NAQC Issue Paper

NAQC's Issue Papers aim to provide critical knowledge on important quitline topics and guidance for decision-making.

Measuring Quit Rates

I. Background

Why Do We Need to Measure Quit Rates?

A strong body of evidence supports the efficacy of telephone counseling for smoking cessation. This evidence comes in the forms of randomized clinical trials and quasi-experimental studies of real-world implementation involving over 30,000 patients. Several major reviews and meta-analyses have drawn favorable conclusions regarding the provision of phone counseling. Most recently, the updated Public Health Service Clinical Practice Guideline, *Treating Tobacco Use and Dependence: 2008 Update*, identified "quitline" counseling (telephone counseling that includes counselor-initiated calls or proactive counseling) as an evidence-based treatment that increased the odds of abstinence by approximately 60% or 1.6 (95% Confidence Interval 1.4-1.8) (Fiore, 2008). This body of research has supported a rapid expansion in public access to quitline services. At present all tobacco users in North America have access to quitline services. In the U.S., a single toll-free number (800-QUIT-NOW) has been established as a portal to state-based quitlines.

With this large body of evidence supporting the efficacy of telephone counseling for smoking cessation, it is reasonable to ask: Why do we still need to measure quit rates? It is important to consider the answers to this question as the reasons to continue measurement of quit rates may provide some insight into addressing methodological issues related to how these rates are measured and reported.

A framework for understanding the needs for ongoing evaluation of quitline quit rates is presented by Glasgow and others in their RE-AIM (Reach, Efficacy, Adoption, Implementation, Maintenance) framework. Reach refers to the percentage and characteristics of individuals who are affected by a policy or program. Efficacy refers to the measurement of a program's outcomes and can include assessment of behavioral outcomes (e.g. smoking cessation) as well as other participant-centered measures such as consumer satisfaction. Adoption refers to the proportional representativeness of settings (e.g. states, provinces, countries, or regions) that adopt a given policy or program. Implementation refers to the extent to which a program is delivered as intended (e.g. adherence to evidence-based counseling protocols). Maintenance measures the extent to which programs are sustained over time.

Within the RE-AIM framework, standard measurement of quitline outcomes is desirable for several reasons. First, the expansion of quitline services has been accompanied by a growth in the number of quitline service providers. While many of the services currently being offered are based upon or modeled after protocols proven in clinical trials, not all these services have been directly tested. A standard measurement of outcomes is therefore needed to assure that outcomes of these services are in the range that might be expected for an evidence-based treatment.

Second, this growth of quitline service providers creates the need for purchasers of quitline services (e.g. employers, states, etc.) to make choices between different potential providers. Standard measurement of program outcomes will help to inform purchasers and service providers during the selection process.

NAQC's Quality Improvement Initiative is made possible with funds from The Centers for Disease Control and Prevention, The American Cancer Society and The American Legacy Foundation. NAQC would like to acknowledge its Advisory Council members for their role in reviewing and approving this Issue Paper as well as the invaluable feedback from NAQC members during the review and comment phases of the process.

NORTH AMERICAN QUITLINE CONSORTIUM

Third, the existence of standard outcome measures is important to inform program stakeholders (e.g. health department administrators, legislators, etc.) and to argue for or justify continued funding. Stakeholders will also have a natural tendency to compare the outcomes of their services to that of their neighbors. In this process, standardization of outcome measures will help prevent unnecessary confusion and possible consternation among stakeholders. Finally, standardized evaluation of quitline services has the potential to facilitate identification and spur adoption of "best practices." The existence of standard measures may even support efforts to move "beyond maintenance" in helping to identify innovative strategies to optimize the effectiveness of quitline services that are discovered during program implementation.

The purpose of this paper is to provide recommendations to guide the assessment of program outcomes for quitline services. The main focus will be on developing standards for measurements of abstinence outcomes. Issues regarding the reach of quitline services are addressed in a companion NAQC Issue Paper, Measuring Reach of Quitline Programs, by Dr. Sharon Cummins. A summary and synthesis paper by Stephen Michael, MA, will address issues related to the relationship between outcomes, inputs and processes critical to improving quitline quality and provide insights into the further evolution of quitline services and evaluation. The reader interested in further background on quitline operation and evaluation is referred to the CDC document, Telephone Quitlines: A Resource for Development, Implementation, and Evaluation (Centers for Disease Control and Prevention, 2004).

Why Do We Need a Standard Measure of Quit Rates?

In the mathematical sense, the determination of quit rates for quitline services is straightforward. Like all other rates, abstinence rates are composed of a numerator and a denominator. In this case, the denominator should represent the number of people who receive (or should have received) quitline services. The numerator represents the number of people who received quitline services and who quit using tobacco.

Abstinence Rate	=	Number of individuals who received quitline services who stopped using tobacco / Number of individuals who receive quitline services

What soon becomes clear is that there are many different ways to define the numerator and denominator in this calculation, and these choices can have a dramatic influence on the calculation of the abstinence rate.

The major questions involving definition of the denominator involve what constitutes "receipt" of quitline services. This issue is addressed in detail by Dr. Cummins in the above cited NAQC Issue Paper, *Measuring Reach of Quitline Programs*. Some of the questions we must answer are:

1. Should this denominator include all callers?
2. Should the denominator be limited to current tobacco users?
3. Should the denominator be limited to tobacco users who are interested in quitting in the near future?
4. What level of engagement with the quitline should occur before the individual is considered to have "received" quitline services (and be included in the denominator)?
5. What level of informed consent should be obtained from callers in the conduct of outcomes evaluation and how should this influence the denominator?

The major questions involving definitions of the numerator involve the definition and measurement of abstinence from tobacco. Some of the questions we must answer are:

1. Who should be included in the follow-up evaluation of quitline outcomes and how should these individuals be selected?

2. When should follow-up evaluation be conducted?

3. How should abstinence be defined at the follow-up (e.g. how long a period of abstinence should be counted as a successful outcome, how should individuals be counted who are not reached on follow-up, etc.)?

4. What questions should be used to assess abstinence?

Finally, we recognize there are important issues that influence the definition of the numerator and denominator in the calculation of abstinence rates. The central issues here are the treatment of missing data and how to count individuals who are not reached as part of the follow-up evaluation.

Why Now?

In many respects, this NAQC Issue Paper is a continuation of work begun by the North American Quitline Consortium in its creation of a Minimal Data Set (MDS) for intake and follow-up across quitlines (NAQC, 2005). The MDS includes items about reasons for calling, how callers heard about the quitline, tobacco use history, quitting history, current tobacco use, nicotine replacement therapy (NRT) use, caller demographics, use of other stop-smoking assistance and satisfaction with quitline services. The items were accompanied by a list of administrative data to collect. Adoption of the MDS was voluntary, and quitlines were encouraged to begin using the MDS by October 2005.

In 2007, NAQC re-formed the Minimal Data Set Workgroup to assess the extent to which North American quitlines had implemented the MDS, possible item revisions and areas in which NAQC might support quitlines in using the MDS (NAQC Minimal Data Set Workgroup, 2008). While this evaluation found a high degree of implementation of intake and follow-up items, there was considerable variation in how individual items were implemented across quitlines. In addition, the assessment identified a lack of standardization of follow-up protocols across quitlines. This variation in follow-up poses considerable challenges to standardized reporting of program outcomes. Some of the highlights in variation of follow-up are:

1. Approximately 20% of quitlines did not conduct follow-up assessments at the seven months time point recommended by the MDS (six quitlines or 11% conducted follow-up at other time periods, and four quitlines or 7% did not conduct follow-up at all).

2. Variation existed in the types of individuals being included in follow-up evaluations. While most quitlines conducted follow-up evaluation with callers over the age of 18, others limited follow-up to individuals who set a quit date or to callers who completed the counseling intervention.

3. Only slightly more than half of quitlines (56%) implemented each of the MDS follow-up questions on average exactly as worded or with minor working changes.

