CANADIAN TOBACCO CONTROL RESEARCH INITIATIVE

CTCR

Identifying Outcome Indicators and Measurement Items for Evaluating Youth Smoking Cessation Interventions:

Background Paper

by
Tobin T. Copley, MA
Chris Y. Lovato, PhD
Steve Manske, EdD

For the Office of Programs and Mass Media, Tobacco Control Programme, Health Canada
In fulfillment of Final Report deliverable, April 30, 2005

To order free copies of this publication, contact:

Canadian Tobacco Control Research Initiative (CTCRI)
10 Alcorn Avenue, Suite 200
Toronto, Ontario M4V 3B1
T: 416.934.5968
F: 416.961.4189
E: info@ctcri.ca

This publication is available online at:

http://www.ctcri.ca/

The material in this publication (with the exception of material in Appendix D) may be copied or reproduced without permission; however, the following citation must be used:

Copley TT, Lovato CY, Manske S (2006). ***Indicators and measurement items for outcome evaluation of youth smoking cessation interventions: Background paper.*** Toronto: Report from the Canadian Tobacco Control Research Initiative to the Office of Programs and Mass Media, Tobacco Control Programme, Health Canada.

© Canadian Tobacco Control Research Initiative (2006) | Printed 2006 | 001

Table of Contents

Acknowledgements i

Executive Summary ii

1 BACKGROUND 1

1.1 Purpose of the report 1
1.2 Anticipated outcomes 1
1.3 Anticipated use 1

2 HISTORY OF THE PROJECT 2

3 SCOPE OF THE PROJECT 2

4 APPROACH 3

4.1 Inclusion and exclusion criteria 3
4.2 Development of the logic model 4
4.3 Review of measurement items 7
4.4 Identification of indicators and measurement items 7
4.5 Identification of measurement issues 8
4.6 Development of recommendations 9

5 FINDINGS 9

5.1 Interventions/surveillance program questionnaires 10
5.2 Indicators and items used to evaluate interventions 12

6 MEASUREMENT ISSUES 14

6.1 Are the items we are using meaningful to youth? 16
6.2 What information is available regarding item reliability and validity? 16
6.3 Should biochemical validation be recommended for both real- world and research evaluation studies? 17
6.4 How should we measure quitting in youth? 17
6.5 How should we measure smoking reduction in youth? 18
6.6 How should we measure abstinence in youth? 19
6.7 How is the concept of slips relevant in youth smoking cessation? 19
6.8 What are self-report error issues that should be considered in measuring youth smoking cessation outcomes? 20
6.9 What are the ideal post-intervention follow-up intervals for use in real-world and research studies on youth smoking cessation? 21
6.10 Which approach should be considered for real- world and research studies: intent to treat or attrition analysis? 22

7 RECOMMENDATIONS 23

7.1 Identify core indicators 23
7.2 Identify items for evaluation 23
7.3 Recommendations related to outcome measurement for consideration by the working group 27
7.4 Questions to be addressed in the next steps 28

References 29

Appendices

Appendix A: Glossary of terms 35
Appendix B: Questionnaire content overview: Intervention evaluation..... 37
Appendix C: Questionnaire content overview: Surveillance surveys..... 40
Appendix D: Copies of reviewed intervention evaluation questionnaires..... 43
Appendix E: Survey references 161
Appendix F: Literature review process..... 163
Appendix G: List of reviewed articles..... 164

List of Figures and Tables

Figure 1: Logic Model for the outcome evaluation of behaviour change related to a youth smoking cessation intervention..... 6

Table 1: Summary of intervention programs using reviewed evaluation questionnaires. 11
Table 2: Summary of selected national population surveillance surveys..... 12
Table 3: Definitions related to quitting..... 18
Table 4: Follow-up interval recommendations..... 22
Table 5: Validated core items recommended for consideration by the working group.... 24
Table 6: Items showing promise needing more testing, and suggested modifications..... 25

The CTCRI

The Canadian Tobacco Control Research Initiative (CTCRI) was initiated in 1997 to increase research relevant to advancing program and policy agendas, thereby reducing the prevalence of tobacco use; to stimulate more research on tobacco; and to draw more Canadian investigators into tobacco research. The CTCRI is a partnership of the National Cancer Institute of Canada (NCIC), the Canadian Cancer Society (CCS), the Canadian Institutes of Health Research (CIHR) and Health Canada.

Acknowledgements

Many thanks for the assistance, expert advice, and support we have received from the project advisory committee:

Julie Green	Health Canada
Murray Kaiserman	Health Canada
Oonagh Maley	CTCRI (Project Manager)
Judy Snider	Health Canada
Tom Stephens	Thomas Stephens & Associates

We gratefully acknowledge the assistance and expert advice of the following people in identifying the latest information available on the topic addressed in this report and providing input regarding key issues:

Cathy Backinger	US National Cancer Institute
Linda Bailey	Center for Tobacco Cessation; North American Quitline Consortium
Dianne Barker	Barker Bi-Coastal Health, National Youth Smoking Cessation Survey
Bill Blatt	American Lung Association, N-O-T program
Sharon Campbell	Centre for Behavioral Research & Program Evaluation, University of Waterloo; North American Quitline Consortium
Corinne Husten	US Centers for Disease Control and Prevention
Anne Malarcher	US Centers for Disease Control and Prevention
Paul McDonald	University of Waterloo
Gloria Meyer	Center for Tobacco Research & Intervention, Madison, Wisconsin
Jennifer O'Loughlin	McGill University
C. Tracy Orleans	Robert Wood Johnson Foundation
Tammy Sims	Center for Tobacco Research & Intervention, Madison, Wisconsin
Amy K. Sporer	University of Illinois at Chicago; Helping Young Smokers Quit
Steve Sussman	University of Southern California

We also wish to thank the Youth Tobacco Cessation Collaborative for their earlier work in this area.

The authors acknowledge the cooperation of the questionnaire/survey developers for permitting the inclusion of the survey items in this report.

We wish to acknowledge Lia De Pauw, research assistant, for her assistance retrieving information, summarizing articles, and doing the background work necessary to create the final product. Thanks to Oonagh Maley who provided assistance in managing the project and coordinating materials during the initial development and earlier stages of this project. Emily J. Alfred, CTCRI administrative assistant, coordinated the final version of the report, provided assistance in layout and general proof reading.

The interpretation and opinions expressed in this publication are the responsibility of the CTCRI and the Working Group.

Executive Summary

BACKGROUND

Current research reveals a lack of evaluation data regarding youth smoking cessation interventions that can help decision-makers and practitioners arrive at informed decisions regarding better practices. The information contained in this document is meant to provide a foundation that anchors future discussions regarding selection of indicators and measures for evaluating the outcomes of smoking cessation interventions for the youth population. In particular, it is meant to serve as a resource document for an expert workshop on youth cessation indicators that will be funded by Health Canada in November 2006. Because this document is oriented towards developments in research and measurement and it is likely to evolve on the basis of deliberations at the expert workshop, in its current form it will not meet the needs of community evaluators or program professionals. A more applied and "user-friendly" document providing guidance on the selection of specific measures will be produced in later phases of this project.

The five "building blocks" provided in this background paper include:

1. A logic model (conceptual framework) illustrating a heuristic for youth cessation that includes a set of definitions of outcome behaviours related to smoking cessation. From this model, behaviours and associated indicators for measurement are identified.
2. A review of measurement items used in intervention and surveillance programs, based on a scan of available measures of tobacco cessation, focusing on those measures related to the evaluation of youth tobacco cessation programs. This is based on a review of published scientific literature, as well as information gathered from key individuals from Canada and the US who were identified as national experts.
3. Identification of a recommended set of indicators and measurement items for evaluation taken from the published and unpublished literature for evaluating youth smoking cessation programs including: (1) validated core items, (2) identification of items that show promise but need further testing, (3) identification of a small number of non-core (optional) items, (4) specification of gaps in the list of items available.
4. Identification of a set of key measurement issues summarizing the relative strengths and limitations of identified measures.
5. A set of recommendations and key questions that can be used to develop consensus regarding a core set of items for evaluation studies and surveillance, as well as key measurement issues where additional work is needed.

APPROACH

The following inclusion/exclusion criteria frame the parameters of this review:

1. *Type of measure:* Only those measures suitable for self-report survey administration are considered.
2. *Population:* "Youth" are defined as individuals between 12 and 18 years of age.[1]

[1] The age range was decided upon by the advisory group.

3. *Type of tobacco consumed:* The review focuses exclusively on cessation of smoking cigarettes. Smoking other forms of tobacco (cigars, pipes, etc.) and non-combustion use (chew, snuff, etc.) is not considered.[2]
4. *Smoking history:* The standardized but arbitrary measure of having smoked 100 or more lifetime cigarettes was used to define a smoker, thereby excluding individuals termed puffers and experimenters.[3]
5. *Smoking pattern:* Due to the wide variety of patterns of youth smoking, the only restriction we have considered is that youth must have smoked at least 100 cigarettes. We note that it is not uncommon for adult cessation programs to be limited to daily smokers.
6. *Geographic scope:* The study was restricted to measures used in Canada and the United States.

The literature on youth smoking cessation is a relatively new field, so the number of reports published are, at present, relatively few but rapidly growing in number. Recognizing this, the literature review conducted for this report builds on the foundation of work represented by key reviews on youth smoking cessation measurement such as those by Sussman et al. (1999), Mermelstein et al. (2002), Sussman (2002), Garrison et al. (2003), and O'Loughlin et al. (2003), and the articles contained within the American Journal of Health Behavior's 2003 special issue on youth smoking cessation. We undertook a search of the 1966-present Medline database using terms such as adolescent, youth, tobacco, smoking, cessation, indicator, measurement, and outcome, as both keywords and MEDLARS descriptors. A total of 102 articles were retrieved and reviewed.

A select group of expert researchers provided suggestions regarding key literature and measures that should be included in this review. Among these individuals we found there was a high level of consensus in the comments made regarding existing gaps. We identified and reviewed key articles from the published literature and relevant work in the "grey" literature (unpublished, disseminated through unconventional means such as internet sites, or works in progress).

To identify the major evaluation questionnaires used in intervention programs in Canada, we drew on the findings of a recent review by O'Loughlin et al. (2003) on the evaluation of youth smoking cessation programs in Canada, and supplemented O'Loughlin's findings through consultation with experts at Health Canada and others in the tobacco control community, in addition to identifying instruments used or referenced in the published and grey literature.

We identified 27 intervention programs or surveillance surveys containing measures appropriate for the measurement of youth smoking cessation. From this set of surveys, we selected 18 intervention programs or surveillance surveys representing 34 individual questionnaires for detailed review based on the following inclusion criteria:

1. *Currency:* The questionnaire is currently used in evaluating youth smoking cessation interventions.

---

[2] We also do not consider the possible substitution of cigarettes for another form of tobacco when individuals are attempting to reduce / quit smoking cigarettes (e.g., shifting from cigarettes to smokeless tobacco).

[3] While this may exclude some very low-dosage youth smokers who may arguably have developed an addiction to smoking with less than 100 cigarettes, given that the focus of this paper is on smoking cessation we feel it is reasonable to use this definition for minimum lifetime consumption. Whether youth can accurately estimate their consumption of 100 cigarettes is the subject of debate in the literature, and is a discussion we do not engage in here.

2. *Widespread use:* The questionnaire is currently used nationally or at least at the provincial/state level.
3. *Identified by an expert:* The questionnaire (or program) was identified favourably in either a recent review of programs or in personal communication with recognized experts in the field.
4. *Cited in the literature:* The questionnaire is referenced in recently published articles by more than one author or group.
5. *Face Validity:* The questionnaire "makes sense" in terms of constructs we are trying to measure.
6. *Canadian context:* Questionnaires used in Canada were given priority.
7. *English language:* French language questionnaires were excluded from this review.

The information gathered from the literature, scans of questionnaires, and discussions with experts were summarized into descriptive tables that were used to conceptually synthesize findings. With respect to surveys, we conducted a review of what measures exist, compiled these measures, and selected from the best among them. Steve Sussman also offered to review all studies from his 2002 paper that were published from 1996 forward, documenting which studies identified the origin of questionnaire measures used, and which papers reported on validity and reliability measures for the measures used.

We developed a set of recommendations by using the information contained in the summary and descriptive tables describing the content of the reviewed questionnaires. This information was combined with findings of the literature reviewed, discussions with experts, and input from the project advisory group. A core set of items was identified for which there appears to be general agreement regarding utility and validity. Existing gaps in items and recommendations regarding ways in which items could be improved were also identified. In the course of synthesizing information from different sources, a set of questions was also developed that should be addressed by the working group who will take this project to its next phase.

