Johns Hopkins Bloomberg School of Public Health
- Matthew Farrelly—Research Triangle Park
- Elizabeth H Gilpin—University of California at San Diego
- LD Graubard—National Cancer Institute
- Anne Hartman—National Cancer Institute
- Dorothy Hatuskami—University of Minnesota
- Lyndon Haviland—American Legacy Foundation
- Robert Hornik—Annenberg School of Communication
- David Levy—Pacific Institute for Research and Evaluation
- David Mendez—University of Michigan
- Giovanni Parmigiani—Johns Hopkins School of Medicine
Terry Pechacek—Centers for Disease Control and Prevention/Office on Smoking and Health
John Pierce*—University of California at San Diego
Everett Rogers—University of New Mexico
Jonathan Samet*—Johns Hopkins Bloomberg School of Public Health
Stephanie Smith*—Johns Hopkins Bloomberg School of Public Health
Frances Stillman*—National Cancer Institute (now at the Institute for Global Tobacco Control)
William Trochim—Cornell University
Kenneth Warner*—University of Michigan
Gonghuan Yang—Chinese Academy of Preventive Medicine
*Steering Committee Member

Steering Committee members unable to attend
Frank Chaloupka—University of Illinois at Chicago
Gary Giovino—Roswell Park Cancer Institute
Scott Zeger—Johns Hopkins Bloomberg School of Public Health

Staff
Shannon Clubb—Johns Hopkins Bloomberg School of Public Health
Roberta Gray—Johns Hopkins Bloomberg School of Public Health

REFERENCES