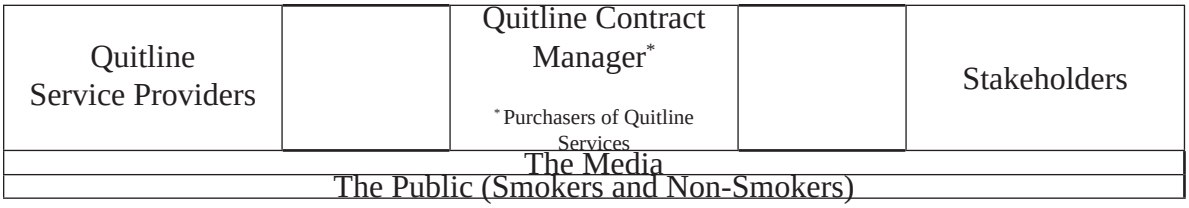
4. On average, nearly half of quitlines (40%) used response categories for each follow-up question (e.g. for assessing duration of abstinence) that differed from MDS recommendations. Less than one quarter (23%) did not use the MDS response categories at all for each question, and 14% used response categories for each question that could not be rolled up to be consistent with MDS.

This variation in data collection methods essentially precludes the determination of a standard quit rate across all quitlines. These and related issues are addressed in the remainder of this paper.

Who is the Audience

There are many potential audiences for this NAQC Issue Paper. A simplified schematic of the quitline community is shown in Figure 1.

Figure 1: The Quitline Community



One part of the quitline community are quitline service providers. This group represents a mix of large and small providers of quitline services. Another part of the quitline community are the stakeholders. While there are a range of stakeholders, in the simplified model above stakeholders are primarily viewed as the individuals or groups who make decisions about quitline funding. Examples of stakeholders include federal agencies, legislators, health department directors, health plan directors or employee or worksite health promotion supervisors. Positioned between the quitline service providers and the stakeholders are a group of individuals referred to as quitline contract managers. Quitline contract managers interact with quitline service providers (seeking bids, selecting vendors, overseeing contracts, etc.) and often report to stakeholders. All of these parties may also at various times have contact with the representatives of the media and thereby the public at large including both smokers and non-smokers.

Each of these parties also has an interest in quitline quit rates. Quitline service providers have an interest in examining the quit rates of their own program as a measure of quality and to identify opportunities for improvement and will also report these rates when they are seeking new contracts for their services. Stakeholders often are interested in quit rates of the services they fund to ensure their programs are having an impact and to support or to justify continued or expanded expenditures for quitline services. Quitline contract managers seek information on quit rates (from quitline providers) and report this information to stakeholders. The public has an interest in quit rates as well. For smokers, this information may influence decisions on whether or not to call the quitline. For non-smokers, this information could theoretically influence the level of encouragement they give to smoking acquaintances to call quitlines. Information on quit rates also has the potential to influence general support for continuation of public funding for quitline services.

This NAQC Issue Paper is intended primarily for quitline contract managers. At present, quitline contract managers are faced with an array of different quit rates calculated in myriad ways. This presents challenges when choosing a quitline service provider or assessing the quality of quitline performance. Many quitline contract managers are also often forced to make difficult choices as to which quit rate to use when communicating with different parties. Arriving at a standard definition for measurement of quitline quit rates is intended to simplify quitline contract managers' communications with quitline service providers and quitline stakeholders.

Several recent papers have provided recommendations on assessment of tobacco cessation outcomes as part of clinical trials (West, Hajek, Stead, & Stapleton, 2005; Hughes et al., 2003). Issues addressed include standards for duration of abstinence that need to be considered a treatment success, the role of rigorous biochemical verification and the implementation of abstinence calculations with sophisticated "blinded" follow-up protocols. While these recommendations are entirely appropriate for the conduct of clinical research, many of these methods are associated with increased respondent burden, can be expensive to implement and require the participation of large and well-developed research teams. Here, the goal is to develop an easy-to-use evidence-

based method that does not have the same degree of respondent burden and implementation cost.

II. Denominator Issues & Recommendation

A denominator used in the quit rate calculation provides a count of those served by the quitline. This section describes different definitions of the denominator and their corresponding advantages and disadvantages[1]. Table 1 below describes the various kinds of subgroups calling a quitline, starting with the most inclusive group then excluding groups one by one. These groups include:

1. All callers to the quitline. In addition to current and former tobacco users, includes proxies, wrong numbers and pranks.
2. All tobacco users.*
3. All tobacco users* *seeking treatment.*
4. All tobacco users* seeking treatment *who register for services.*
5. All tobacco users* seeking treatment who register for services and *consent to the evaluation.*
6. All tobacco users* seeking treatment who register for services and consent to the evaluation and *receive at least minimal evidence-based treatment.*

Figure 2 schematically illustrates potential denominator groups. The shaded portion of the figure defines the recommended denominator group. The non-shaded sections define groups recommended for exclusion from the denominator.

Table 1. Potential Denominator Groups and Their Pros and Cons

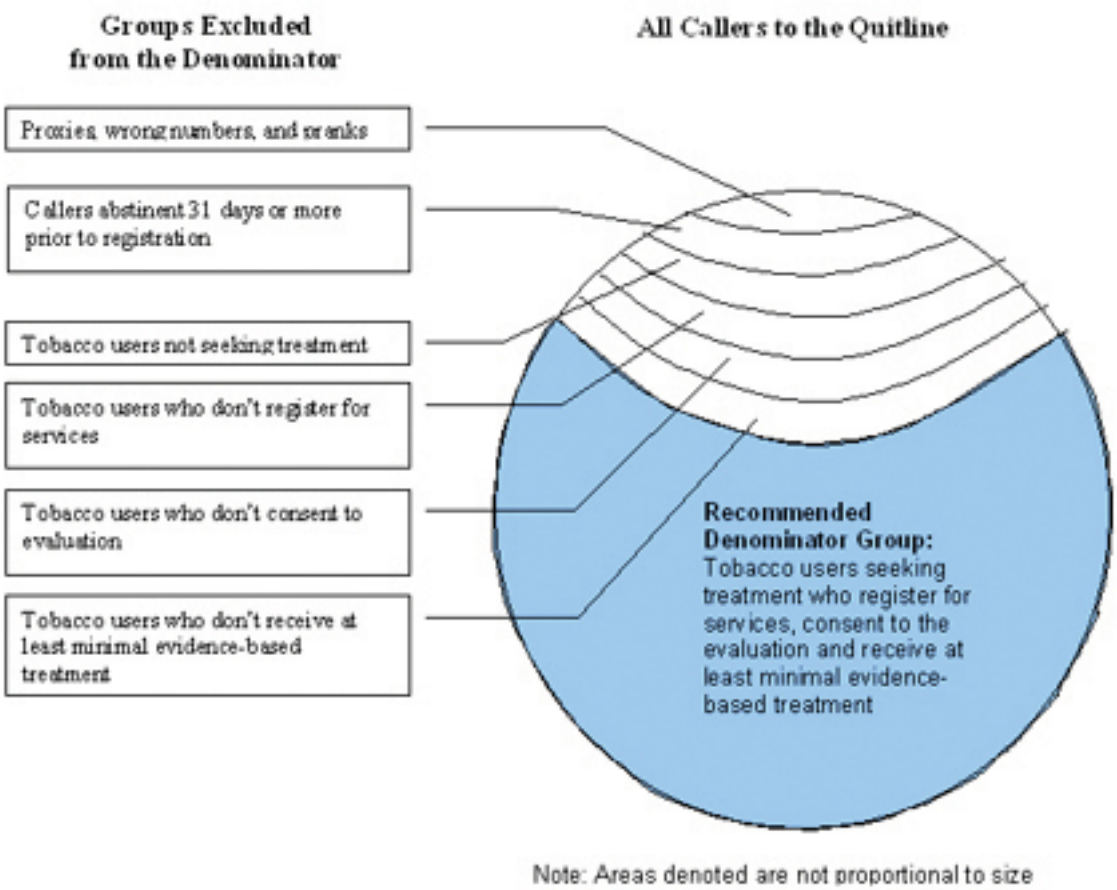
Group	Those Excluded	Pros of the Denominator	Cons of the Denominator
All callers to the quitline	None.	Simple to count and use.	Very heterogeneous group including those not seeking treatment, who, if included would artificially lower quit rate.
All tobacco users	All non-tobacco users including those calling to assist friends or family; medical providers; the public seeking information.	The rate reflects only those who use tobacco, including those who have recently quit within 30 days of intake or registration, the central audience of quitline and cessation follow-up.	Very heterogeneous group including those not seeking treatment, who, if included would artificially lower quit rate. If a goal of the quitline is to educate the general public about quitting tobacco, this rate excludes non-tobacco users.
All tobacco users seeking treatment	Non-tobacco users. Tobacco users who 1) are not interested in quitting and 2) were already quit at intake for more than 30 days.	This group reflects a target audience for most quitlines: tobacco users ready to quit. Research suggests including those quit for more than 30 days would bias quit rates upwards, so this group is excluded.	If a program specifically seeks to serve those not ready to quit or those already quit, these groups should be included in the denominator.
All tobacco users seeking treatment who register for services	Non-tobacco users. Tobacco users who 1) are not interested in quitting, 2) were already quit for more than 30 days, and 3) failed to register (complete intake).	This group reflects the target audience for most quitlines. Excluding those without intake is important because intake is necessary to provide treatment, so individuals without intake may not receive intervention. This would artificially lower abstinence rates.	Some callers may not complete intake due to technological or other problems. A quitline may have intended to serve those without intake, even if it may be unable to do so.