FINDINGS

In general, the literature on methodological issues in the evaluation of smoking cessation interventions for youths is relatively recent and is not well developed; researchers also continue to grapple with fundamental issues such as definitions and validity of measurements for youth. In contrast, the literature on the evaluation of smoking interventions for adults is well developed and mature. The published literature acknowledges that there are many problems associated with attempting to develop survey measures to characterize youth smoking behaviour, particularly with regard to cessation. This was underscored in our discussion with experts in the field and reaffirmed in our review of documents. The existing measures are not ideal; however, they still provide useful information in situations where decisions must be made regarding better practices in youth smoking cessation.

Indicators and items used to evaluate interventions. In the intervention program questionnaires reviewed, indicators of smoking behaviour change include: I) Abstinence (cessation or zero consumption since intervention), II) Reduction (consuming fewer cigarettes over a period of time than at baseline), or III) Quit attempts (actual attempts made to quit smoking). These three indicators are also widely used in surveillance.

Measurement Issues. From the literature review process and subsequent synthesis emerged a set of measurement questions. If the question is discussed substantially within the literature or among the experts we consulted, the question is addressed in the full document. Some of these questions are very

broad issues that represent gaps in the current knowledge or require extensive research to address. Following are ten key questions about measurement for which there is some discussion and guidance in the literature:

1. Are the items we are using meaningful to youth?
2. What information is available regarding item reliability and validity?
3. Should biochemical validation be recommended for both real-world and research evaluation studies?
4. How should we measure quitting in youth?
5. How should we measure smoking reduction in youth?
6. How should we measure abstinence in youth?
7. How is the concept of slips relevant in youth smoking cessation?
8. What are self-report error issues that should be considered in measuring youth smoking cessation outcomes?
9. What are the ideal post-intervention follow-up intervals for use in real-world and research studies on youth smoking cessation?
10. Which approach should be considered for real-world and research studies: intent to treat, or attrition analysis?

When considering these questions, it is important to keep in mind some of the challenges related to youth smoking that can make the measurement of cessation-related behaviours particularly challenging (these challenges primarily relate to the patterns of smoking among youth).

RECOMMENDATIONS

Core indicators. We recommend that the working group adopt the following three indicators to evaluate behaviour change in youth smoking cessation interventions:

I. *Abstinence* (cessation or zero consumption since intervention),

II. *Reduction* (consuming fewer cigarettes per unit time than at baseline), or

III. *Quit attempts* (actual attempts made to quit smoking)

Validated core items. Amongst validated items available, six items from the Helping Youth Smokers Quit project are recommended as part of a set of core measurement items for the evaluation of youth smoking cessation interventions. Taken together, these items permit the estimation of cigarette consumption level, both over the last 7 and 30 days. Determination of daily or non-daily smoking status can be made using either seven-day or 30-day time periods. The items selected will also permit description of patterns of smoking over the course of the previous seven days, abstinence over the previous 7 or 30 days, continuous abstinence, and the number and duration of failed quit attempts.

Items that show promise and/or need more testing and suggested modifications to existing items. We did not identify any items requiring more testing that differ substantially from the recommended core items above; however, four items have been identified that may be able to be improved upon through modification. The items address 30-day consumption, quit attempts, and longest quit attempt.

Non-core (optional) items. No suitable optional items for the evaluation of smoking behaviour change in youth smoking cessation programs were identified, other than those already described.

Gaps to address in current items. There are a number of significant gaps relevant to the measurement of youth smoking cessation intervention outcomes. For our purposes we consider gaps to be areas in which there is either no consensus or there is no item currently developed to measure a concept of interest. Following are a set of recommendations that would address key gaps:

1. Develop a standard phraseology for questions asking about quit attempts, in which the respondent is instructed to only include those events in which their intention was to quit either temporarily or permanently.
2. Undertake further research to determine the applicability and nature of slip events for youth.
3. Identify strategies for measuring slip events, including items to describe how “long” and “deep” the event(s) are.
4. Consider further research to improve our understanding of the patterns, timing, and predictors of adolescent smoking relapse events.
5. Consider further research to improve our understanding of associations between physical, social, temporal, emotional and situational contexts and smoking patterns for youth smokers.
6. Address how standards of acceptability should be developed and applied for evaluation of community interventions that use single-group designs.
7. Develop guidelines for the implementation of biochemical validation and the use of control groups for youth smoking cessation research studies.
8. Further research should be considered to investigate youth’s recall of cigarette consumption in the recent past.
9. Adopt 30 days non-smoking as the criterion for a successful “quit” for youth smokers.
10. Consider developing a standard for a successful quit based on three criteria being met: i) intention to quit permanently, ii) 30-days smoke free, and iii) self identifies non- or former smoker.
11. Record values as continuous data, avoiding imposed categorization in data collection forms whenever possible.
12. Extend the time period used in consumption recall questions to 30 days, but continue collecting seven-day recall data concurrently.
13. Modify the definition of a quit attempt from minimum criterion of 24 hours abstinent to one week abstinent for non-daily smokers.
14. Develop standard phraseology for questions asking about quit attempts, in which the respondent is instructed to only include those events in which their intention was to quit permanently.
15. Specify minimum pre- and post-test assessment for outcome evaluation. Both tests should use a core set of items that are directly comparable between pre and post tests.
16. Promote minimum 3- and preferably 6-month follow-up for youth smoking cessation interventions. Encourage longer follow-ups when possible.

Finally, based on our review of the literature, discussions with colleagues and experience in similar projects related to indicators in tobacco control, the following issues/questions regarding next steps should be considered by the working group:

1. Does the working group agree, in principle, with the core indicators recommended?
2. Does the working group agree, in principle, with the measurement items recommended?
3. Does the working group feel comfortable recommending moving forward with a set of indicators / measures identified here? If not, what modifications are necessary?
4. What are the most useful processes to encourage adoption of the recommended indicators and measurement items? Is there opportunity to recommend policy requiring use of the measures at Federal, Provincial, and Territorial levels, or perhaps as review criteria for cessation intervention funding applications? If yes, how can we best move toward such policy? How would this be disseminated?
5. Should we move towards developing a standard (or "core") set of measures for intervention implementation and participant characteristics?
6. What background information on evaluation issues (design issues, process issues, implementation issues) should be distributed with core indicators and measures?
7. What infrastructure is required to take this initiative to the next phase?
8. We do not recommend that community-based evaluations of youth smoking cessation interventions use biochemical validation as a practice standard. The working group should provide guidance regarding in what situations biochemical validation should be used.
9. Would it be worthwhile to put some effort towards developing workable measures for slips as it may apply to the youth population, or is a better understanding of this important cessation concept as it applies to youth smokers required before we can develop measures to assess it?
10. The advisory committee for this paper suggested we focus on behaviours; however, the working group may wish to consider the utility of including other types of questions (i.e., non-behavioural questions) in future investigations.
11. A limitation of this paper is the exclusion of French language evaluation questionnaires. The extent to which French language questionnaires currently exist, or the issues associated with translating and validating French versions based on English language questions is not addressed in this paper, but is an issue the working group may wish to discuss.

In summary, this document provides a review on the "state-of-the-art" in measurement of youth smoking for the purposes of evaluating smoking cessation interventions. This information will provide the basis for further discussion and consensus building at an expert workshop to be held in November 2006, funded by Health Canada. It is hoped that the results of that workshop will contribute to the development of core indicators for evaluating youth cessation programs, an area in which there is currently a lack of evidence to inform best practices.

1 BACKGROUND

1.1 Purpose of the report

Current research reveals a lack of evaluation data regarding youth smoking cessation interventions that can help decision-makers and practitioners arrive at informed decisions regarding better practices. The Youth Tobacco Cessation Collaborative (YTCC) is a collaboration of major US and Canadian organizations that fund research, program, and policy initiatives in youth tobacco control. While the efforts of the YTCC have focused on coordinating intervention activities and building capacity to deliver effective cessation interventions to youth, there is an imperative need for evaluation data that will guide policy and decision-makers in identifying effective approaches. The YTCC is working to address knowledge gaps in evaluation of youth tobacco control initiatives, and this project seeks to address these gaps.

1.2 Anticipated outcomes

This document provides "building blocks" that can be used as a starting point for identifying and facilitating consensus building regarding the adoption of standardized indicators and measurement items for evaluating youth smoking cessation programs. The indicators identified in this report will be taken forward to a second project phase, in which the indicators will be tested and validated among both youth smokers and a panel of tobacco control experts. The final set of items resulting from this process is intended to be used in the evaluation of youth smoking interventions within the real-world context, as well as, potentially, for studies conducted in the research context, where greater control of confounding variables is feasible.

Ultimately, the identification and adoption of a common set of measures—perhaps even the establishment of a registry of youth smoking cessation programs and the measures they employ—would facilitate comparison of results across programs implemented by schools, and government and non-government organizations in community settings. A common set of measures could also be adopted for national tobacco surveys. The indicators could also be made accessible to researchers conducting efficacy trials in more controlled research settings. We also feel it is important that, when considering the applicability of common indicators, the scope of interventions include macro-level policy initiatives such as taxation policy or smoke-free legislation in addition to more targeted interventions such as quitlines or school-based stop-smoking programs.

1.3 Anticipated use

The background information contained in this document is meant to provide a foundation that anchors future discussions regarding the selection of indicators and measures for evaluating the outcomes of smoking cessation programs or interventions for the youth population. In particular, it is meant to serve as a resource document for an expert workshop that will be held on November 16, 2006 entitled *Indicators and Measurement Items for Outcome Evaluation of Youth Smoking Cessation Interventions*.

In the next phase of this project, a working group, committee, or a similar group will review, develop and oversee the process of specifying a set of indicators and measurement items for which there is general consensus and commitment. The indicators and their associated measures are intended to guide the collection of outcome data for evaluation studies of youth smoking cessation programs implemented in schools, communities, and clinical settings.

Because this document is oriented towards developments in research and measurement and it is likely to evolve on the basis of deliberations at the expert workshop, in its current form it will not meet the needs of community evaluators or program professionals. A more applied and "user-friendly"

document providing guidance on the selection of specific measures will be produced in later phases of this project.

The product of this work contributes to the YTCC's two-year goals and objectives of the National Blueprint for Action for 2005-2007[4] in the following areas (the YTCC's goal and objective numbering is used in the list below):

Research goal 2: Increase national and state-based surveillance of youth and young adult tobacco-use cessation services, behaviour (including use of treatments and services), and policies.

Objective 1. Identify, assess and recommend current measures on surveys, revise measures or include additional measures, and conduct surveys of youth and young adult tobacco use cessation behaviours.

Research goal 3: Develop and test tobacco-use treatment programs, services, and interventions for youth and young adults.

Objective 2. Develop standard program for evaluation methods; evaluate the efficacy of various types of services, support and other interventions.

Implementation goal 2: Increase the capacity and capability to deliver effective youth and young adult tobacco-cessation interventions and services.

Objective 1. Develop and promote the use of quality control strategies in the delivery of cessation interventions.

Objective 2. Identify interim criteria for measuring the effectiveness of existing cessation interventions until guidelines become available.

2 HISTORY OF THE PROJECT

This project follows up on the January 2004 Youth Cessation Roundtable convened by Health Canada to discuss current issues in youth smoking cessation research, programming, and evaluation. The Roundtable process and subsequent discussion with expert researchers in the field led to the conclusion that not only is there a lack of information on the effectiveness of youth tobacco cessation programs, but that, even more fundamentally, there is a lack of a clear and shared understanding of what measures are and have been used, what the strengths and limitations of these measures are, and what measures should be considered for which applications.

3 SCOPE OF THE PROJECT

The five "building blocks" produced in this background paper include:

1. *Logic model[5]:* The logic model (conceptual framework) illustrates a heuristic for youth smoking cessation that includes a set of definitions of outcome behaviours related to smoking cessation. From this model, behaviours and associated indicators for measurement are identified.

[4] http://www.youthtobaccocessation.org/blueprint/index.html

[5] Appendix A contains a glossary of terms and list of acronyms used in the report, so that the reader and authors will share a common understanding of intent.

