Challenges of Evaluating Complex Policy Interventions

Robert Schwartz, Shawn O’Connor, Bo Zhang
Tel Aviv University
Tel Aviv, Israel
December 17th, 2013
OUTLINE

- Evaluand – comprehensive & complex strategies
- Evaluation Challenges
- Solutions in *Evaluating the Complex*
- Inadequacy of those solutions
- System dynamics modeling
- Combining SimSmoke & IPCA
- Learning from comparison – the MPOWER study
Complexity Puzzle

Mixes

Interactions

Synergies

Loops

Sub-populations

Multiple locations

Multiple players
Non-linearity (Uphoff, 1992)

- The distance between causes and effects can be long in time, and sometimes short,
- There are usually several causes for any change that occurs
- There is no proportional relationship between the size of causes and effects
- Change occurs through qualitative leaps and bounds, and is not incremental and proportional
Feedback Loops

Feedback loops can be positive (amplifying deviation from an equilibrium state) or negative (reducing deviation, bringing the system back towards the original equilibrium state)
Synergy

- CDC, Health Canada, IOM, others promote comprehensive strategies and the expectations for synergy

- Empirical evidence is sparse

- Modeling – multiply coefficients?
Why not evaluate complex strategies

- **Evaluability Assessment:**
  - Theory of change? Likelihood of activities leading to desired outcomes
  - Measurability: Can change be measured and attributed to the intervention(s)?
  - Intended user: people with desire and authority to make needed changes as identified in the evaluation
Why Evaluate Complex Strategies?

1. Accountability Reaction

2. Results-based Management

3. Evidence informed Policymaking
Why Evaluate Complex Strategies 2

- Inform policy choices about allocation of resources and interventions:
  - How much of what, when?
  - What to do with more resources?
  - What to cut when resources decrease?
  - Identify intervention gaps for general population and for sub-population groups (see next slide)
CSE Challenges

- Program evaluators generally trained in evaluating single program interventions
- Determining population level outcomes (paucity of good data)
- Obtaining data on inputs, activities and outputs
- Daunting task of obtaining quality evaluative information about a multitude of interventions – i.e. determining program effectiveness at the micro-level
And the biggest challenges of all?

Attributing population level outcomes to micro-level interventions

Determining the relative contribution of each micro-level intervention to changes in macro-level outcomes – especially when interactions and synergies are expected
Critique of classic approaches

- Black box focus on final outcomes
- Lack of attention to synergies
  - What mixes?
  - In what sequence?
- Lack of attention to feedback loops
- Lack of attention to multiplier effects
  - Little evidence to inform strategy development
EVALUATING THE COMPLEX SOLUTIONS

Evaluating the Complex
Attribution, Contribution, and Beyond
Comparative Policy Evaluation
Volume 18

Kim Forss, Mita Marra, and Robert Schwartz, editors
Thematic Evaluation

Identify themes which are relevant to policy and evaluate the diverse interventions and programs against these themes.

Look across multiple interventions and often across multiple regions and countries.

Define ‘themes’ in terms of sectors or policy areas (e.g. agriculture, education, environment) OR in terms of cross-cutting (or horizontal) issues (e.g. gender, sustainable development)
Complex Evaluation Strategy (Wilkins)

- Performance measurement system
- Strategic investment of program evaluation resources
- Evaluation support and quality assurance
- Macro-level surveillance of outcomes using key indicators
Contribution Analysis for Complex Strategies (Mayne)

- **Step 1**: Set out the cause-effect issue to be addressed
- **Step 2**: Develop the postulated theory of change and risks to it
- **Step 3**: Gather the existing evidence on the theory of change
- **Step 4**: Assemble and assess the contribution story, and challenges to it
- **Step 5**: Seek out additional evidence
Step 6: Revise and strengthen the contribution story

Step 7: In complex settings, assemble and assess the complex contribution story

Develop the contribution story for each additional sub-theory of change

- (using Steps 2–6)

- Develop the contribution story for any general theory of change (using Steps 2–6)
Quantified Intervention – Path Logic Models (*Toulemonde*)

An innovative technique for helping understand if and how complex strategies are meeting their goals.

Helps ‘tell the story’ to policymakers

Useful for identifying needs for further evaluative information
EC supported activities and their most direct effects

1. Conducive development environment
2. Economic activities where the poor engage
3. Reduced vulnerability
4. Equitable access to basic services
5. Expanded services whilst protecting quality

Poverty reduced
Quantified Intervention – Path Logic Models (*Toulemonde*) (cont’d)

- 5 key paths toward poverty reduction
- 200 + components gathered into 9 clusters
- 9 specific components selected, each being typical of a cluster
- Impact assessed for each component
- Profile of each given component assessed
  (assumption that the overall potential impact on poverty is 100)
- Amount sliced between the 5 paths
1 – conducive development environment

2 – growth of economic activities where the poor engage

3 – reduced vulnerability

4 – equitable access to basic services

5 – expanded services whilst protecting quality

Government

EC
Intervention Path Contribution Analysis (IPCA)
# PREVENTION LOGIC MODEL

**Strategy goal:** To prevent smoking initiation and regular use among Ontario’s children, youth, & young adults in order to eliminate tobacco-related illness and death.