Cont. on page 6.

**Including those who have recently quit within 30 days of intake or registration.*

All tobacco users seeking treatment who register for services and consent to the evaluation	Non-tobacco users. Tobacco users who 1) are not interested in quitting, 2) were already quit for more than 30 days, 3) failed to register (complete intake), 4) declined to participate in the evaluation.	This group reflects those who the quitline intends to target and who have agreed to follow-up. This group represents the "intention-to-treat" group and closely approximates the group that a clinical trial model would define as the "treatment" group. Those agreeing to follow-up will be more amenable to follow-up, likely resulting in a higher ITT rate. Consent is required in clinical trials (see section 4 on response rates).	Non-consenters include participants quitlines would like to follow-up; however, they have declined to participate. This is not a clinical trial but a real-world evaluation group; there has been no randomization to be selected for this treatment, callers have self-selected themselves for this intervention. This group may or may not have actually gotten the treatment.
All tobacco users seeking treatment who register for services and consent to the evaluation and receive the minimal evidence-based treatment	Non-tobacco users. Tobacco users who 1) are not interested in quitting, 2) were already quit for more than 30 days, 3) failed to register (complete intake), 4) declined to participate in the evaluation, and 5) did not receive minimal treatment.	This group got some of the intended treatment, would more accurately allow assessment of the treatment itself.	Given a proven, strong dose-response relationship, this group will overestimate the quit rate. It also fails to meet criteria for the "Intention To Treat" model because the quitline intended to serve more callers than are included in this group.

Figure 2. Potential Denominator Groups: Excluded Subsets in Non-Shaded Areas and Recommended Denominator Group in Shaded Area



Recommendations

- **Include in the denominator all tobacco users who register for services, consent to follow-up*, receive some evidence-based treatment, and have not been quit at intake or registration for more than 30 consecutive days.**
- **Individuals who complete registration and intake but do not receive any counseling should not be included in quit rate calculations.**
- **A reasonable minimum for having received some treatment is the receipt of at least one telephone counseling session.**
- *The definition of treatment is expected to evolve as quitlines become more involved in provision of a range of tobacco cessation services (e.g. pharmacological therapy, Web-assisted tobacco interventions, etc.).*
- *Failure to deliver counseling to individuals who register for services is a quitline quality issue that should be addressed in companion papers to this report.*

III. Numerator Issues & Recommendation

The numerator in abstinence rates defines abstinence, or success in quitting tobacco. Collecting data for the numerator requires surveying specific quitline callers at specific times, with specific questions. Therefore, several key issues must be resolved about how the numerator in quit rate calculations is measured. First, a time point must be selected to conduct follow-up evaluation. Second, duration of abstinence at follow-up must be specified as meeting the definition of "successful" treatment. Third, question wording must be specified. Fourth, validation of self-reported abstinence by biochemical measures should be considered. This section explores each of these considerations.

Timing & Follow-Up

Specifying the timing of follow-up is an integral part of defining the numerator and requires two decisions. The first is the reference point, or the time from which follow-up begins; second is the length of follow-up, or the time from the reference point to follow-up. Each of these is discussed below.

Reference Point

The options in selecting a reference point for follow-up include beginning follow-up at enrollment, after a quit date is set or after treatment ends. The advantage of beginning follow-up at enrollment is that the date of enrollment is usually easily available for all callers. The disadvantage is that by beginning follow-up at enrollment, the period during which abstinence is calculated includes the time when the caller was supported by the quitline and may not yet have made a quit attempt. The second option for a reference point is after a quit date is set. While this approach has intuitive appeal, reliably tracking the true quit date for each caller is difficult in practice, and many callers do not set a quit date at all. Therefore, the abstinence rate resulting from this reference would not be valid.

Finally, a reference point may be defined after treatment ends. The advantage of this option is that it measures the amount of time a caller is abstinent without quitline support. The disadvantage is that determining when treatment ends and calculating a call date can be difficult because it requires specific information about the date of a callers' last call. Also, this option does not count as abstinence the time from when a quit date is achieved during treatment to after the treatment ends.

**With regards to obtaining consent for follow-up evaluation, quitlines (or their evaluation team) should strongly consider consultation with Human Subjects Research or Institutional Review Board depending on plans to publish or disseminate evaluation findings. Consent rates should also be reported with the aim to achieve at least an 85% rate of consent.*

Length of Follow-Up

A Society for Research on Nicotine and Tobacco (SRNT) workgroup charged with setting a uniform abstinence measure was unable to come to consensus in recommending either a six- or 12-month follow-up. In its MDS, NAQC recommends follow-up at seven months post enrollment, assuming a one-month treatment and follow-up six months post-treatment. The disadvantage of this approach is that brief, one-call programs are followed up later than is desirable. Despite this consideration, NAQC's recommended seven-month follow-up corresponds fairly well to the six month time point recommended by the SRNT workgroup while also allowing an initial one month grace period to initiate both treatment and a quit attempt. The preference for a longer length of follow-up must be weighted against attrition concerns. Attrition is manageable but often the solutions are resource intensive and expensive.

Recommendation

- **Conduct follow up seven months following quitline enrollment.** (This recommendation builds upon the existing MDS system and has the advantage that a majority of quitlines already conduct follow-up at this time point (NAQC Minimal Data Set Workgroup, 2008). The desire for longer-term follow-up is outweighed by attrition concerns or increased costs needed to mitigate attrition.)

Duration of Abstinence

The definition of abstinence and duration of abstinence needing to be counted as a treatment "success" have received considerable attention in the tobacco research community (Velicer, Prochaska, Rossi, & Snow, 1992; Hughes et al., 2003, West et al., 2005). Initial definitions centered around the concept of "continuous" abstinence as measured from the beginning of the intervention (or quit date) through to the final follow-up evaluation. More recently the concept of "prolonged" abstinence has been introduced. Prolonged abstinence is similar to continuous abstinence with the exception that a grace period is allowed (typically two weeks) to establish initial abstinence. In contrast to continuous and prolonged abstinence, point prevalence abstinence measures describe the proportion of callers who are abstinent for a shorter period of time immediately prior and through the follow-up evaluation. The length of time required to be abstinent for each measure varies, but the most common periods are 24 hours, seven days and 30 days.

A workgroup formed by SRNT reviewed the literature on outcome measures and attempted to generate consensus around uniform measures of abstinence within clinical trials (Hughes et al., 2003). Both the workgroup and a 2005 proposal by West and colleagues supported the use of prolonged abstinence, based on its greater stability. While there are merits of more prolonged abstinence measures as part of clinical trials, several points suggest consideration should be given to shorter-term point prevalence measures as part of ongoing program evaluation.

The first point of consideration in favor of utilizing a shorter-term abstinence measure is that there is actually a high degree of correlation between point prevalence and prolonged abstinence measures. For example, Velicer and Prochaska (2004) found a high correlation between both seven-day point prevalence and prolonged abstinence (r=0.85) and 30-day point prevalence and six-month prolonged abstinence (r=0.85).