2. *Review of measurement items used in intervention and surveillance program:* This report is derived from results of a scan of available measures of tobacco cessation, focusing on those measures related to the evaluation of youth tobacco cessation programs. This is based on a review of published scientific literature, as well as information gathered from key individuals from Canada and the US identified as national experts.
3. *Identification of indicators and measurement items for evaluation:* This set of recommended measures identified from the published and unpublished literature for evaluating youth smoking cessation programs includes: (1) validated core items, (2) identification of items that show promise but need further testing, (3) identification of a small number of non-core (optional) items, and (4) specification of gaps in the list of items available.
4. *Identification of key measurement issues:* This set of issues summarizes the relative strengths and limitations of identified measures.
5. *Recommendations:* These recommendations and key questions can be used to develop consensus regarding a core set of items for evaluation studies as well as key measurement issues where additional work is needed. This work will also be useful for surveillance purposes.

4 APPROACH

This section outlines our approach, including how we have defined inclusion and exclusion criteria, developed our logic model, conducted a review of the literature, consulted experts, identified indicators and items, identified key measurement issues for consideration, and developed recommendations.

4.1 Inclusion and exclusion criteria

This paper focuses specifically on behavioural outcomes—changes in smoking behaviour—by youth smokers who have enrolled in an intervention program with the intention of quitting. There was consensus among the advisory group that the process leading up to enrolment in an intervention, issues related to participation, and non-behavioural outcomes—including nicotine dependence, stages of change, skills building, knowledge, attitudes, perceptions and self-identification—are outside the scope of this document. This paper focuses exclusively on action taken to quit, and does not address the trajectory of youth smoking leading to this action.

The following inclusion/exclusion criteria frame the parameters of this review:

1. *Type of measure:* Only those smoking behaviour measures suitable for self-report survey administration are considered.
2. *Population:* "Youth" are defined as individuals between 12 and 18 years of age.[6]
3. *Type of tobacco consumed:* The review focuses exclusively on cessation of smoking cigarettes.[7] Smoking other forms of tobacco (cigars, pipes, etc.) and non-combustion use (chew, snuff, etc.) is not considered.[8]

---

[6] The age range was decided upon by the advisory group.

[7] We recognize that, ultimately, development of measures for cessation from all forms of tobacco use would be a worthwhile objective; however, at this early stage we have chosen to focus our efforts on cigarette smoking cessation specifically.

4. *Smoking history:* The standardized but arbitrary measure of having smoked 100 or more lifetime cigarettes was used to define a smoker, thereby excluding individuals termed puffers and experimenters.[9]
5. *Smoking pattern:* Due to the wide variety of patterns of youth smoking, the only restriction we have considered is that the individual must have smoked at least 100 cigarettes. We note that it is not uncommon for adult cessation programs to be limited to daily smokers.
6. *Geographic scope:* The study was restricted to measures used in Canada and the United States.

4.2 Development of the logic model

To help frame the range of behavioural outcomes relevant to the identification of specific indicators for measuring youth smoking cessation, we developed a logic model in which the macro-level inter-relationships between a cessation intervention program and behavioural outcomes are illustrated. The logic model providing the framework for this report is shown below in Figure 1.

In specifying the logic model, we are focusing only on changes in smoking behaviour following participation in an intervention. The focus of the present paper is on measures for outcome evaluation; consequently, neither program outputs nor factors related to an individual's propensity to participate in an intervention fall within the scope of this paper. However, it is important to recognize that program characteristics are equally important to describe in any evaluation (e.g., target population, type of treatment, format, number of sessions, leader characteristics).

The logic model describes the flow of individuals through an intervention program. The model is situated within an existing policy and environmental context. These external factors interact with treatment ingredients to influence outcomes. The left-most box acknowledges the process that needs to occur before individuals are ready to participate in an intervention program with the intent to quit smoking. The program inputs used in the execution of the program result in a number of program outputs (the logic model shows a selection of such outputs as examples). Ideally, effective programs will result in a desired set of outcomes for individuals entering the program. This paper's concern is with measurement of behavioural intervention outcomes, specifically quitting smoking. Recognizing that behavioural changes may not occur immediately, and that individuals may not be able to maintain the intended change, outcomes in the model are characterized as immediate, medium, and long-term, so that the timing and duration of behavioural changes following the intervention can be assessed.

One of the limitations of this logic model is that it outlines common outputs and outcomes, but does not describe the intervention approach.

The arrows exiting each outcome box illustrate that individuals may either continue their successful change into the future, or they may fail to maintain the change. Program success is achieved with

---

[8] We also do not consider the possible substitution of cigarettes for another form of tobacco when individuals are attempting to reduce / quit smoking cigarettes (e.g., shifting from cigarettes to smokeless tobacco).

[9] While this may exclude some very low-dosage youth smokers who may arguably have developed symptoms of nicotine dependence having smoked fewer than 100 cigarettes, given that the focus of this paper is on smoking cessation, we feel it is reasonable for our present purposes to use this definition for minimum lifetime consumption, although we recognize that this criterion may exclude some novice smokers who wish to stop smoking and need help to do it. Whether youth can accurately estimate their consumption of 100 cigarettes, or whether 100 cigarettes has a fundamental validity as a marker of habituation or nicotine dependence, is the subject of debate in the literature and is a discussion we do not engage in here.

continued smoking cessation, as represented by the right-most box. Intervention failure is represented in the bottom box, which incorporates both relapse (restarting smoking after a quit attempt) and reduction.[10] Relapsed quitters or individuals who had reduced consumption but have not attempted to quit would have opportunities to participate in another intervention program in the future (or to take more time preparing to do so). We should note that the logic model is a simplified, idealized representation of the intervention process; consequently, program drop-outs and loss to follow-up are ignored. The model is intended to provide a conceptual framework to describe an individual's pass through a single cessation intervention.

The scope of items we considered is represented by the items contained in the large grey box in the logic model. While there is some evidence that other factors occurring earlier in the change process (e.g., behavioural intention) may lead to cessation, the focus of the present paper is on outcomes in smoking behaviour following a cessation intervention. We found that logic models were not discussed in the literature reviewed for this paper; however, the model presented is consistent with the approaches described in the literature on the measurement of behavioural outcomes from youth smoking cessation programs.[11] The model is situated within the smoking cessation logic model described in the US CDC's guide to evaluation of tobacco control programs (MacDonald et al., 2001).

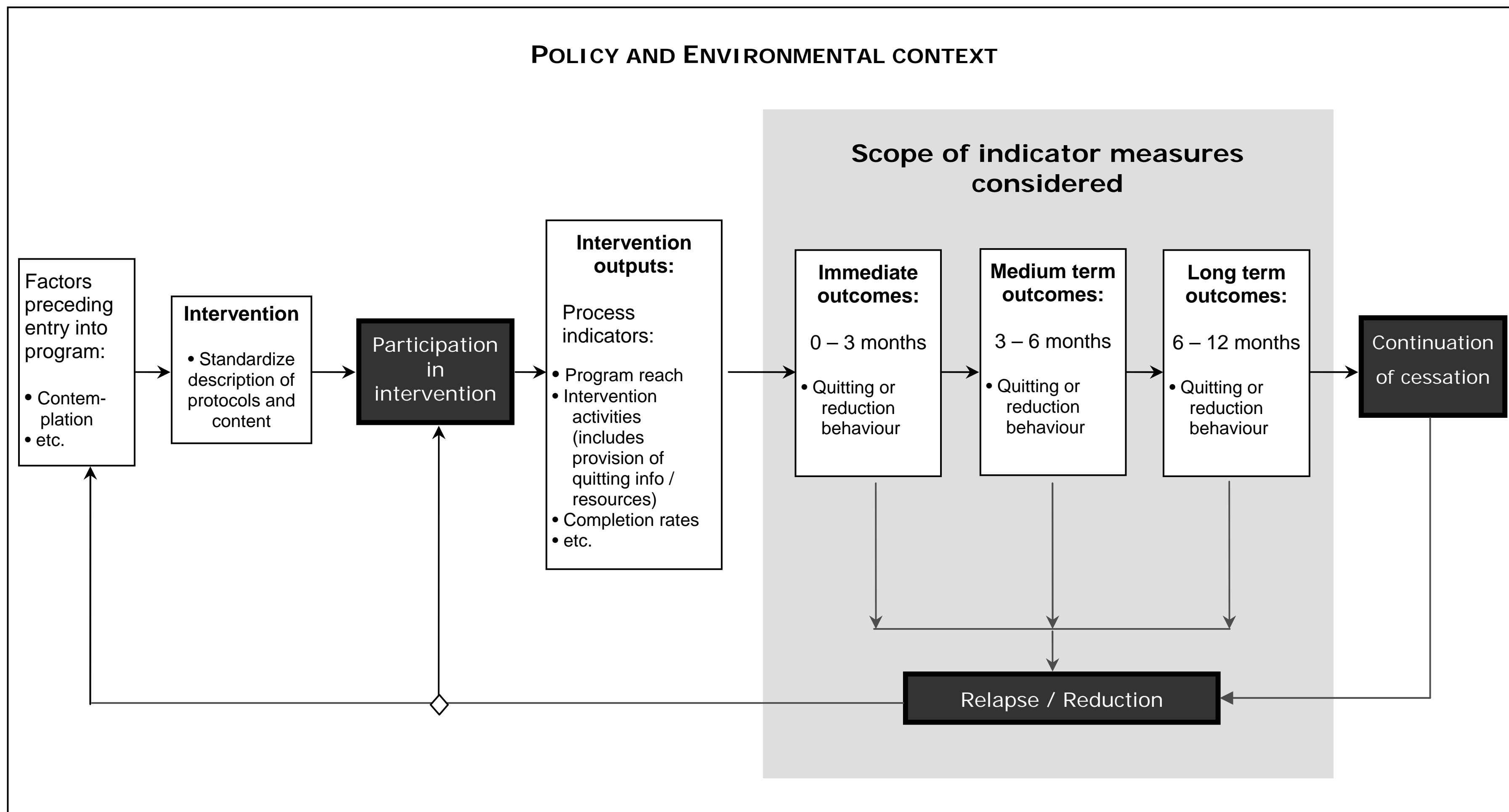
The model presented shows immediate-, medium-, and long- term outcomes. While there is no clear guidance from the youth smoking cessation literature in terms of specifying the boundaries between these three time frames, they are intended to indicate relative rather than absolute time intervals, ranging from initial changes in consumption level or quitting attempts towards long-term continuous abstinence.

A preliminary version of the model was presented for discussion at a meeting of the project advisory group in Ottawa on March 9, 2005. The current model incorporates many of the changes suggested at that meeting.

[10] While reduction is included in this paper as a behavioural indicator of interest for measurement of cessation program outcomes, it must be recognized that the desired outcome for cessation programs is smoking cessation, not reduction. While reduction may be viewed as a positive step—and perhaps an intermediate one on the road to eventual cessation—we ultimately view reduction alone as a treatment failure for cessation programs.

[11] We should note that our literature review did not include a review of the theoretical literature on youth smoking cessation.

Figure 1: Logic Model for the outcome evaluation of behaviour change related to a youth smoking cessation intervention



4.3 Review of measurement items

The literature on youth smoking cessation is a relatively new field, so the number of reports published are, at present, relatively few but rapidly growing in number. Recognizing this, the literature review conducted for this report builds on the foundation of work represented by key reviews on youth smoking cessation measurement such as those by Sussman et al. (1999), Mermelstein et al. (2002), Sussman (2002), Garrison et al. (2003), and O'Loughlin et al. (2003), as well as the articles contained in the American Journal of Health Behavior's 2003 special issue on youth smoking cessation. We undertook a search of the 1966-present Medline database using terms such as adolescent, youth, tobacco, smoking, cessation, indicator, measurement, and outcome, as both keywords and MEDLARS descriptors. The search strategies we employed in this literature review are discussed in more detail in Appendix F. A total of 102 articles were retrieved for further review. A list of the articles we collected for our literature review is contained in Appendix G.

A select group of expert researchers provided suggestions regarding key literature and measures that should be included in this review; however, the number of individuals contacted was necessarily limited and selective, due to our timeline and the availability of experts recommended. Thus, it is possible that the advice provided does not represent the full range of views held by all experts in the field. We did discover, however, a high level of consensus in comments made regarding existing gaps.

We identified and reviewed key articles from the published literature and relevant work in the grey literature (unpublished, disseminated through unconventional means such as internet sites, or works in progress). To locate grey literature, we employed three main strategies. First, we contacted researchers who are recognized as experts or have access to specialized knowledge in youth smoking cessation, using contacts gained through attendance at the National Alliance on Tobacco Use Cessation (NATUC) meeting in Washington, DC on December 16, 2004. Secondly, we contacted selected authors of papers in the American Journal of Health Behavior's 2003 special issue on youth smoking cessation. CTCRI, as a participant in that issue, facilitated such contacts. Thirdly, we used CRISP[12], a searchable database of research projects funded by various branches of the US federal government, as one means of identifying recently funded youth smoking cessation projects. Because our objective in using CRISP was to identify recent projects that were unlikely to have yet published findings in peer-reviewed journals, we restricted our search of the CRISP database to projects funded in 2003 and 2004.