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Interventions</th>
<th>Expected Contribution</th>
<th>Known Contribution</th>
<th>Paths</th>
<th>Lt Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership, Coordination, Collaboration</td>
<td>Youth Action Alliances</td>
<td></td>
<td></td>
<td>Availability</td>
<td>Reduce Initiation of tobacco use among children, youth &amp; young adults</td>
</tr>
<tr>
<td></td>
<td>Lungs are for Life</td>
<td></td>
<td></td>
<td>Visibility / Marketing</td>
<td></td>
</tr>
<tr>
<td>Capacity Building</td>
<td>High School Grants</td>
<td></td>
<td></td>
<td>Price</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leave the Pack Behind</td>
<td></td>
<td></td>
<td>Social Climate</td>
<td></td>
</tr>
<tr>
<td>Technical Assistance</td>
<td>Youth Access Restrictions</td>
<td></td>
<td></td>
<td>Knowledge / Awareness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stupid.ca</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research, Evaluation, Monitoring, Knowledge Exchange</td>
<td>Public Education/Media Campaigns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Youth Vortal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Taxes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ontario Tobacco-free Network</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advocates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advertising &amp; POS Restrictions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advertising Bans</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ontario Tobacco Research Unit
### PREVENTION LOGIC MODEL

**Strategy goal:** To prevent smoking initiation and regular use among Ontario's children, youth, and young adults in order to eliminate tobacco-related illness and death.

#### Interventions
- **Youth Action Alliances**
- **Lungs are for Life**
- **High School Grants**
- **Leave the Pack Behind**
- **Youth Access Restrictions**
- **Stupid.ca**
- **Public Education/Media Campaigns**
- **Youth Vortal**
- **Taxes**
- **Ontario Tobacco-free Network**
- **Advocates**
- **Advertising & POS Restrictions**
- **Advertising Bans**

#### Expected Contribution
- Large
- Medium
- Small
- Unknown

#### Known Contribution
- Large
- Unknown

#### Paths
- Social Climate

#### Lt Outcomes
- Reduce Initiation of tobacco use among children, youth, and young adults
- Reduce tobacco use prevalence among children, youth, and young adults

---

**THE ONTARIO TOBACCO RESEARCH UNIT**

**UNITE DE RECHERCHE SUR LE TABAC DE L'ONTARIO**

**Ontario Tobacco Research Unit**
PREVENTION LOGIC MODEL

To prevent smoking initiation and regular use among Ontario’s children, youth, & young adults in order to eliminate tobacco-related illness and death

Interventions | Expected Contribution | Known Contribution | Paths | Long-Term
---|---|---|---|---
Youth Action Alliances | ? | ? | Large | Reduce Initiation of tobacco use among children, youth & young adults
 | Small | Small | Medium | Reduce tobacco use prevalence among children, youth, and young adults
 | Large | | Some |
Ontario Tobacco Research Unit

Organization Learning

Statement of Purpose / Problem:
Complex policy initiatives challenge evaluators in new and daunting ways, beyond the scope of what existing tools of the trade can manage. Two distinguishing characteristics set evaluation of complex policy initiatives apart:
1) Attribution;
2) Collaboration.

Complex strategy initiatives would almost always fail standard evaluability assessment. Their objectives are seldom clear-cut and measurable. And there is little evidence to support a causal linkage, between unique components (programs and projects) or mixes of these and expected outcomes. Even when macro-level outcomes are measurable, it is very difficult to attribute changes in outcomes to the implementation of the strategy.

Context:
The Smoke-Free Ontario Strategy is a complex construct that includes several province-wide policy initiatives and numerous program interventions delivered through 36 Public Health Units and 20 non-governmental organizations. The primary evaluation clients are interested in learning about the overall effectiveness of the strategy in meeting such outcome objectives as decreased cigarette consumption and smoking prevalence. Moreover, the clients want to understand which aspects of the Strategy are more successful in what contexts.

Some Current & Imminent Work
- Youth programs – Youth Action Alliances, Youth sub-strategy
- Enforcement – Smoke Free Ontario Act implementation
- Cessation – sub-strategy, supply versus needs, innovative demonstration projects

Approach:
Learning orientation
Participatory
Realistic:
- Contexts
- Mechanisms
- Outcomes

Implications:
A mix of evaluation approaches is needed to provide stakeholders with knowledge for learning about needed changes in order to maximize performance. Approaches range from micro-level ongoing performance measurement to macro-level surveillance and include innovative mezzo-level assessments of sub-strategies and strategy components. Central coordination of Smoke-Free Ontario Strategy Evaluation efforts using a collaborative learning approach shows promising signs of having positive impact on learning for better practice and policy. This approach provides a useful model for other strategies in Public Health and other areas.