A second point in favor of the use of a point prevalence abstinence measure for quitline evaluations is that this is consistent with the selection of outcome measures used to conduct the meta-analyses for the Public Health Service Clinical Practice Guideline, *Treating Tobacco Use and Dependence: 2008 Update*. The Guideline authors indicated point prevalence was preferred to continuous abstinence for several reasons. These included the frequency of studies reporting point prevalence abstinence (vs. more prolonged measures), the potential of continuous abstinence measures to underestimate the percentage of individuals who eventually obtain abstinence (i.e. recycle) and fact that most relapse begins soon after the initial quit attempt and these individuals

are most likely to report continued tobacco use at later follow-up. Several studies suggest the smoking cessation relapse curve is steepest in the initial few weeks following a quit attempt. Employing a 30-day point prevalence measure assures that individuals who are still in this very steep section of the relapse curve are not counted as "successes" in the quit rate calculation.

Recommendation

- **Measure and report 30-day point prevalence abstinence.**

Intake & Follow-Up Questions

In order for abstinence rate calculations to be comparable between quitlines, the specific wording for survey items used to gather data for abstinence rates must be identical. In particular, on both the intake and follow-up surveys, items must have the same question phrasing for the stem and item (the question) and the response options available to the respondent. One of the most successful efforts to standardize question wording is the MDS (described more fully in Section 3 above). While many quitlines have implemented at least a portion of the MDS intake and follow-up items since 2005, a report of the NAQC MDS Workgroup (NAQC Minimal Data Set Workgroup, 2008) found variability in administration of the questions (phrasing, timing) and follow-up eligibility criteria. Before results of follow-up surveys can be aggregated across quitlines and before comparisons can be made between quitlines, intake and follow-up questions need to be identical and used more consistently. Additionally, the definition of eligibility for follow-up must be applied identically across quitlines (see Section 2 for a fuller discussion of follow-up populations).

This section specifically reviews MDS items critical to calculating quit rates and makes recommendations for precise question wording for the one outcome measure proposed in this NAQC Issue Paper—30-day point prevalence. The items in Table 2 are necessary for determining eligibility for inclusion in 30-day point prevalence calculations and for calculating the quit rate. Table 2 below also provides comments on the purposes, uses and special concerns of those items.

Table 2. MDS Items Necessary for Determining Eligibility for Follow-Up and Calculation of Quit Rates

MDS Questions	Comments on the Purpose, Use and Special Concerns of Items
Intake:	
5a. USA: Do you currently smoke cigarettes every day, some days, or not at all? *(CHECK ONE)* o Everyday o Some days (if less than 7 days per week or less than 1 cigarette per day) o Not at all: o When was the last time you smoked a cigarette, even a puff (dd/mm/yyyy)? *(if exact day is unknown, use the 15th of the month)* **5b. Canada: Do you currently smoke cigarettes daily, occasionally, or not at all?** *(CHECK ONE)* o Daily o Occasionally (if less than 7 days per week or less than 1 cigarette per day) o Not at all: o When was the last time you smoked a cigarette, even a puff (dd/mm/yyyy)? *(if exact day is unknown, use the 15th of the month)*	o The purpose of this question is to assess who, at intake, are "everyday" or "someday" or "occasional" smokers. Note that in the US the term "some days" is used and in Canada the term "occasional" is used to refer to the same population of smokers (those that smoke less than 7 days per week or less than 30 days per month, or less than 1 cigarette per day). o An area of controversy in calculating quit rates is whether or not to include those who have already quit in the denominator. We address this in section 4. o This MDS item can be used to determine if a caller has already quit at intake, and would indicate their eligibility for follow-up or inclusion in quit rate calculations. o Item of note: The prompt after the date field: "if exact data is unknown, use the 15th of the month", was added as a solution to handling respondents who often have problems recalling the exact date they last smoked a cigarette.

Cont. on page 10

7. Do you currently use other tobacco products such as…*(check all that apply)* o Cigars o Pipes o Chewing tobacco or snuff o Other tobacco products (e.g. bidis)	o The purpose of this question is to assess whether callers are using other tobacco at intake. This question, taken together with Q5a, assesses all tobacco use. o Those who are using cigarettes **and / or** other types of tobacco at intake should be considered non-abstinent at intake. Callers that report using **any** type of tobacco, therefore, would be eligible for follow-up and inclusion in quit rate calculations.
Follow-up:	
10. Have you smoked any cigarettes or used other tobacco, <u>even a puff</u> , in the last 30 days? o Yes o No o Don't know o Refused	This item measures whether quitline users have abstained from using **any** type of tobacco in the 30 days prior to being administered the follow-up survey. This measure of duration of abstinence is called a 30-day point prevalence estimate. Those who report having used any form of tobacco in the last 30 days at follow-up are considered non-abstinent. Note that the threshold for tobacco use is very stringent (no use, not even one puff).

Recommendation

- **Use Intake items 5a or b and 7 and follow-up item 10 in Table 2 above to measure 30-day point prevalence, the proposed outcome measure in this paper.**

Biochemical Validation

For most telephone counseling interventions, tobacco use status at follow-up is derived from participants' own self-reported quit rates. One concern with using self-reported quit rates is that participants may falsely report themselves as having quit, thereby inflating the quit rate. Biochemical validation can be used to assess smoking behavior when it is believed that using self-reported quit rates will not yield accurate results. Biochemical validation can reduce inflation of quit rates caused by dishonest self-report, however, it is costly (especially when used with large populations), more invasive than telephone, Internet and mail follow-up surveys, and according to the literature, not always necessary.

According to a SRNT subcommittee charged with assessing the utility of biomarkers and making recommendations regarding their applications, it is suggested "in large-population, low-intensity trials, biochemical verification is neither feasible nor necessary" (Benowitz et al., 2002). According to their report, while it is likely that quit rates are inflated when using self-reports, the amount of inflation is small. This is supported by many reviews of the literature including Velicer et al. (1992), which found self-reported quit rates for intervention studies were only slightly more inflated than those of untreated volunteer samples. Patrick et al. (1994) also found generally high levels of specificity and sensitivity for self-report.

However, it is important to note there are some special populations for which misreporting may be higher (Benowitz et al., 2002). These populations include adolescents who may misreport because accurate reporting would in some cases be admitting to illegal activity or because of elevated pressure to quit. In addition, pregnant smokers and medical patients may feel an elevated need to misreport since these special populations are likely to feel an increased pressure/expectancy to quit (Benowitz, et al., 2002).

Recommendations

- **Do not conduct biochemical validation. It is not recommended and the literature shows that self-reported smoking behavior is an adequate means to measure quit rates for tobacco cessation programs.**
- *Remain aware that self-reported quit rates are likely to include some small amount of inflation, and that this inflation is likely to be higher if the special populations you measured are facing an elevated expectancy to quit or feel a greater need to hide smoking behaviors.*

IV. Shared Issues in Defining the Numerator & Denominator

A handful of issues impact the definition and calculation of the numerator and denominator. How to handle missing data, specialized interventions within a quitline and participants who are already quit are each discussed below. How to handle missing data is a particularly important consideration and has a substantial impact on the calculation of quit rates.

Reducing Missing Data

The first and most important strategy for dealing with the problem of missing data is to attempt to minimize the amount of missing data. For the purposes of quit rate evaluation, this translates directly into increasing the response rates on follow-up evaluation surveys. However, telephone surveyors face new barriers to reaching high response rates. Response rates are declining nationally across many populations and for many survey topic areas. For example, since 2002, 70% of states participating in the Behavioral Risk Factor Surveillance System (BRFSS) have experienced a decline in response rate of a median 2.2 percentage points (Link & Mokdad, 2006).

The changing landscape of telephone ownership is one factor influencing surveyors' ability to reach participants. Many U.S. households are switching from landline telephones to cell phones and the pace of this change is accelerating. Data from the National Health Interview Survey show the percentage of adults living in wireless-only households has been steadily growing. In 2003, only one out of every 28 adults lived in wireless-only households. In 2005, one out of every 13 adults lived in wireless-only households and by 2006 it was one out of every eight (Bloomberg and Luke, 2007). This trend is likely to affect the ability to successfully contact quitline callers for follow-up because quitline participants' move from landline to wireless-only households degrades a survey sample. Even when callers' telephone numbers remain active, the increased practice of screening calls using caller ID and voicemail reduces contact with potential survey respondents (Tuckel & O'Neill, 2001). Moreover, achieving a high response rate is critical to the precision and accuracy of quit rates (see the next section, *Missing Outcome Data*, for a fuller discussion of this issue).