4.4 Identification of indicators and measurement items

To identify the major evaluation questionnaires used in intervention programs in Canada, we drew on the findings of a recent review by O'Loughlin et al. (2003) on the evaluation of youth smoking cessation programs in Canada, and supplemented O'Loughlin's findings through consultation with experts at Health Canada and those in the tobacco control community, in addition to identifying instruments used or referenced in the published and grey literature.

We identified 27 intervention programs or surveillance surveys containing measures appropriate for the measurement of youth smoking cessation. Although the primary focus was on items used in evaluation studies, we also examined measures used for surveillance because they are a source of standardized items for evaluating interventions and can be used as a standard of comparison. From this set of surveys, we selected 18 intervention programs or surveillance surveys, representing 34 individual questionnaires, for detailed review based on the following inclusion criteria:

---

[12] CRISP is Computer Retrieval of Information on Scientific Projects, a database funded by the US National Institutes of Health (NIH). http://crisp.cit.nih.gov/

1. *Currency:* The questionnaire is currently used in evaluating youth smoking cessation interventions.
2. *Widespread use:* The questionnaire is currently used nationally or at least at the provincial/state level.
3. *Identified by an expert:* The questionnaire (or program) was identified favourably in either a recent review of programs or in personal communication with recognized experts in the field.
4. *Cited in the literature:* The questionnaire is referenced in recently published articles by more than one author or group.
5. *Face Validity:* The questionnaire "makes sense" in terms of constructs we are trying to measure.
6. *Canadian context:* Questionnaires used in Canada were given priority.
7. *English language:* French language questionnaires were excluded from this review.

The questionnaires selected for review in this paper met all of the above criteria.[13]

In order to be considered for inclusion in our review, evaluations of intervention programs must have used at least two evaluation questionnaires: a baseline version that is completed by program participants prior to the intervention (the pre version), and a post version that is completed by participants following the intervention. While this paper is concerned with the identification of measurement items for the evaluation of behavioural outcomes in cessation interventions, we acknowledge that there has been a large investment in the development of smoking behaviour measures for surveillance surveys. Surveillance items may potentially be adapted for use as evaluation measures. Thus, we have provided an overview of smoking behaviour items in questionnaires used in major surveillance projects related to tobacco control. As with the evaluation questionnaires, we apply the criteria outlined above to select the questionnaires used in surveillance. Cross-sectional surveys naturally do not have pre and post versions, although prospective longitudinal surveys do offer an opportunity to follow individuals forward in time similar to pre/post evaluation questionnaires, with the notable difference that there is no intervention associated with the survey itself.

A select number of expert researchers were identified through the published literature, through contacts made at the NATUC meeting in Washington, DC, recommendations from members of the project advisory committee, and referral from one expert to another. All experts were asked for their input on the identification of appropriate indicators and measures for youth smoking cessation, key works in the published literature, relevant unpublished studies, and their expert opinions on a series of specific questions regarding measurement issues pertaining to youth.

4.5 Identification of measurement issues

We summarized all the information gathered through our review of the literature, scans of questionnaires, and discussions with experts, into descriptive tables. These tables were then used to perform a conceptual synthesis of our findings. With respect to survey measures specifically, we conducted a review of what measures exist, compiled these measures, and selected from the best of these measures.

[13] See Appendix D for complete copies of the evaluation questionnaires reviewed, and Appendix E for reference information.

Steve Sussman also helped by reviewing all studies from his 2002 paper that were published from 1996 forward, documenting which studies identified the origin of questionnaire measures used, and which papers reported on validity and reliability for the measures used.

To facilitate our conceptual synthesis of the questionnaires, we created tables that abstracted the questionnaire content into a form that would be directly useful for our purposes. We did this by taking the collected information, and re-grouping and organizing the summary content into higher-level categories to provide an effective overview of coverage, gaps, and variations in measurement between the different questionnaires reviewed.

For the intervention evaluation questionnaires, we provide a table with a high-level summary in Appendix B. Copies of the intervention evaluation questionnaires are provided in Appendix D. A high-level summary of the smoking behaviour questions for the population surveillance surveys is provided in Appendix C, and the grouping of measurement areas parallels that used for the intervention questionnaires as much as possible. Because the focus of this paper is on evaluation, details of the surveillance survey question wording is not provided in a separate table. Reference information for both the evaluation and surveillance surveys themselves is provided in Appendix E.

The two first authors of this paper identified the main issues for discussion, with input from other members of the advisory group.

4.6 Development of recommendations

We developed the set of recommendations provided in Section 7 using the information contained in the summary and detail tables describing the content of the reviewed questionnaires. We combined this information with findings of the literature reviewed, discussions with experts, and input from the project advisory group. A core set of items was identified, for which there is general agreement regarding utility and validity. Existing gaps in items available were also identified, as well as recommendations regarding ways in which items could be improved. In the course of synthesizing information from different sources, we also developed a set of questions that should be addressed by the working group who will take this project onto its next phase.

5 FINDINGS

This section presents the findings from our synthesis of the published and grey literature, with a focus on both measurement issues and the content of questionnaires used in evaluation studies and surveillance. Section 5.1 provides an overview of the interventions and surveillance programs that used the surveys we reviewed. In Section 5.2, the questionnaire content, in terms of the indicators and items represented in the intervention programs reviewed, is described.

In general, the literature on methodological issues in the evaluation of smoking cessation interventions for youths is relatively recent and is not well developed; researchers also continue to grapple with fundamental issues such as definitions and validity of measurements for youth. In contrast, the literature on the evaluation of smoking interventions for adults is well developed and mature. Research activity on youth smoking cessation did not develop significantly until the 1990s (Backinger et al. 2003a), and a shortage of evidence regarding effective treatment for youth smokers persists (Curry 2003).

The published literature acknowledges that there are many problems associated with attempting to develop survey measures to characterize youth smoking behaviour, particularly with regard to cessation. This was underscored in our discussion with experts in the field and reaffirmed in our review of documents. These limitations are discussed in more detail below. Clearly, the existing

measures are not ideal; however, they still provide useful information in situations where decisions must be made regarding better practices in youth smoking cessation.

The great majority of published and grey literature does not report on reliability and validity of the questions used in the evaluation. Of the 77 articles included in our review, only 22 provide any information regarding the reliability and validity of measures used; this number drops to just 6 if papers that do not describe item validation other than biochemical validation or bogus pipeline methods are excluded.

Relatively few questionnaires intended for a youth population and containing items related to cessation were identified in our discussions with experts or are referenced in the literature. There is broad similarity in the topics covered related to change in smoking behaviour and in the general thrust among the different instruments that we reviewed on youth smoking. Within this general similarity, however, there are significant differences in question wording and in when and how data collection is implemented. Such differences can affect the results obtained in evaluation studies. Below, we enumerate the main measurement approaches found in the reviewed questionnaires. Major measurement issues in the literature are incorporated into the discussion in Section 6.

5.1 Interventions/surveillance program questionnaires

This section describes the characteristics of the intervention programs or surveillance, which use the questionnaires reviewed in subsequent sections of this paper. We offer this description in order to better understand the context in which the reviewed questionnaires are implemented.

Table 1 describes the general characteristics of the programs using the questionnaires reviewed here. For convenience, we refer to the questionnaires by relating them to the intervention programs that use them. For example, we refer to the questionnaires used to evaluate the Kick the Nic (KTN) program as the KTN pre and KTN post questionnaires. It should be mentioned here that the Helping Young Smokers Quit (HYSQ) questionnaires were developed as part of an initiative to develop a common set of measures for use in a large number of intervention programs. Currently, the HYSQ evaluation questionnaires are being used in 42 intervention programs in the United States, although the evaluation protocol varies between programs.

With the exception of the North American Quitline Consortium's (NAQC) Minimal Dataset (MDS), all the evaluation programs using the reviewed questionnaires specifically target this paper's age range of interest. All but the Towards No Tobacco (TNT) program and the School Smoking Profile (SSP), require that individuals either self-identify as a current smoker or indicate that they have smoked at least one cigarette within the last month. In all programs, pre-test questionnaires are administered either before or at the first session of the intervention (or at initial contact in the case of the NAQC-MDS) and the post-test is typically administered at the last regular session or very shortly afterward.[14] A number of programs have follow- up questionnaires as either a standard or optional part of the program, and the intervals at which the follow-ups are to be administered vary or may be at the discretion of the program facilitator.

A summary of the evaluation questionnaire content can be found in Appendix B, and copies of the evaluation questionnaires can be found in Appendix D.

---

[14] Towards No Tobacco (TNT) administers the post-test one week after the intervention; the NAQC-MDS' guidelines ask for a follow-up seven months after the intervention.

Table 1: Summary of intervention programs using reviewed evaluation questionnaires

Program name	Where used	Age range	Smoking eligibility	Pre-test	No. of sessions	Post- test	Follow-up
HYSQ (U. of Illinois at Chicago: Curry)	U.S.	14-18	Varies by program	Before 1st session[15]	7 - 8[16]	Within 2 weeks of last session	6 and 12 months after 1st session
Kick the Nic (BC Ministry of Health)	Canada	13-19	All smokers	1st session	10	Last session	Optional
NAQC-MDS (NAQC: Bailey/Campbell)	Canada / U.S.	Varies by quitline[17]	All smokers[18]	Intake	Varies by quitline	Varies by quitline	7 months after program
No More Butts![19] (NS Dept of Health)	Canada	15-19	Daily smokers only	1st session	15	Last session	No
N-O-T (NCI, ALA)	U.S.	14-19	Regular[20] smokers only	1st session	10	Last session	3-6 months after last session “suggested”
Project EX (U of South California: Sussman)	U.S.	14-18	Smoked in last 30 days	Intake	8	Last session	3 months after program
Quit4Life (Health Canada)	Canada	12-18	All smokers	1st session	10[21]	Last session	Optional[22]
SSP[23] (U of Waterloo: Manske)	Canada	11-18	All students	n/a	n/a	n/a	n/a
TAP (Community Interventions Inc)	U.S.	12-18	All smokers	1st session	8	Last session	3,6, & 12 months “recommended”
TNT (U of Southern California: Sussman)	U.S.	14-18	Smokers and non-smokers	1 week before program	10	1 week after program	At 1 and 2 year booster sessions

In a broad sense, there are two applications for measures of youth smoking cessation behaviour: population surveillance, and evaluation of specific interventions. While this review focuses on the

---

[15] For a few programs, baseline questionnaire was completed between 1st and 2nd sessions.

[16] This is an average figure; actual numbers vary by program.

[17] Because quitlines within the NAQC have different target populations and offer different kinds of services, the MDS does not provide a definition of “eligible callers” for quitlines. While oriented towards adult smokers, members within the NAQC provide services to individuals in this paper’s age range of interest.

[18] While one would assume that only current smokers would call a quitline looking for assistance to quit smoking, the eligibility for use of quitline services is not uniform across the NAQC. For our present purposes we assume that eligibility is limited to current smokers who are looking to quit.

[19] Personal communication: Nancy Hoddinott, Coordinator of Tobacco Strategy, Tobacco Control Unit, Nova Scotia Department of Health.

[20] “Regular” smoking is left for participant to self-identify.

[21] Q4L offers an additional 4 optional sessions.

[22] Maintenance focus may be offered one or more times, at facilitator’s discretion. The interval after last session or between follow-ups is not prescribed.

[23] The School Smoking Profile is a monitoring tool that can be used in the evaluation of programs or policy changes. It is not part of a particular intervention program. Please note: the current version of the SSP is a module of the School Health Action Planning and Evaluation System (SHAPES).

latter, items used for surveillance can potentially be adopted for use in questionnaires evaluating specific interventions. Table 2 provides an overview of the population surveillance surveys selected for review, with the year of the survey reviewed, the age range eligible to participate, and where the survey is used. For surveillance, we have limited our review to major national level current surveys either with significant smoking content, or that are commonly used in the youth smoking literature. With the exception of the Behavioural Risk Factor Surveillance System (BRFSS) and Canadian Community Health Survey (CCHS), the target age range for all the selected surveys falls within the range of interest for this paper. While the Canadian Tobacco Use Monitoring Survey (CTUMS) is a general population survey, CTUMS deliberately over-samples individuals in the 15-24 age group. The CCHS, CTUMS, and Youth Smoking Survey (YSS) are the primary surveys for the surveillance of smoking in Canada. The questions asked in the reviewed surveys are summarized in Appendix C. Broadly speaking, the behaviour questions asked in these surveys do not differ dramatically from the items we documented above in the evaluation questionnaires.