Acknowledgements:
The Ontario Tobacco Research Unit receives funding from the Ontario Ministry of Health Promotion and the Ontario Ministry of Health and Long-Term Care.

Evaluating the Smoke-Free Ontario Strategy

Micro
Performance Measur.
OTRU Evaluation
Local Evaluation
Quality Assurance

Mezzo
Sub-Strategy Eval
Strategy Comp. Eval
Cluster Eval
Multi-level Eval.

Macro
Surveillance
Population Surveys
OTS longitudinal survey

Knowledge Exchange
Policymakers, Practitioners, Advocates, Media,
General Public, Research Community
LIGHT CA

Complex Contribution story assembled and assessed by key actors

BUT

So complex that nobody really knows
IPCA Challenges

Lack of evidence on expected effects of each intervention on each path

It is not possible to measure actual effects of each intervention over the course of any given year

Doesn’t actually measure synergies, feedback loops, jump points.

Complex contribution story too hard to figure out and verify
System Dynamics Models

- A methodology and mathematical modeling technique to understanding the behavior of complex systems over time
- Addresses internal feedback loops and time delays that affect the behavior of the entire system
- Models non-linear relationships
<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Looking at the dynamic interrelationships of whole system, rather than an isolated part</td>
<td>Very complex model; need data from many different sources</td>
</tr>
<tr>
<td>Improving understanding of system behavior over time (esp. long-term); gaining insights more important than numerical prediction</td>
<td>Generally not able to control for individual-level covariates (though age group and sex taken into account)</td>
</tr>
<tr>
<td>Taking into account dynamic population structure (e.g., population moving over time by age and sex)</td>
<td>Majority models: no consideration of immigration, SES, tobacco industry influence, no parameter estimates (e.g., ORs, RRs) or variance</td>
</tr>
<tr>
<td>Dealing with feedback loops (e.g., former smokers → current smokers via relapse)</td>
<td>No objective or clear instructions on model calibration</td>
</tr>
<tr>
<td>Suitable for non-linear relationship</td>
<td>Suitable for large areas (state or country) but not for smaller areas (community, hospitals, or schools)</td>
</tr>
<tr>
<td>Simulating policy impact even if numerical data not available</td>
<td>Simulating the potential impact, rather than deliver the “truth”</td>
</tr>
</tbody>
</table>
SimSmoke Model
(Levy, et al., 2000-2012: over 30 publications)

- **Purpose**: to *evaluate & predict* impact of tobacco policies
- **Policy evaluated**: tax, clean air laws, mass media policies, advertising & marketing bans, health warnings, youth access, & cessation treatment; *individual policy at a time or all policies at a time*
- **Outcome**: smoking prevalence & smoking-related deaths
- **Taking into account**: dynamic population structure (age, gender, births, deaths, & smoking status)
- **Ignoring impact of**: migration, SES, SHS & tobacco industry
- **Software**: Excel Spreadsheet & C++ macros
- **Applied in**: over 30 countries, incl. Asian, European, South & North American countries, & 6 states in US
SimSmoke: Basic Structure of Model

- Population model begins with initial year population (by age and gender) and moves through time (by year) with births and deaths (1st order discrete Markov process)

- Smoking model distinguishes population in never smokers, smokers, and ex-smokers and moves through time with initiation, cessation and relapse (Markov)

- Smoking-attributable deaths depend on smoking rates and RRs from CPS-II, in total and by cause

- Policy modules- one for each policy with interdependent effects on smoking rates
SimSmoke: The Impact of Specific Policies

Role of Individual Policies in Reducing Smoking Prevalence

Male smoking prevalence reduction from 1991 (60%) to 2006 (42%) in Thailand
SimSmoke: Projection of Smoking Prevalence by Different Policies

Male Smoking Prevalence, Netherlands 2010 – 2040

(Nagelhout et al., 2012)
SimSmoke: Challenges

- Keep sample methods and definitions consistent for surveys on smoking prevalence
- Get better data to examine initiation and cessation behaviors
- Develop objective and clear instructions on model validation & calibration
- Get better estimates on synergistic effects among policies (currently assuming multiplicative)
- To incorporate the impact of SES & migration
- To develop modules for other policies (e.g., point-of-sale display ban)
- To consider influence of tobacco industry, SHS, & contraband
General Sticking points

- Assigning value to intensity of policy intervention
- Choosing coefficients – do American estimates apply to Canada?
- Sensitivity analysis needed as state of evidence for effects of most policies is weak and effect sizes vary widely amongst studies (see Dutch study).
- Synergistic effects not evidence based

- Tax Policies – 25% above or below
- Smoke-free Policies – 25% above or below
- Media Campaigns – 50% above or below
- Advertising Bans – 50% above or below
- Health Warnings – 100%
- Cessation Treatment Policies – 100%
What is next?

IPCA

SIMSMOKE

MPOWER