Taken together, it is more important than ever that quitlines carefully consider telephone survey follow-up protocols in order to achieve the highest possible response rate with the resources available for evaluation. The recommended target goal for follow-up response rate is 50% understanding of course that this may not be attainable based on the population served, available budget or other factors. Conversely, some quitlines (e.g. those conducting research studies) may obtain response rates much higher than 50%. In any case it is critical quitlines report the response rate along with the quit rate to help with interpretation of the quit rate.

Fortunately, much literature exists to guide surveyors in achieving a high response rate (Lepkowski et al., 2008, Dillman, 1991). Dr. Don Dillman has been conducting scientific research over the past 25 years on increasing response rates and is considered to be one of the most influential figures in developing a scientific basis for survey research methodology. A sociologist, Dillman's work crosses a wide variety of topic areas. ➢
His recent books include *Survey Nonresponse* (2002, eds., Groves, Dillman, Little and Eltinge), and *Mail and Internet Surveys: The Tailored Design Method,* 2nd ed. (2007 Update). The field of public opinion research has

also contributed a large body of literature to optimize random digit dial telephone surveys (Groves, 1988). It is important to note, however, the literature on random digit dial telephone surveys has somewhat limited applicability to the listed samples used in quitline follow-up evaluations.

The following section provides highlights from the literature on survey methodology in order to highlight areas where recommendations for increasing survey response rates may be drawn. In other areas where the literature and experience provide somewhat weaker evidence, suggestions for quitlines to consider are provided.

The literature supports the use of pre-notification or advance letters, sent by mail, as a cost-effective way to increase response rates. Sending advance letters helps to reduce suspicion about a telephone call from an unknown source, as well as legitimize the survey call and differentiate the survey from a sales call (Dillman, Clark & Sinclair, 1995). Additionally, letters can also be used to evoke the principles of social exchange and reciprocation, thus increasing response (Dillman, 2007). The letters should be as brief as possible and printed on the quitline and/or sponsoring agency letterhead and envelope to capitalize on participants' affiliation with the service.

Survey methodology research also demonstrates that advance letters can produce a sizeable increase in response. A meta-analysis published in 2007 reviewed 29 independent studies on the influence of advance letters on response to telephone surveys. Findings indicate that sending an advance letter improves the cooperation rate (agreeing to take the survey) in random digit dial telephone surveys by about 11 percent on average and answering the telephone by about eight percent. The effect is even larger when using a listed sample of identified persons (De Leeuw, Callegaro, Hox, Korendijk, & Lensvelt-Mulders, 2007), which is the case in quitline follow-up surveys.

In addition to boosting response advance, mailings are also cost effective. A recent study tested the use of advance mailings as part of a 2002 survey on nutrition and physical activity. The targeted sample size was 3,500 interviews with randomly selected adults, following protocols established by the CDC for the BRFSS. The study found that sending advance mailings in a random-digit-dial survey reduced the number of attempts needed to close out each case, resulting in lower interviewer and telephone costs. This produced a net savings despite the increased costs associated with the advance mailings* (Hembroff, Rusz, Rafferty, McGee, & Ehrlich, 2005).

Literature also supports the use of monetary incentives to improve response rates in public opinion surveys (Singer, Groves, & Corning, 1999; Singer, 2002; Cantor, O'Hare, & O'Connor, 2007). Incentives can be provided in advance by enclosing a token amount ($2 to $5 cash or gift card) inside the advance letter. The token advance incentive is considered part of the social exchange; it is a thank-you or reward for the time and effort expended to complete the survey. Incentives can also be promised, by mentioning at time of consent and again in the advanced letter that a thank-you gift will be provided upon completion of the survey (such as $10 or $20 or an opportunity to be entered into a drawing for a larger prize). This type of incentive is an example of an economic exchange model or paying participants to complete the survey. Advance incentives have been found to be more effective than promised incentives in raising response rates. An experiment comparing several incentive protocols in a survey of consumer attitudes (Singer, Hoewyk, & Maher, 2000) found that promised incentives did not increase response rates, while an incentive of $5 enclosed with the advance letter increased response rates by at least 10 percentage points. The advance incentive also reduced the number of call attempts needed to close out a case, thus contributing to the control of survey costs.

However, lack of an available mailing address, transiency among the quitline population and costs of providing incentives to every caller eligible for follow-up can make advance incentives problematic or even impossible for

**This study compared three advance mail strategies in a random-digit-dial survey and found that to obtain approximately 3,000 completed surveys, sending advance letters cost $11,018 less than no mailing at all.*

some quitlines. In these cases promised incentives may be a good alternative. Promised incentives are less costly since they are only paid to those who complete the survey and can be used even when the address is not available. Use of an incentive is recommended if response rates are low. The type and amount of incentive should depend on the unique qualities of a quitline's population, the contact data available for that population and the budget allocated for follow-up.

One substantial barrier to achieving high response rates is the number of quitline participants whose telephone number is not in service at follow-up. One reason for this is many U.S. households are changing from landline telephones to cell phones, and the pace of this change is increasing (Bloomberg and Luke, 2007). Quitlines can mitigate the damage of this trend by collecting as much data as is feasible during the first call. This includes asking for more than one telephone number (cell phone, work, landline) and a mailing address, which is necessary for sending an advance letter (see the above discussion of the effectiveness of advance letters). Weighing the collection of additional information at intake against the length of the intake process and available data fields, as much contact information should be collected as is feasible.

The introduction of the survey to evaluation respondents presents an additional opportunity to increase response rates. In an effort to increase cooperation rates in large-scale telephone surveys the National Organization for Research at the University of Chicago (NORC) developed and tested five introductory scripts for use during an ongoing evaluation of the Racial and Ethnic Approaches to Community Health (REACH) 2010 Program. This experiment found that the inclusion of an introductory statement identifying a government sponsor (which many quitlines have) increased participation (Dew et al., 2004). Dillman (2007) has postulated the following principles regarding the survey introduction: it should be brief, identify the survey's topic and length, as well as a reminder of any incentive or previous agreement to participate. Quitlines should carefully consider the follow-up survey introduction as an opportunity for boosting response.

Converting soft refusals is a critical element of obtaining the highest response rate possible. It is also common practice among surveyors. However, little literature exists to guide quitline evaluators in developing a soft refusal policy. Since survey response is correlated with tobacco use status, quitline participants may decline to take the survey, giving reasons such as "I didn't really use the quitline" or "I didn't quit smoking." Therefore, survey protocols may be developed to convert this type of soft refusal by assuring potential respondents that the quitline is interested in hearing about the experiences of all callers, regardless of how much or how little they used the quitline, and whether or not they are still using tobacco. Quitlines should carefully consider their own protocols in converting soft refusals.

Finally, the number of attempts surveyors make before closing out a contact can influence response rates. The CDC recommends 15 attempts for the BRFSS. However, the number of attempts needed to produce an adequate response rate varies by population. One evaluation of the Minnesota QUITPLAN Helpline found that 80% of the eventual respondent total was completed after six attempts, and 90% was completed after ten attempts. Continuing to make 15 attempts yielded another 10% of the total respondent group (based on 25 total attempts). Making more than 15 attempts was found to have diminishing returns (Rainey & Huggins, 2007). Therefore, quitlines should make as many attempts as possible given budget constraints and adjust that number as needed to achieve an adequate response rate.

Recommendations

- **Select a combination of strategies appropriate to the quitline's unique resources and needs in order to obtain a follow up response rate of 50%.**
- **Report the follow-up survey response rate along with the quit rate.**

- *When response rates fall below the goal of 50%, utilize advance letters and incentives which literature shows to be cost effective.*
- *Consider obtaining consent and additional contact information at intake and carefully attend to policies and protocols regarding survey introductions, conversion of soft refusals and the number of attempts in efforts to increase response rates.*
- *Quit rates should be interpreted with caution.*

Missing Outcome Data

Even after substantial efforts are made to increase follow-up survey response rates there will almost certainly be some missing data. How to handle missing outcome data is a key decision in calculating abstinence rates. There are two main approaches to calculating quit rates in the face of missing data. First, one may base quit rates only on those that respond to the survey, which is referred to as a responder rate (RR). Conversely, one may employ an Intention-To-Treat (ITT) approach where all non-respondents are considered to be smoking. The major advantages of both the RR and ITT rate is they are simple to implement and explain. The importance of these advantages should not be overlooked when this information needs to be shared with quitline stakeholders, the media and the public.