Table 2: Summary of selected national population surveillance surveys

Survey name	Age range	Where used
Behavioral Risk Factor Surveillance System (BRFSS)	18+	U.S.
Canadian Community Health Survey (CCHS) 2.1 (2003)	12+	Canada
Canadian Tobacco Use Monitoring Survey (CTUMS) 2004	15+	Canada
Monitoring the Future (MTF) 2003	Grades 8,10,12	U.S.
National Youth Smoking Cessation Survey (NYSCS)[24]	16-24	U.S.
National Youth Tobacco Survey (NYTS)	9-21	U.S.
Youth Risk Behavior Survey (YRBS)	Grades 9-12	U.S.
Youth Smoking Survey (YSS) 2004-05	Grades 5-12	Canada

5.2 Indicators and items used to evaluate interventions

This section describes the content of the questionnaires used to evaluate youth smoking cessation interventions. In the intervention program questionnaires reviewed, indicators of smoking behaviour change include:

I. *Abstinence* (cessation or zero consumption since intervention),

II. *Reduction* (consuming fewer cigarettes per unit time than at baseline), or

III. *Quit attempts* (actual attempts made to quit smoking)

These three indicators are also widely used in surveillance. While intervention questionnaires can assess changes in consumption levels prospectively through pre- and post-test questions, in surveillance a survey must use either a longitudinal design (as used by the NYSCS) or ask about consumption change as a retrospective question in cross-sectional surveys. Surveillance surveys also differ from intervention questionnaires in that cessation questions cannot work forward from a known, intended quit event, and consequently may only ask about quit attempts within some specified past time period and about the length of time since the respondent last quit. Because our present focus is identifying and describing indicators and items for the evaluation of youth smoking cessation interventions, we do not provide a detailed commentary on the items used in the surveillance questionnaires; however, for the reader's interest we have produced a summary table of surveillance items, similar to that provided for the evaluation questionnaires, in Appendix C.

[24] NYSCS is a longitudinal survey that is part of a research initiative entitled "Assessing Youth Smoking Cessation Needs and Practices."

In the sections that immediately follow, we characterize the ways in which the evaluation questionnaires measure each of the three indicators (cessation, reduction, and quit attempts).

5.2.1 Abstinence (cessation)

For this paper, we use the term abstinence to identify a successful (or not-yet-failed) quit attempt. All the questionnaires used to evaluate interventions included questions on abstinence in the post-intervention questionnaire. Most questionnaires asked for the length of time since the respondent's last cigarette (KTN, NMB, NOT, HYSQ), providing a measure of continuous abstinence, although the start point for the reported abstinent period may not have been the intervention's planned quit date. The abstinence question in the Q4L questionnaire asks respondents to identify a time period corresponding most closely with their longest quit attempt since starting the Q4L program, and an option for "quit completely" is provided.[25] The NAQC-MDS asks if the respondent has smoked any cigarettes in the last 30 days. Both the NAQC-MDS and the HYSQ questions specifically ask about smoking even a few puffs.

5.2.2 Reduction

Determining whether an intervention program has achieved a reduction in a participant's cigarette consumption is done by comparing pre- and post-test cigarette consumption, as reported by the participant. Such a comparison may find that consumption levels have decreased, stayed the same, or even increased.

All questionnaires collect information about the respondent's cigarette consumption at the time the questionnaire is administered, although the approach used and the level of detail regarding current consumption varies between instruments. Most questionnaires use various questions to ascertain usual, average, or typical smoking levels (Q4L, KTN, NMB, NOT) and/or ask for an estimate of the amount smoked in the last week (Q4L, KTN). A few questionnaires adopt the 7-day "wheel" seen in the YSS and CTUMS (KTN pre, HYSQ), asking for the actual number of cigarettes smoked in each of the previous 7 days. Interestingly, given that it is widely recognized that youth smoking patterns can often be irregular or sporadic, only the NOT, HYSQ, and NAQC-MDS questionnaires ask about consumption over the last 30 days—and the NOT only asks if respondents had smoked on 20 or more of the last 30 days (and does not ask about number of cigarettes over that period).

The NOT questionnaires are unique in explicitly asking about typical weekday and weekend day consumption level separately. The HYSQ and NAQC-MDS share the familiar approach of asking first on how many of the previous 30 days the respondent smoked, followed by a question asking how many cigarettes they smoked on the days they did smoke. None of the questionnaires parallel this approach for the previous 7 days. The KTN, NMB, and HYSQ post-intervention questionnaires all have questions in which the respondent can self-identify directly if they have reduced their smoking and maintained this lower level following the intervention, if they reduced their smoking for a while but were unable to sustain it, or if they did not change their smoking consumption after the intervention. Only the KTN asks respondents for the length of time that they have reduced their smoking. As consumption level questions are asked both pre- and post-intervention for most programs, these questions can be used to identify and estimate the magnitude of changes in consumption levels following the intervention.

---

[25] It is not clear if individuals who have made more than one quit attempt since starting the program would indicate a period of time or the "quit completely" option, since they may consider themselves to have quit completely even if they slipped or failed in their first quit attempt(s). See discussion of quitting and slips in Sections 6.4 and 6.7.

Questionnaires sometimes request consumption information in ways that may not yield valid information. For example, in the KTN baseline questionnaire, respondents are asked how many cigarettes they smoke "on an average day." This question is asked of all respondents, so very low dose, non-daily smokers could be expected to have great difficulty providing a meaningful answer to this question. A question asked in both the pre- and post-intervention Q4L questionnaire may be difficult for individuals who do not smoke every week to answer: if they elect to respond no to a question asking if they "usually smoke every week," their consumption will be coded as zero; if they choose to answer yes to the question, then each option in the follow-up question assumes that they smoke to some degree every week.[26] In the KTN post questionnaire, post-intervention consumption level is asked of all respondents except those with failed quit attempts; consumption level information should be collected from all individuals at post-test, including those who continue to smoke after a quit attempt. This gap in the KTN prevents the evaluation of change in smoking behaviour for those participants that fail to quit.

5.2.3 Quit attempts

In this review, the term quit attempt describes a failed attempt to quit. All of the intervention questionnaires we reviewed, with the exception of the NAQC-MDS, ask about previous quit attempts in the baseline questionnaire. Most ask some combination of number of attempts ever or in the previous year, and the length of the longest or most recent attempt. Attempts typically have to have been at least 24 hours in duration to qualify as a quit attempt. Three questionnaires (Q4L, KTN, and HYSQ) specifically ask about quit attempts post-intervention. The Q4L and HYSQ questionnaires simply ask what the longest period of time the respondent has gone without smoking since starting the intervention (the shortest response option provided in the Q4L is "a few hours"; HYSQ counts only days). The KTN asks respondents to complete the sentence, "I stopped smoking for ___ days before I started smoking again." None of the questionnaires asks about the length of time preceding and following quit attempts, nor about consumption levels during these smoking periods (except for current consumption level). Only the HYSQ instrument asks about the number of attempts the respondent has made to quit smoking since starting the intervention.

6 MEASUREMENT ISSUES

Throughout the research process for this paper, we have identified and considered measurement issues in the evaluation of youth smoking cessation interventions. Key issues were identified at the advisory group meeting in Ottawa (March 2005), in our discussions with Canadian and American experts in youth smoking cessation, in further reviews of the literature, and during our synthesis of questionnaire items. Out of this process and the subsequent synthesis of our findings on questionnaire content, a list of measurement questions emerged. We have not addressed all of these questions in this discussion section; rather, we speak to those questions that are discussed substantially within the literature or among the experts we consulted. We do not mean to infer that questions that have not yet been addressed in the literature are not worthy of discussion, but we are leaving such gap issues for future research. Some of these questions encompass very broad issues that represent gaps in the current knowledge or require extensive research to address; consequently, we do not see our comments on measurement issues as final, and we anticipate that further discussion in future phases of this project will improve our understanding of these measurement questions, In cases where we have identified a measurement issue but have not found adequate information to offer direction, we include the issue in Section 7 at the end of this paper for the working group's consideration.

---

[26] The top end of the consumption scale for this question is perhaps deliberately imprecise, as it reads: "more than 15 cigarettes every week (around a pack a week)." This phrasing may produce ambiguity for the respondent, since cigarettes are actually sold in packs of 20 or 25 (33%-66% greater than the 15 cigarettes a week in the question).

The following are ten key questions about measurement related to smoking behaviour of which there is some discussion in the literature to guide us:

1. Are the items we are using meaningful to youth?
2. What information is available regarding item reliability and validity?
3. Should biochemical validation be recommended for both real-world and research evaluation studies?
4. How should we measure quitting in youth?
5. How should we measure smoking reduction in youth?
6. How should we measure abstinence in youth?
7. How is the concept of slips relevant in youth smoking cessation?
8. What are self-report error issues that should be considered in measuring youth smoking cessation outcomes?
9. What are the ideal post-intervention follow-up intervals for use in real-world and research studies on youth smoking cessation?
10. Which approach should be considered for real-world and research studies: intent to treat, or attrition analysis?

When considering these questions, it is important to keep in mind some of the issues related to youth smoking that can make the measurement of cessation-related behaviours particularly challenging. These challenges include:

1. *Non-daily smoking:* Non-daily smoking is more common among youths than adults; in comparison to adults, some youths may smoke on relatively few days in a given month.
2. *Irregularity:* Many youths exhibit an irregular or sporadic pattern in terms of number of cigarettes smoked in a day, whether or not they are daily smokers.
3. *Low dosage:* The number of cigarettes consumed per unit time may be very low. Youths may also be more likely to smoke only partial cigarettes, or to share a cigarette with someone else, so asking for "number of cigarettes smoked" can be problematic.
4. *Identification:* A youth's identity as a smoker may be in flux, and may not correspond to actual smoking behaviour.
5. *Intentions:* A youth's intentions behind observed changes may not be what we would expect. For example, a common thread in the literature is the observation that youths may "quit" for a period of time with the intention to re-start when the situation prompting the quit changes.
6. *Constraints on behaviour:* Youths may not always be free to choose when they can smoke. An implication of this is that while we are only recommending behavioural outcome items, the evaluator may need to track other contextual items in order to attribute changes in behaviour to the intervention.
7. *Substitutions for Smoking:* Measurement of substitute behaviours should be considered. For example, switching to smokeless tobacco use or use of marijuana, and other problems among teen smokers such as weight management or mood.

While we need to recognize the challenges above, we also need to keep in mind that not all youth smokers exhibit non-daily or irregular smoking patterns or are constrained in their opportunities to smoke. For example, there are certainly many adolescent smokers who have a more or less usual amount that they smoke on a daily basis. What is known, however, is that the proportion of such

regular daily smokers is significantly less among youths than in the adult population. Furthermore, because adolescence is a time of maturation and change, patterns of smoking for any individual youth is subject to change over time. Such challenges represent common threads in the literature, and need to be kept in mind for the discussion items that follow.

In the sections below, we raise issues related to quitting, abstinence, and slips. For clarity, in this paper we define quitting as being the relatively short-term process of stopping smoking; abstinence as a smoke-free state that can extend for a longer period of time following a quit; and slips as episodes of limited smoking following a quit from which an individual may recover to avoid failure of the cessation attempt.

6.1 Are the items we are using meaningful to youth?

In order for outcome measures to be meaningful, they must match the patterns of the target group (Mermelstein et al., 2002). O'Loughlin observes that we do not understand how youths conceptualize smoking concepts, or what terms are meaningful to them.[27] Mermelstein (2002) and O'Loughlin (2005) are consistent with our general findings from the literature review, the common thread being that the terms used to describe smoking-related behaviours and changes in smoking pattern may have significantly different meanings for adolescents than for adults. This is one of the fundamental reasons why measures developed and validated for an adult population cannot simply be transferred uncritically for use among adolescents, and why it is important to attempt to validate smoking measures within the intended audience.

An additional challenge in developing meaningful measures for youths is that youths are by no means a developmentally homogeneous group. The adolescent period is one of great development and change; the conceptualizations and meanings of 12 year olds and 18 years olds may have less in common than between 18 years olds and adults. Mermelstein (2003) and O'Loughlin (2005) make the point that items should be consistent with the language youths use to describe cigarettes and smoking (Backinger et al. (2003b) refer to youth "vernacular"). However, because youth culture is dynamic and youth terminology can change so rapidly, efforts to keep terminology current would present a formidable challenge. Modifying terminology to use the most current or geographically appropriate terminology is likely not desirable in any event, since doing so would result in a series of measures using different metrics.[28] O'Loughlin (2005) and others reason that an ideal set of questions does not currently exist in the literature (a common view among the experts we talked with), but she is confident that a workable set can be developed if the required developmental research is implemented. Bercovitz et al. (1999) found that "youth regard smoking and what it means to be a smoker more in terms of the particular behaviours and/or the situations in which smoking takes place than in terms of the number of cigarettes smoked per day, or the frequency of smoking per month," and that "there needs to be more work on the development of items that can capture the differing situations in which [youth] smoke" (p. 12).

6.2 What information is available regarding item reliability and validity?

As described earlier in this paper, testing to assess the reliability and validity of items used is seldom reported in the research studies we reviewed. Published articles on methodological issues point out the inherent problems with the assessment of the test-retest reliability of smoking state or consumption level in a youth population due to irregular or unstable smoking patterns and the high prevalence of

[27] Personal communication, March 22, 2005.

[28] One option to consider would be using "standard" terminology, but including youth vernacular in parentheses. The effects of this approach on responses would need to be assessed with research investigating the comparability of responses from different cultural or geographic youth populations using different vernacular terms thought to have equivalent meanings for the respective groups.

non-daily smoking among youth. Because consumption can be highly variable from one time period to the next, even if cigarette consumption is reported accurately, natural variability in youth smoking behaviour makes interpretation of test-retest data difficult. As Bercovitz et al. (1999) point out, "it is not possible to separate inconsistencies from day to day or week to week in actual smoking patterns from the lack of reliability of the item" (p. 5).

Bercovitz et al. (1999) conducted a careful evaluation of the reliability of questions used to assess youth tobacco use in Canada. Among their findings was that, in general, test-retest agreement was higher for older respondents than for younger respondents, although they did not offer an explanation for the association. They also found that responses to the last 7 days smoking recall question (the "wheel") were consistent with responses to a question asking on how many days respondent smoked in the last 30 days. They did not identify a valid measure for youth smoking cessation.

6.3 Should biochemical validation be recommended for both real- world and research evaluation studies?

Biochemical validation is often used in research studies to determine if a study participant has been smoking, by testing carbon monoxide levels in expired air or by testing body tissues for metabolites of nicotine.[29] There are a number of problems with the use of biochemical validation in youth.

Sussman (2002) makes note of a "recent debate concerning applicability of biochemical validation to adolescent tobacco users considering that teens may metabolize nicotine differently (e.g., more quickly) than adults" (p. 51). Consequently, the usefulness of biochemical validation is limited to attempting to validate whether or not an individual has smoked in recent past; it is not useful for validating reported consumption level, other than as a binary indication of smoked/did not smoke. The potential problem presented by youths metabolizing nicotine differently than adults may be further compounded by the fact that many youths may have irregular smoking patterns, and expired CO's short half-life makes it unsuitable as a validation measure for non-daily smokers. Paul McDonald feels that salivary cotinine may be a tractable measure to confirm smoking status; however, using cotinine is quite expensive and is therefore not practical outside of research studies.[30]

Finally, biochemical validation measures may be seen to be invasive. Sussman et al. (1999) note that the use of biochemical validation may interfere with the development of trust between the researcher (or program facilitator) and participants. The use of biochemical validation measures could also be expected to present challenges for participant and parental consent for community interventions involving youth, especially for programs in schools.

In summary, biochemical validation is unlikely to be practical in real-world interventions, but using longer half-life measures such as cotinine is appropriate in research settings.

6.4 How should we measure quitting in youth?

The impact of variations in the measurement of behaviours related to cessation is non-trivial. Although quitting may seem like a relatively straightforward concept to some, how and when the concept is measured can produce widely differing results. For example, Backinger et al. (2003) note a study by McDonald in which quit rates varied from 2.5% to 17% depending on how quitting smoking was assessed. Table 3 below gives some sense of the range of definitions used to define quitting.

---

[29] Biochemical validation is a complex issue; we have left discussion of issues—such as strengths and weakness of various measures, their respective sensitivity and specificity, marker duration, ease of use, and cost—for the working group.

[30] Personal communication, March 18, 2005.

Table 3: Definitions related to quitting

Article	Definition
Mermelstein et al 2002	Not smoking with the intention to quit. (They do not articulate a minimum abstinent time period, but do recommend using both 7 and 30 day abstinence rates.)
Ockene et al. 2000	7 consecutive days of non-smoking (This is considered a consensus measure citing USDHHS 1989 and Shiffman et al. 1986.) It is important to note that this criterion is intended for daily smokers, and is not specific to adolescents.
Ossip-Klein et al 1986	24 hours non-smoking; 48 hours is a more conservative measure
Stanton et al. 2002	24 hours non-smoking
Sussman et al. 1999	Distinguishes between *quit* and *cessation*. They find definitions of quit in reviewed studies to be highly varied and ambiguous. Most studies reviewed were concerned with daily smoking, and defined cessation as no smoking in the last 7 or 30 days.
Sussman 2001	30 days abstinent at follow-up
Sussman 2002	"Immediate quit": quitting tobacco for at least 7 days immediately post-test

Many of the definitions in Table 3 are based on single days or on 7 or fewer consecutive days of non-smoking. Such short time periods are of limited use for adolescents, who commonly exhibit non-daily or sporadic smoking patterns. Furthermore, some adolescents may go for periods of many days without smoking as part of their normal pattern of smoking. Sussman (2002) notes that many adolescent smoking cessation trials use smoking at least once a month for the baseline eligibility criterion; in this case, quit definitions such as 7 consecutive days of non-smoking can create confusion, since some low-frequency individuals will be considered quit at baseline, even though these non-smoking days may not have been quit attempts. Different measures for quit may be appropriate for daily versus non-daily smokers. In discussion with experts, an emerging popular opinion appears to be 30-day abstinence as a criterion for a successful quit for youth smokers.

It is widely recognized that the concepts of quitting, and consequently of a quit attempt, do not necessarily have the same meaning for youths as they do for adults. This makes developing valid measures and interpreting data collected using those measures problematic. Youths may be using the term to describe an intentional period of abstinence, but in which their intention is to quit for a period of time and ultimately resume smoking in the future, rather than intending to stop smoking permanently. While a number of the quitting questions in the reviewed questionnaires included a qualification limiting the response to only those attempts in which the respondent intended to quit permanently, we found that unqualified use of the term quit was still common.

6.5 How should we measure smoking reduction in youth?

The measurement of smoking reduction has some problems similar to identifying quitting, as described in the section above. For non-daily, irregular smokers in particular, the time frame used in the consumption recall questions may not be long enough to collect sufficient information from which to estimate a useful average consumption level. Setting aside any issues related to potential recall error or bias,[31] because many adolescents have irregular and variable smoking patterns, accurate short-term recall of cigarettes consumed could describe a period in which either more or fewer cigarettes were consumed than the true long-term average would indicate.

While one solution might seem to lie in increasing the length of the period in which respondents are asked to describe their smoking consumption, this approach rather quickly reaches the practical limits for accurate respondent recall. Prospectively tracking the consumption of participants prior to entry into a program would be both impractical to implement for a community intervention, and difficult to

[31] See Section 6.8 for a discussion of issues related to error in self-reported smoking consumption among youths.

design in such a way as to avoid changing participant behaviour, thereby failing to obtain an improved baseline measure of consumption. A commonly accepted measure in both the published literature and in communication with the experts contacted is consumption over the last 30 days, in combination with asking about consumption on each of the last 7 days.

6.6 How should we measure abstinence in youth?

Mermelstein et al. (2002) argue that 30-day abstinence measures are more appropriate for adolescents given the variability of smoking patterns exhibited, but that researchers should collect and report both 7- and 30-day prevalence measures. Mermelstein (2003) points out that 7-day prevalence may not be a sufficient time period for low frequency non-daily smokers.

Furthermore, both point prevalence and prolonged abstinence should be reported; point prevalence rates alone do not capture relapse, as they can include individuals who cycle in and out of abstinence. (This is of particular relevance for adolescent populations, given the frequency of irregular or sporadic smoking patterns observed in the population.) It is worth noting that some of the evaluation questionnaires reviewed ask about the length of abstinence not continuously from the intervention's designated quit date but from the date the respondent last quit or last smoked a cigarette. Either phrasing provides a measure of the length of continuous abstinence.

A number of authors felt it important to include intention and self-label in the definition of abstinence (e.g., Mermelstein et al. 2002); that is, did the individual initiate and attempt to maintain a period of non-smoking with the intention to voluntarily quit for the long-term? It may be necessary to clarify the objective of quit events for youth, and discriminate between periods of abstinence in which 1) they quit with the aim of achieving life-long abstinence, 2) they voluntarily quit for a certain period of time (e.g., a sporting event or season, duration of a relationship with a non-smoker, etc.) with the intent of eventually resuming smoking, or 3) their abstinence period is a result of a forced quit (e.g., reasons of health, parents, or school)?

Smoking cessation programs in the community typically use a single-group design. Sussman et al. (1999) note that one difficulty in assessing the success of adolescent cessation programs is that even when quit rates are high, there is uncertainty and a wide range of values regarding estimates of the natural quit rates in the population. While acknowledging that spontaneous quit rate is difficult to measure, Mermelstein (2003) states that there are relatively low spontaneous quit rates among adolescents. This is consistent with O'Loughlin's[32] preliminary observations regarding quit rates in her study of youth smokers in Quebec. Mermelstein (2003) refers to a 1996 study by Stanton et al. as one of the better studies addressing spontaneous quit rates. One approach that has been used is to compare quit rates observed in a single group study against a "standard of acceptability" drawn from the literature. For example, an evaluation of the KTN program in British Columbia (Lovato et al., 2002) compared the program's quit rate to the average quit rate of all programs reported in Sussman's 2002 review of youth smoking cessation trials. An evaluation of Health Canada's Q4L program compared the program's quit rate to age-specific quit rates in the population estimated from CTUMS data.[33]

6.7 How is the concept of slips relevant in youth smoking cessation?

A slip is an isolated event or series of events in which an individual who is attempting to quit smoking resumes smoking for a short period of time followed by a return to non-smoking. A slip is intended to

[32] Personal communication, March 22, 2005.

[33] Personal correspondence from Julie Green, Office of Programs and Mass Media. Tobacco Control Programme, Health Canada, April 06, 2005.

allow an individual to have short episodes in which he or she smokes without the behaviour being considered a failure of the cessation intervention. Such events are also termed lapses in the literature. A significant gap in our understanding exists around understanding the cessation trajectory for youth, and what slips (or "slip-like") behaviour might look like. The impact of slip events on the long-term prospects for the success of a youth's quit attempt is not known.

There is little common ground regarding how to define a slip for youth smokers. It is also widely recognized that such a definition is likely not meaningful for many non-daily youth smokers, as their smoking patterns can be highly irregular, with gaps of days between cigarettes not being unusual. For individuals with very sporadic low frequency smoking patterns, their usual pattern of smoking could resemble a continual pattern of slips, even though they are not making any efforts to quit smoking. If such individuals were to try to quit, what would a meaningful slip look like for them?

If measurement items do not explicitly ask about continuous abstinence since the intervention quit date, detection and interpretation of any evidence of slip events is difficult or may not be possible. Typically, items do not ask about remaining smoke-free since the time of quitting, but ask about a specified time period such as the last 30 days. None of the reviewed questionnaires included items that capture the slip's duration or the number of cigarettes consumed during a slip event, nor is information collected about the durations of the abstinence periods that precede and follow a slip event. Furthermore, there is no consensus on what constitutes a slip. Definitions include Ossip-Klein et al.'s (1986) definition for adults of up to 7 consecutive days with at least one puff per day. Hughes et al. (2003) use a less precise definition of "a few cigarettes" following a quit (p. 15),and Backinger et al. (2003b) provides a youth-specific definition of smoking on three consecutive days or smoking on at least one day per week for three consecutive weeks.

The terms slip or lapse are contrasted with relapse, which is characterized in the literature as a period of smoking following abstinence represented by a longer-term resumption of regular smoking. Relapses are considered to constitute treatment failure. Mermelstein (2003) argues that we currently know very little about adolescent relapse, and need to know more. She also states that a significant part of this problem is that we need to know more about how adolescents interpret relapses. These problems with terminology and concepts can be traced in large part to the transfer of the concepts from the adult smoking literature. Terms such as slip and relapse were developed originally to describe events for the adult population, in which individuals were typically required to be daily smokers in order to qualify for inclusion in the intervention being evaluated. Concepts such as these cannot be transferred uncritically from the adult literature for use with adolescents. For example, how can a researcher distinguish between a slip and a relapse if the smoker never had "regular" pattern of smoking to resume? If the smoker's pattern was sufficiently infrequent and highly irregular, it may not be easy to determine whether an individual has resumed their usual pattern of smoking, or is abstinent, or slipped, or relapsed. For such reasons, while we can appreciate that slips are an important concept that provides individuals who are attempting to quit some room for minor infractions, the concept may not be measurable. The term slips is used in the adult literature to characterize "near failure" events; however, in a youth population, such events may actually be relatively unexceptional and may in fact be consistent with a pre-existing pattern of irregular non-daily smoking.