At the same time, it is important to acknowledge limitations of both the RR and ITT approaches in quit rate calculation. It is unavoidable that each of these simple approaches results in some degree of systematic error. It is well established in the conduct of clinical research that individuals who do not complete the treatment protocol or are otherwise "lost to follow-up" have worse outcomes than individuals who complete all aspects of the treatment and follow-up evaluation. The RR approach ignores this association and thus likely leads to a systematic overestimate of the true quit rate. The ITT approach addresses this issue by adopting an extreme position that all individuals who do not complete the follow-up evaluation are considered treatment failures. This likely leads to a systematic underestimate of the true quit rates.

In the conduct of clinical trails, the ITT approach has long been the dominant (though not exclusive) strategy for reporting quit rates. This is based upon the belief that the ITT approach is the most conservative when testing the hypothesis that a new treatment or program is efficacious compared to a different treatment. While this may be the appropriate approach for clinical trails, it is not at all clear this is the best approach for the evaluation of real-world operational programs such as quitlines.

In considering this issue further, it is important to recognize there are substantial downsides to the overestimation and underestimation of the true quitline quit rates. The major disadvantage of overestimation of the true quit rate is the creation of unrealistic expectations regarding program outcomes and the risk of the loss of credibility (of the evaluation and perhaps of the overall program) if there is a large gap between reported and true program outcomes. The disadvantages of underestimation of true quit rates are the creation of overly pessimistic attitudes about quitline services that could negatively influence participation rates (i.e. tobacco user's interest in calling the quitline) and decisions by stakeholders regarding the continuation of funding or expansion of quitline programs.

Recognizing the risk of over- and under-estimation of quit rates, this NAQC Issue Paper approaches the handling missing data by attempting to identify an approach that, in a simple manner, most closely approximates the true quit rate. In considering this issue further, it is apparent that the relative accuracy of the RR and ITT approaches to quit rate calculation is dependent upon the nature of the relationship between the follow-up survey response rate and the quit rate among respondents.

To illustrate these issues, the following three hypothetical cases are presented. Each of these cases deals with the follow-up evaluation of 1,000 imaginary quitline participants. For the purposes of this exercise, let us imagine we have "perfect knowledge" of the quitline outcomes for the entire 1,000 participants and the overall quit rate is 25% (250 out of the 1,000 participants have quit). What will vary for each case is the strength of the relationship between the survey response rate and quitting. For each hypothetical case, the follow-up surveys are conducted in waves representing, for example, a call attempt.

Table 3. Three Hypothetical Cases of Quitline Evaluations

Follow-up Wave	Number Reached	Number Smoking	Number Quit	Quit Rate within Wave	Cumulative Response Rate	Cumulative RR Quit Rate	Cumulative ITT Quit Rate
CASE 1							
Wave 1	200	142	58	29.0%	20.0%	29.0%	5.8%
Wave 2	200	146	54	27.0%	40.0%	28.0%	11.2%
Wave 3	200	150	50	25.0%	60.0%	27.0%	16.2%
Wave 4	200	154	46	23.0%	80.0%	26.0%	20.8%
Wave 5	200	158	42	21.0%	100.0%	25.0%	25.0%
CASE 2							
Wave 1	200	126	74	37.0%	20.0%	37.0%	7.4%
Wave 2	200	138	62	31.0%	40.0%	34.0%	13.6%
Wave 3	200	150	50	25.0%	60.0%	31.0%	18.6%
Wave 4	200	162	38	19.0%	80.0%	28.0%	22.4%
Wave 5	200	174	26	13.0%	100.0%	25.0%	25.0%
CASE 3							
Wave 1	200	110	90	45.0%	20.0%	45.0%	9.0%
Wave 2	200	130	70	35.0%	40.0%	40.0%	16.0%
Wave 3	200	150	50	25.0%	60.0%	35.0%	21.0%
Wave 4	200	170	30	15.0%	80.0%	30.0%	24.0%
Wave 5	200	190	10	5.0%	100.0%	25.0%	25.0%

In Case 1 we imagine with the initial and each follow-up wave that we are able to complete follow-up evaluation with an additional 200 participants. For example, after our first follow-up call attempt, we are able to reach 200 of our 1,000 participants (cumulative response rate 20%). The second attempt reaches an additional 200 of our 1,000 participants (cumulative response rate 40%). After the third and fourth calls, we are able to reach an additional 200 of our 1,000 participants (cumulative response rate 60%) and so on. It is easier to reach the first 200 participants than the last 200 participants, and thus, we expect there to be a progressive decline in the observed quit rate within these sequential groups of 200.

The specifics for Case 1 are shown in Table 3 and include the number reached, the projected number still smoking, the projected number quit, the quit rate within each group of 200, the cumulative response rate and the cumulative quit rates calculated using the RR and ITT approaches. In Case 1 we hypothesize a relatively flat relationship between the response group and the quit rate within each group. In Case 1, the quit rate within each group decreases by only two percentage points with each subsequent follow-up wave. So for the first 200 individuals reached as part of the follow-up evaluation, 29% will report being abstinent. For the next 200 individuals reached, 27% will report being abstinent. For the next 200, 25% will report being abstinent, etc.

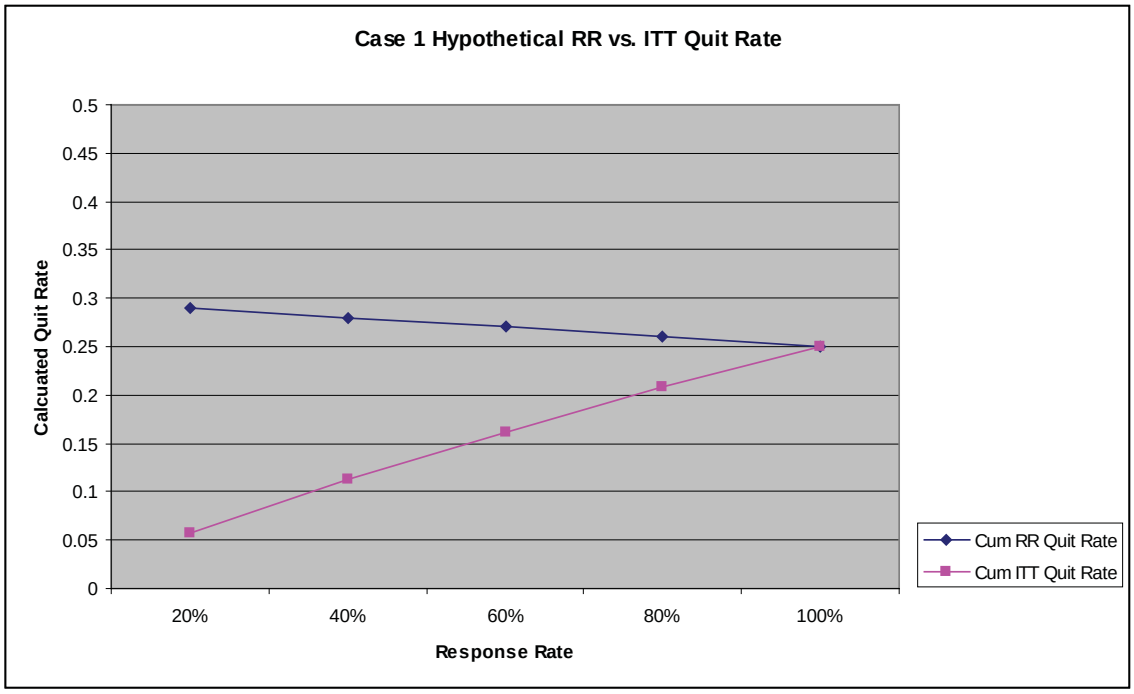
For the purposes of illustration, we then calculate a cumulative response rate after adding the follow-up surveys of each wave of 200 participants as they are completed. After the first wave of follow-up we have 200 completed follow-up surveys, a cumulative response rate of 20% (200/1000), and cumulative RR quit rate of 29% (58/200) and a cumulative ITT quit rate of 5.8% (58/1000, all non-respondents considered to be smokers). After the second follow-up wave we have 400 completed follow-up surveys, a cumulative response rate of 40% (400/1000), a cumulative RR quit rate of 28% (58+54/400) and a cumulative ITT quit rate of 11.2% (58+54/1000). With each successive follow-up wave the cumulative response rate increases and eventually