6.8 What are self-report error issues that should be considered in measuring youth smoking cessation outcomes?

We identified six concerns related to self-report error in the measurement of youth smoking behaviours. First, the error may result from deception on the part of the respondent. Second, it may be difficult to recall consumption patterns accurately over a period of 30 or even 7 days; that is, the recall period exceeds the respondent's retrospective memory capability. Third, it may be very difficult for respondents to calculate a "usual" amount smoked if their smoking pattern is irregular. The magnitude of this problem increases with increasing irregularity of the respondent's smoking pattern.

Fourth, as Paul McDonald has observed, individuals have a tendency to characterize longer term past patterns based on the patterns they have experienced in the more recent past.[34] This recent recall bias results in a consumption recall averaging problem for individuals with highly variable consumption patterns, in which individuals with irregular smoking patterns, who smoked more than their (actual) usual amount, will over-estimate their usual consumption level, and vice-versa. McDonald is exploring whether it is possible to estimate a correction factor for recalled consumption level, based on responses to “amount usually smoked” and “amount smoked yesterday” responses. Fifth, individuals may have difficulty accurately recalling the timing of events of interest, such as quit dates. Finally, individuals may exhibit a selective (but unintentional) memory, so that not all smoking events of interest (e.g., slips) may be recorded for the evaluation. An important observation to make here is that biochemical validation measures[35] can really only address the first and last sources of error.

6.9 What are the ideal post-intervention follow-up intervals for use in real-world and research studies on youth smoking cessation?

How and when post-intervention measures are employed should be considered, as this will affect the information gathered. The timing of the first and subsequent follow-up tests is a significant issue that can affect results, perhaps most importantly with respect to the timing of the first post-intervention follow-up. For this reason, consistency is highly desirable for both what measures are used and when the measures are used relative to the intervention.

Mermelstein et al. (2002) argue that frequent follow-ups will help gather information on the process of behaviour change, relapse, and seasonality effects. A commonly recommended interval for follow-up among adolescent cessation studies is six months (see Table 4), although there is a large amount of variation around this value. For example, five studies reviewed by Sussman (2002) had follow-up for only 1 month, while 15 of the studies he reviewed conducted follow-up for 1 year or more, including one with a 3-year follow-up. Backinger et al (2003b) note that sample retention becomes a more significant factor as the follow-up interval increases. Milton et al. (2004) recommend measuring program outcomes post-intervention and again at 6 and 12 months post-intervention to determine if the intervention had a lasting effect. The follow-up timing for adult programs is based on the survival curve for smoking relapse. This relapse survival curve is not well understood for youth; nonetheless, the recommended 3- and 6-month timing intervals for post-intervention follow-ups is based on our understanding of relapse rates for adults, with the hope that they will also be relevant for youth smokers (Backinger et al., 2003). Generally speaking, it will be more difficult for community interventions to conduct longer-term, multiple follow-up than for research studies. There is no clear consensus regarding precisely when the immediate post-test should be conducted: for example, Sussman’s TNT program conducts the immediate post-test 7 days after the last session; the HYSQ protocol specifies “within two weeks” of the last session, and follow-up intervals used in the North American Quitline Consortium vary by quitline. It is important to recognize that, in the real world, follow-up often does not take place at a specific point in time, but rather occurs over a period of time as program staff work to recontact and get responses from program participants. In longer studies, this follow-up work may occur over a period of months, to specific target date objectives need to be tempered with a recognition of the limitations often faced with follow-up in the real world.

---

[34] Personal communication, March 18, 2005.

[35] Alternative approaches to addressing respondent recall error include the use of smoking diaries, ecological momentary assessment, or the use of email or cellular phone text messaging. However, as these approaches constitute a significant methodological departure from standard survey methods, they are not addressed in this paper (not to diminish the potential value of alternative data collection approaches such as these).

Table 4: Follow-up interval recommendations

Article	Definition
Backinger 2004	Recommends follow-ups up to 6 to 12 months after end of intervention.
Backinger 2005[36]	Immediate post-test 4 weeks after intervention.
Garrison et al 2003	Follow-up of at least 6 to 12 months is essential
Mermelstein et al. 2002	3, 6, 9, 12 months. Includes both point prevalence and prolonged abstinence to that point.
Milton et al. 2003	At least 6 months after the end of intervention
Milton et al. 2004	6 and 12 months post-intervention

The articles we reviewed from the published literature typically describe follow-ups being recommended at various intervals post-intervention, but precisely when the interval should begin is often unspecified or vague. The follow-up interval could conceivably start from any of these times: the last session of the intervention; the target quit date; a participant's actual quit date; the first session of the program; or when a participant entered the program. Consistency in follow-up measurement timing would facilitate comparison and interpretation of evaluation findings across programs; consequently, this is an issue the working group may wish to consider.

There is a general view that longer follow-up intervals and multiple follow-ups are preferred, with 3 and 6 months emerging as a recommended standard to use in community interventions. Mermelstein (2003) recommends that 3 month follow-ups be conducted out to 6 months post-intervention, and if possible up to 12 months follow-up. It is widely recognized, however, that conducting any post-program follow-up can be very difficult, and program facilitators may not have the resources required to execute follow-up surveys months after the intervention has occurred. In addition, the constraints of operating within the school year and the school environment may preclude achieving this objective for school-based interventions. For research studies, longer-term follow-ups in the range of 1 year post-program were supported, in addition to follow-up at 3 and 6 months. The relapse survival curve is not well understood for youth. The 3- and 6-month timing for post-intervention follow-ups is based on our understanding of relapse rates for adults. The exact start period for follow-up intervals varies in the literature.

6.10 Which approach should be considered for real- world and research studies: intent to treat or attrition analysis?

Not everyone who starts a smoking cessation program completes it and can be contacted to complete a follow-up questionnaire. A methodological question in outcome evaluation is whether individuals who do not complete a program differ in a systematic way from individuals who remain in the program. When calculating quit rates at the end of a program, intent-to-treat (ITT) analysis considers all individuals lost to follow-up to be program failures; that is, it is assumed they did not successfully quit, and they remain in the denominator when calculating the program's quit rate. Attrition analysis is an alternative and less conservative approach, in which individuals who are lost to follow-up are also removed from the denominator when calculating the program's quit rate. Characteristics of the individuals lost to follow-up and those remaining in the program may be assessed to determine if there are any significant differences in the characteristics of the two groups.

Many authors (Milton et al., 2003; Backinger et al., 2003, Garrison, 2003) recommend calculating intervention success rates using ITT analysis. These authors point out that one cannot assume that individuals who either did not complete the intervention or were lost to follow-up actually quit smoking. For example, as Garrison (2003) notes, individuals who are lost to follow-up might be more

[36] Email communication March 29, 2005

likely than those who continue with the program to feel that the intervention is not beneficial; they may experience stronger desires to resume smoking; or they might have fewer social supports. An additional rationale for using ITT, offered by Malarcher, is to minimize the potential for problems associated with programs manipulating data to enhance reported quit rates.[37]

However, if loss to follow-up is large, using ITT analysis can significantly under-power a study. McDonald is of the opinion that ITT analysis is not helpful in a population intervention, and feels a more reasonable approach is to look at those participants one could reasonably be expected to contact, and compare their characteristics with the characteristics of those who were lost to follow-up.[38] Milton et al. (2003) also note the importance of assessing the similarity in characteristics between those who were and were not re-contacted following an intervention.

ITT analysis may be most appropriate for research studies, but in practical application in a community setting with a youth population, ITT is not always feasible.

7 RECOMMENDATIONS

The following set of recommendations focus on specific core indicators, measurement items, issues related to outcome measurement, and specific questions that should be addressed by the working group, who will be using this background document to move this initiative to its next phase.[39] A desirable goal would be to ultimately develop a set of guidelines for measurement of youth smoking behaviours, cessation intervention characteristics, and guidelines for the reporting of these measures.

7.1 Identify core indicators

We recommend that the working group adopt the following three indicators to evaluate behaviour change in youth smoking cessation interventions:

I. *Abstinence* (cessation or zero consumption since intervention),

II. *Reduction* (consuming fewer cigarettes per unit time than at baseline)

III. *Quit attempts* (actual attempts made to quit smoking)

7.2 Identify items for evaluation

To reiterate, the purpose of this background paper is to identify a set of measurement items for the evaluation of youth smoking cessation interventions including:

1. Validated core items
2. Items that show promise but need more testing
3. Non-core (optional) items, and
4. Gaps in the items currently available.

---

[37] Email communication April 20, 2005.

[38] Personal communication March 18, 2005.

[39] Phase II of this project will involve a series of qualitative focus groups with youth smokers to better understand how youth interpret the words and concepts that researchers use when trying to measure youth smoking cessation. A workshop with experts in youth tobacco control to consider these recommendations and the results of the youth focus groups will also be part of phase II.

In the sections below we enumerate, for the working group's consideration, the specific items for each of the above categories. While many of these items are not ideal, these recommendations are made based on our assessment of the best available items at the present time.

7.2.1 Validated core items

The six items in Table 5 are recommended as part of a set of core measurement items for the evaluation of youth smoking cessation interventions. Taken together, these items permit the estimation of cigarette consumption level, over both the last 7 (item 1) and 30 (items 2 and 3) days. Determination of daily or non-daily smoking status can be made using either 7-day (item 1) or 30-day (item 2) time periods. Patterns of smoking can be described over the course of the previous 7 days with item 1.

Items 1 through 3 also can be used to determine abstinence over the previous 7 (item 1) or 30 (item 2) days. Continuous abstinence can be assessed using item 6 for individuals who are currently abstinent. Information about the number and duration of failed quit attempts can be collected with items 4 and 5 respectively.

The recommended items are selected from the Helping Young Smokers Quit (HYSQ) project, because the questionnaire is comprehensive, question wording is among the most clear of the questionnaires reviewed, and the items consistently specify the behaviour of interest, the time period of interest, and what specific quantity is being asked for. The specific question numbers for each recommended item are identified. For some items, we see opportunities to further improve the clarity of the selected questions, and provide notes accordingly.

Table 5: Validated core items recommended for consideration by the working group

Validated core items[40]	Source	Cessation	Reduction	Quit attempts
1. 7-day prevalence: frequency and quantity smoked on each of last 7 days • *Recommended wording:* "Please think about your smoking during THE PAST WEEK. Using the form below, please write in how many cigarettes you smoked on each day." Form: "During the past week, how many cigarettes did you smoke each day?" Columns labelled "6 days ago" – "Today." Instructions: "If you did not smoke on a day, enter '0.' If you smoked, but less than 1 cigarette, enter '1.'" Respondents are instructed to work backwards from "today." • *Telephone administration:* The CTUMS implementation of the "wheel" is recommended.	HYSQ Baseline Q38	•	•	
2. 30-day prevalence: number of days smoked in last 30 days • *Recommended wording:* "During the past 30 days, on how many days did you smoke cigarettes?"	HYSQ Baseline Q35	•	•	
3. 30-day prevalence: number of cigarettes smoked on days smoked in last 30 days • *Recommended wording:* "During the last 30 days, on the days you smoked, how many cigarettes did you smoke per day?"	HYSQ Baseline Q36	•	•	

[40] Validation of measures is often not reported. See Section 5.

Validated core items[40]	Source	Cessation	Reduction	Quit attempts
4. Number of quit attempts since intervention • *Recommended wording:* "Since you started the stop-smoking program, how many times have you stopped smoking for one day or longer because you were trying to quit smoking?"	HYSQ Follow-up Q39			•
5. Longest quit attempt since intervention • *Recommended wording*: "Since you started the stop-smoking program, what is the longest number of days in a row that you have gone without cigarettes?"	HYSQ Follow-up Q37			•
6. Time since last cigarette • *Recommended wording*: "When was the last time you smoked a cigarette, even one or two puffs?"[41]	HYSQ Follow-up Q10	•		

7.2.2 Items showing promise, needing more testing, and suggested modifications to existing items

We did not identify any items requiring more testing that differ substantially from the recommended core items above; however, a few items have been identified that may be able to be improved upon through modification. We identify four such items in Table 6 below. For the first item, we propose modifying the wording to explicitly instruct respondents to estimate their average consumption over the reference period. In the second and third items, we propose modifying the wording to make it clear that the question is limited to cigarettes and to restrict the question to non-smoking periods in which the intention was to quit smoking permanently. In the last item, we propose modifying the response field to permit recording the response using units other than days. For follow-up questionnaires in which longer quit periods could reasonably be expected, weeks or months would be more appropriate time units to use.