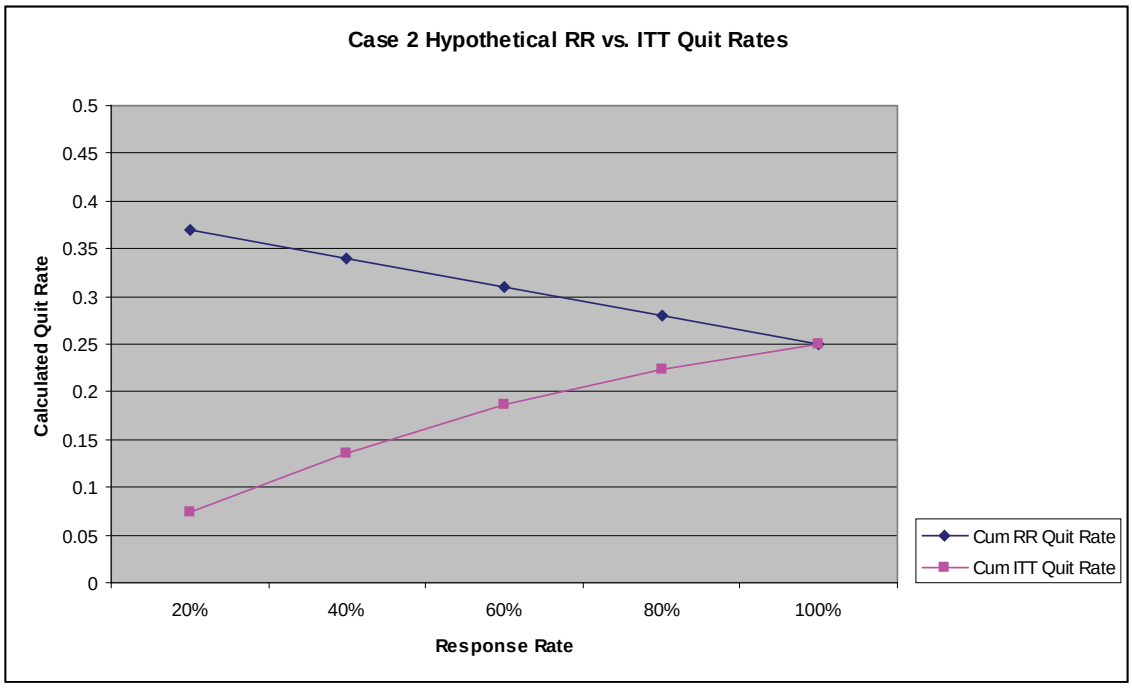
reaches 100%. When the response rate reaches 100% the RR and ITT quit rates are identical at 25%, which represents the "true" quit rate in this hypothetical group of 1,000 quitline participants.

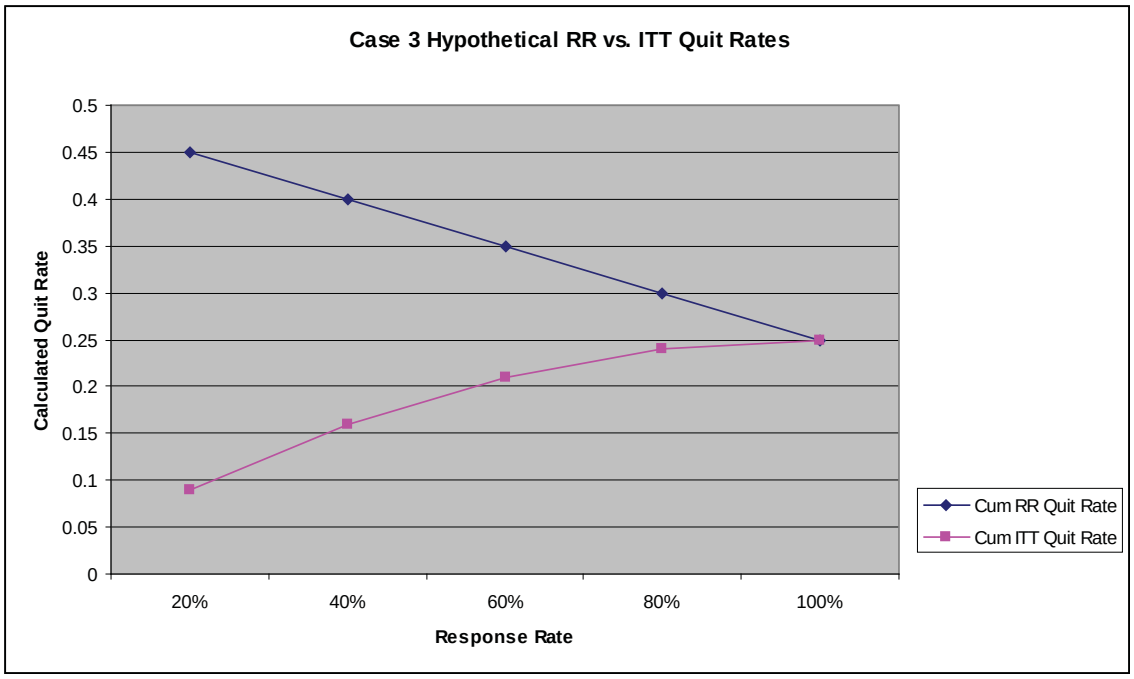
The specifics for Case 2 are calculated the same way except the quit rate decreases more sharply (6% points vs. 2% in Case 1) within each successive follow-up wave. Case 3 is handled the same way expcept the quit rate decrease even more sharply within each follow-up wave (10% for each wave).

To assist in visualization of these cases, Figure 3 plots the relationship between the cumulative response rate and the cumulative RR and ITT quit rates.

Figure 3: Three Hypothetical Cases of the Relationship Between the Cumulative Response Rate and the Cumulative RR and ITT Quit Rates







What these figures show is that in Case 1 (2% decrease in quit rate per 20% increase in response rate), the RR quit rate is closer to the true quit rate of 25% than the ITT quit rate across nearly the entire response rate range. In Case 2 (6% decrease in quit rate per 20% increase in response rate), the RR and ITT quit rates are nearly equal distant from the true quit rate in the response rate range from 40% to 80%. In Case 3 (10% decrease in quit rate per 20% increase in response rate), the ITT quit rate is closer to the true quit rate than the RR quit rate in the response rate range from 40% to 80%.

Based upon these examples, one would conclude in Case 1 the use of a RR quit rate is superior to use of the ITT quit rate for quitline outcome evaluation (i.e. closer to the true value and changing less over a response rate range from 40% to 80%). In Case 2, the RR and ITT quit rates are similar in their accuracy and stability. In Case 3, the ITT quit rate is superior to the RR quit rate.

Let us examine how real data fits into this schema. Data presented here are from separate evaluations of quitline services in Minnesota and California. The Minnesota findings are from evaluation of 670 of Minnesota's QUITPLAN® Helpline participants conducted by Professional Data Analysts, Inc. This data is presented by each wave of follow-up corresponding to increasing efforts to complete the follow-up evaluation. The final response rate at the end of Wave 6 (11+ attempts to reach) was 62.2% (416/667). The number of individuals reached at the initial and subsequent attempts, abstinence rates among these individuals and the cumulative quit rates RR and ITT quit rates are presented in Table 4. Similar to the hypothetical cases, the quit rate decreases as we progress from individuals reached with the initial to subsequent attempts.

For example, in this evaluation the first call attempt reached 94 individuals, of whom 59 reported smoking and 35 reported they had quit (quit rate 37%). The second attempt reached 70 individual of whom 47 reported smoking and 23 reported being quit (quit rate 33%). This general trend (decreasing response rate with increasing attempts to contact) so that among those reached after 11 or more attempts (i.e. Wave 6, n=71) 51 reported smoking while only 20 reported quitting (quit rate within wave of 28%). On average, these data show a decrease in the quit rate of 5.1% for every 20% increase in response rate.