Table 6: Items showing promise needing more testing, and suggested modifications

Item	Cessation	Reduction	Quit attempts[42]
1. Modify 30-day consumption question to allow concept of average consumption level over the time period of interest. • *Wording for initial consideration*: "During the last 30 days, on the days you smoked, how many cigarettes did you smoke per day on average?"[43]	•	•	
2. Modify quit attempt question for youths who smoke every day to distinguish between situational quits and intentions to quit "for good," and to ensure question exclusivity to cigarettes: • *Wording for initial consideration*: "Since you joined[44] the stop-smoking program, how many times have you stopped smoking cigarettes for one day or longer because you were trying to quit smoking for good?"[45]			•

[41] We recommend this question wording only if the question is only asked of individuals who have quit or are in a quit attempt, and if the response is modified to something like, " ___ weeks, ___ days ago OR date last smoked: ______)."

[42] Suggested modification for any quit attempt questions: eligibility for a quit attempt is non-smoking for at least 24 hours *with the intention of trying to quit smoking for good.*

[43] Also consider alternative terms such as *usually* or *typically*.

Item	Cessation	Reduction	Quit attempts [42]
3. Modify quit attempt question to make it more appropriate for youths who *do not* smoke every day to distinguish between situational quits and intentions to quit "for good," and to ensure question exclusivity to cigarettes. Ask this question only of non-daily smokers: • *Wording for initial consideration*: "Since you joined the stop-smoking program, how many times have you stopped smoking cigarettes for one week or longer because you were trying to quit smoking for good?"			•
4. Modify the HYSQ "longest quit attempt since intervention" question (follow-up to Q37) to permit the respondent to give their answer in units in addition to days (e.g., months, weeks, and/or days).[46]			•

7.2.3 Non-core (optional) items

No suitable optional items for the evaluation of smoking behaviour change in youth smoking cessation programs were identified, other than those already described.

7.2.4 Address gaps in current items

For our purposes, we consider gaps to be areas in which there is either no consensus or there is no item currently developed to measure a concept of interest. Because youth smoking behaviour, and how to appropriately measure it, remain poorly understood in comparison to adult smoking, there are a number of significant gaps relevant to the measurement of youth smoking cessation intervention outcomes. One of the themes of this section will be the gaps that are unique to measurement of behaviours for youth; that is, measures may be reasonably well developed in items intended for the adult population, but simply do not work or are somehow invalid when applied to the youth population. In this section, we choose to focus on measurement gaps as they apply to quit attempts, slips, relapse, and our understanding of the situational contexts that influence when and how much youths smoke.

1. *Standard phrasing for quit attempts*
    - Developing standard phraseology for questions asking about quit attempts, in which the respondent is instructed to only include those events in which their intention was to quit either temporarily or permanently.
2. *Applicability and measurement of slips*
    - Undertake further research to determine the applicability and nature of slip events for youth.
    - Identify strategies for measuring slip events, including items to describe how "long" and "deep" the event(s) are.

---

[44] We suggest considering the wording "…joined the stop-smoking program…" in place of the potentially awkward "…started the stop…" This comment also applies to the question for non-daily smokers below.

[45] We suggest undertaking some validation testing to determine the most appropriate phrasing. Quitting "for good" has a potential unintended implication of an opposition to an alternative of quitting "for bad." Perhaps terms such as *permanently* or *forever* would be preferable. This comment also applies to the question for non-daily smokers below.

[46] An alternative option, which would necessarily be exclusive, would be to record the quit date (and failure date, if applicable) if the respondent is able to recall this information. More testing would be required to assess the viability of such an approach for this item.

3. *Understanding relapse*
    - Consider further research to improve our understanding of the patterns, timing, and predictors of adolescent smoking relapse events.
4. *Situational context of youth smoking*
    - Consider further research to improve our understanding of associations between physical, social, temporal, and emotional situational contexts and smoking patterns for youth smokers.

7.3 Recommendations related to outcome measurement for consideration by the working group

7.3.1 Validity

Address how standards of acceptability should be developed and applied for evaluation of community interventions that use single-group designs.

Develop guidelines for the implementation of biochemical validation and the use of control groups for youth smoking cessation research studies.

Further research should be considered to investigate youths' recall of cigarette consumption in the recent past.

7.3.2 Behaviour-specific issues

I. *Abstinence (cessation)*
- Adopting 30 days non-smoking as the criterion for a successful quit for youth smokers.
- Consider developing a standard for a successful quit based on three criteria being met: i) intention to quit permanently, ii) 30-days smoke free, and iii) self identifies non- or former smoker.

II. *Reduction*
- Record values as continuous data, avoiding imposed categorization in data collection forms whenever possible.
- Extend the time period used in consumption recall questions to 30 days, but continue collecting 7-day recall data concurrently.

III. *Quit attempts*
- Modify the definition of a quit attempt from minimum criterion of 24 hours abstinent to 1 week abstinent for non-daily smokers.
- Develop standard phraseology for questions asking about quit attempts, in which respondents are instructed to include only those events for which their intention was to quit permanently.

7.3.3 Intervention outcome evaluation design

Specify minimum pre- and post-test assessment for outcome evaluation. Both tests should use a core set of items that are directly comparable between pre and post tests.

Promote minimum 3- and preferably 6-month follow-up for youth smoking cessation interventions, and encourage longer follow-ups when possible.

7.4 Questions to be addressed in the next steps

1. Does the working group agree, in principle, with the core indicators recommended?
2. Does the working group agree, in principle, with the measurement items recommended? Are there revisions to the recommended items that should be made?
3. Does the working group feel comfortable recommending moving forward with the set of indicators/measures identified here? If not, what modifications are necessary?
4. What are the most useful processes to encourage adoption of the recommended indicators and measurement items? Is there opportunity to recommend policy requiring use of the measures at federal, provincial, and territorial levels, or perhaps as review criteria for cessation intervention funding applications? If yes, how can we best move toward such a policy? How would this be disseminated?
5. Should we move towards developing a standard (or "core") set of measures for intervention implementation and participant characteristics?
6. What background information on evaluation issues (design issues, process issues, implementation issues) should be distributed with core indicators and measures?
7. What is recommended as the ideal timing for collecting data post-intervention?
8. Should items related to temporary quits be included in the core items? If so, how should intention to quit temporarily versus permanently be measured?
9. What infrastructure is required to take this initiative to the next phase?
10. We do not recommend that community-based evaluations of youth smoking cessation interventions use biochemical validation as a practice standard. The working group should provide guidance regarding in what situations biochemical validation should be used.
11. Would it be worthwhile to put some effort towards developing workable measures for slips as they may apply to the youth population, or is a better understanding of this important cessation concept as it applies to youth smokers required before we can develop measures to assess it?
12. The advisory committee for this paper suggested we focus on behaviours; however, the working group may wish to consider the utility of including other types of questions (i.e., non-behavioural questions related to withdrawal) in future research.
13. A limitation of this paper is the exclusion of French language evaluation questionnaires. The extent to which French language questionnaires currently exist, or the issues associated with translating and validating French versions based on English language questions is not addressed in this paper, but is an issue the working group may wish to discuss.

In summary, this document provides a review on the "state-of-the-art" in measurement of youth smoking for the purposes of evaluating smoking cessation programs. This information will provide the basis for further discussion and consensus building at an expert workshop sponsored by the Canadian Tobacco Control Research Initiative which will take place in November 2006. It is hoped that the results of that workshop will contribute to the development of core indicators for evaluating youth cessation programs, an area in which there is currently a lack of evidence to inform best practices.

References

Backinger CL, Fagan P, Matthews E, and Grana, R (2003a). Adolescent and young adult tobacco prevention and cessation: Current status and future directions. *Tob Control*, *12* (Suppl 4), IV46-53.

Backinger CL, McDonald P, Ossip-Klein DJ, Colby SM, Maule CO, Fagan P, Husten C, Colwell B (2003b). Improving the future of youth smoking cessation. *Am J Health Behav, 27* (Suppl 2), S170-S184.

Bercovitz K, Brown SK, Leis A, Loiselle J, McDonald P, Poland B, Potvin L, Smith P (1999). *Evaluation of existing questions used to assess youth tobacco use*. Report to Laboratory Centre for Disease Control, Health Canada.

Copley TT, Lovato CL (2004). *Survey-based indicators of tobacco control strategies in Canada: A resource for national and provincial level decision-makers.* Report to the Canadian Tobacco Control Research Initiative.

Curry SJ (2003). Youth tobacco cessation: Filling the gap between what we do and what we know. *Am J Health Behav, 27* (Suppl 2), S99-S101.

Garrison MM, Christakis DA, Ebel BE, Wiehe SE, Rivara FP (2003). Smoking cessation interventions for adolescents: A systematic review. *Am J Prev Med, 25*(4), 363-7.

Hughes JR, Keely JP, Niaura RS, Ossip-Klein DJ, Richmond RL, Swan GE (2003). Measures of abstinence in clinical trials: Issues and recommendations. *Nicotine Tob Res*, *5*, 13-25.

Lovato CY, MacLeod-Williams H, Lisonkova S (2002). *Evaluation of Kick the Nic: A teen smoking cessation program.* Victoria, BC: Report to the BC Ministry of Health.

MacDonald G, Starr G, Schooley M, Yee SL, Klimowski K, Turner K (2001). *Introduction to program evaluation for comprehensive tobacco control programs.* Atlanta, GA: US Centers for Disease Control and Prevention.

Mermelstein R (2003). Teen smoking cessation. *Tob Control, 12*(Suppl 1), i25-34.

Mermelstein R, Colby SM, Patten C, Prokhorov A, Brown R, Myers M, et al (2002). Methodological issues in measuring treatment outcome in adolescent smoking cessation studies. *Nicotine Tob Res, 4*(4), 395-403.

Milton MH, Maule CO, Backinger CL, and Gregory DM (2003). Recommendations and guidance for practice in youth tobacco cessation. *Am J Health Behav, 27*(Suppl 2), S159-69.

Milton MH, Maule CO, Yee SL, Backinger C, Malarcher AM, Husten CG (2004). *Youth Tobacco Cessation: A Guide for Making Informed Decisions*. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention.

O'Loughlin J, Gervais A, Pabayo R, Schaeffer A (2003). *Background paper on the evaluation of youth cessation initiatives*. Paper prepared for the Youth Cessation Roundtable, January 22, 23, 2004. Ottawa: The Office of Prevention, Cessation and Education, Tobacco Control Programme, Health Canada.

Ockene JK, Emmons KM, Mermelstein RJ, Perkins KA, Bonollo DS, Voorhees CC, et al (2000). Relapse and maintenance issues for smoking cessation. *Health Psychol*, *19*, 17-31.

Ossip-Klein DJ, Bigelow G, Parker SR, Curry S, Hall S, Kirkland S (1986). Classification and assessment of smoking behavior. *Health Psychol*, *5*(Suppl), 3-11.

Shiffman, S., Shumaker, S. A., Abrams, D. B., Cohen, S., Garvey, A., Grunberg, N. W. & Swan, G. E (1986). Models of smoking relapse. *Health Psychol*, *5*(Suppl), 13-27.

Stanton WR, Smith KM (2002). A critique of evaluated adolescent smoking cessation programmes. *J Adolesc*, *25*(4), 427-38.

Starr G, Rogers T, Schooley M, Porter S, Weisen E, Jamison N (2005). *Key outcome indicators for evaluating comprehensive tobacco control programs.* Altanta, GA: US Centers for Disease Control and Prevention.

Sussman S (2002). Effects of sixty-six adolescent tobacco use cessation trials and seventeen prospective studies on self-initiated quitting. *Tobacco Induced Diseases, 1*(1), 35-81.

Sussman S, Dent CW, Lichtman KL (2001). Project EX: Outcomes of a teen smoking cessation program. *Addict Behav*, *26*(3), 425-38.

Sussman S, Lichtman K, Ritt A, Pallonen UE (1999). Effects of thirty-four adolescent tobacco use cessation and prevention trials on regular users of tobacco products. *Subst Use Misuse*, *34*(11), 1469-503.