Table 4. Minnesota QUITPLAN Helpline Evaluation of 670 Clients

Follow-up Wave	Number Reached	Number Smoking	Number Quit	Quit Rate within Wave	Cumulative Response Rate	Cumulative RR Quit Rate	Cumulative ITT Quit Rate
1 (1st attempt)	94	59	35	37%	14%	37.2%	5.2%
2 (2nd attempt)	70	47	23	33%	25%	35.4%	8.7%
3 (3rd attempt)	50	36	14	28%	32%	33.6%	10.7%
4 (4-5th attempt)	62	44	18	29%	41%	32.6%	13.4%
5 (6-10th attempt)	69	56	13	19%	52%	29.9%	15.4%
6 (11+ attempts)	71	51	20	28%	62%	29.6%	18.4%

Data from an evaluation of the ClearWay Minnesota℠ QUITPLAN Helpline, April, 2007.

Data from the California Smokers' Helpline is presented in a similar fashion.* The California findings are from 4,629 smokers who called the California Smokers' Helpline and were randomly selected for an evaluation seven months after their intake date and were not part of another study. The total final response rate after completion of attempts to contact was 68.2% (3158/4629). Specifics from this evaluation in terms of the number of individuals reached at the initial and subsequent attempts, abstinence rates among these individuals and the cumulative quit rates RR and ITT quit rates are presented in Table 5. On average, these data show a decrease in the quit rate of 2.9% for every 20% increase in the response rate.

Table 5. California Smokers' Helping Evaluation of 4629 Clients

Follow-up Wave	Number Reached	Number Smoking	Number Quit	Quit Rate within Wave	Cumulative Response Rate	Cumulative RR Quit Rate	Cumulative ITT Quit Rate
1 (1st attempt)	602	432	170	28.2	13.0	28.2	3.7
2 (2nd attempt)	470	351	119	25.3	23.2	27.0	6.2
3 (3rd attempt)	304	217	87	28.6	29.7	27.3	8.1
4 (4-5th attempt)	428	326	102	23.8	39.0	26.5	10.3
5 (6-10th attempt)	526	397	129	24.5	50.3	26.1	13.1
6 (11+ attempts)	272	221	51	18.8	56.2	25.3	14.2
7 (17+ attempts)	556	437	119	21.4	68.2	24.6	16.8

Data from evaluation of callers from 1/1/2004-12/31/2007.

Both these real world cases (Minnesota with 5.1% decrease in quit rate for 20% increase in response rate and California with 2.9% decrease in quit rate for 20% increase in response rate) show a relationship between response rate and calculated quit rates that is somewhere between the hypothetical Case 1 (2% decrase in quit rate per 20% increase in response) in which the RR quit rate was superior to ITT quit rate and Case 2 (6% decrease in quit rate per 20% increase in response rate) in which RR and ITT quit rates were equally accurate. We conclude from this that across a reasonable range of response rates that might be achieved in quitline evaluations, the RR quit rate is likely to be at least as accurate or perhaps more accurate than the ITT quit rate.

There are obviously several limitations to this analysis. The main limitation is that we do not know from the data the "true" quit rate or how increasing the response rate past the actual rates observed would influence calculated quit rates. In the hypothetical models, a linear relationship between response rates and quit rates among additional respondents reached (i.e. decrease of X% per 20% increase in response rate) was assumed. The conclusions are obviously dependent upon these assumptions. While these limitations are certainly valid, the consistent findings from two states that achieved respectable overall response rates lends credence to the

**We would like to thank Dr. Sharon Cummins, Christopher Anderson, and Dr. Shu-Hong Zhu for sharing data and analyses regarding the California Smokers' Helpline.*

proposed recommendations.

There are certainly more sophisticated ways of handling missing data. These approaches involve imputation of missing outcome data based on the available data – such as intake data. Hall et al. (2001) discuss a number of different imputation models including "missing at random" (MAR) and "missing completely at random" (MCAR). Both models seem inappropriate for missing follow-up data in tobacco cessation modeling because the outcome data is clearly not missing at random.

A more sophisticated approach considers "non-ignorable non-response" (Hall, page 199), which means that the "missingness" is related to the outcome being measured. The ITT approach is an extreme case of this "non-ignorable non-response" in assuming a perfect relationship exists between using tobacco and having missing data.

Finally a "selection model," which is based on generating propensity scores to predict who is missing and then using these scores in a model to help predict quitting, (Hall et al., 2001) could be employed. This type of model may use intake data such as respondent age, stage of readiness to quit and level of addiction to predict outcomes. The advantage of a propensity score model is it may be more accurate than the other calculation strategies mentioned above. The disadvantage of the propensity score imputation is that it requires a sophisticated research team and many hours to develop these models, which have many additional assumptions that must be carefully delineated. Therefore, the use of propensity scores may not be practical in the evaluation of quitline quit rates in general practice.

Recommendations

- **Use the Responder Rate (RR) Quit Rate (number quit/number of follow-up survey respondents) as the primary measure for reporting quitline outcomes.**
- **Always report the response rate alongside the Responder Rate quit rate.**
- *Do not use more complex imputation-based methods for estimating quit rates (i.e. imputation) as part of standard quitline evaluation.*

Quitlines may offer a variety of programs with varying intensity to their callers based on eligibility criteria such as insurance status or income. For example, certain callers wanting to quit may receive written materials only, such as brochures or instructional booklets. A more intensive intervention is brief one-call telephone counseling, while multi-call telephone programs provide an even stronger intervention. In any of these programs, callers may be eligible to receive pharmacotherapy, which increases the intensity of the intervention further. Pharmacotherapy provision may also be an intervention by itself. Additionally, quitlines may also provide specialized programming to priority populations, such as pregnant women, youth, smokeless tobacco users and those with chronic disease. These interventions may also differ in their intensity, such as a 10-session program serving pregnant women versus a referral-only program for youth tobacco users.

Specialized Programs & Dose Response

Quitlines may offer a variety of programs with varying intensity to their callers based on eligibility criteria such as insurance status or income. For example, certain callers wanting to quit may receive written materials only, such as brochures or instructional booklets. A more intensive intervention is brief one-call telephone counseling, while multi-call telephone programs provide an even stronger intervention. In any of these programs, callers may be eligible to receive pharmacotherapy, which increases the intensity of the intervention further. Pharmacotherapy provision may also be an intervention by itself. Additionally, quitlines may also provide specialized programming to priority populations, such as pregnant women, youth, smokeless tobacco users and .

those with chronic disease. These interventions may also differ in their intensity, such as a 10-session program serving pregnant women versus a referral-only program for youth tobacco users.

Literature supports the efficacy of telephone counseling and suggests a strong dose-response relationship. For example, Rabius, Pike, Hunter, Wiatrek, and McAlister (2007) demonstrated that counseling was significantly more effective than materials only and the meta-analysis conducted by Stead, Perera and Lancaster (2008) found an association between intended program intensity and effect size, as did the Public Health Service Clinical Practice Guideline, *Treating Tobacco Use and Dependence: 2008 Update* (Fiore et al., 2008). A dose-response relationship may be seen in the type of services quitlines provide to their callers, each of which has a different intensity of intervention. Therefore each intervention is unique and may be expected to produce a different quit rate.

The differences in quitline programming have important implications for reporting quit rates. Calculating one abstinence rate that combines participants in an intensive, multi-call program with a brief, one-call intervention, for example, is simple. However, it masks the impact of both interventions. Likewise, if one abstinence rate is reported for a standard five-session program for adults and for a low-enrollment, intensive 10-session program for pregnant women, the impact of the program for pregnant women is likely to be overshadowed.

Recommendations

- **Report information about program intensity with quit rates (e.g. duration of programming, availability of pharmacological aids, special counseling strategies and content).**
- *Calculate numerators and denominators separately for each program. (For example, the callers participating in a multi-call program providing NRT could be grouped in separate numerators and denominators from those participating in a one-call brief counseling intervention.)*

Those Already Quit

A special group of participants are those already quit. The literature suggests those already quit at intake are less likely to be smoking at follow-up, so would bias quit rates upwards. However, Hughes et al. (2003) suggests those who have been in a quit attempt for up to several weeks are in a transition period of occasional tobacco use, so should be given a "grace period" when calculating prolonged abstinence. At present, no consensus has emerged as to how to treat those who have recently quit or quit for a long period of time.

A key issue in determining how to treat those quit at intake is defining who is and is not a smoker. This NAQC Issue Paper recommends abstinence is defined as those who have been quit for the 30 days prior to the follow-up period. We recommend this same logic be applied to intake. If participants have smoked in the 30 days prior to calling the quitline, they should be considered smokers and be eligible to be included in quit rates. If they have been abstinent for 31 days or more, they should be considered abstinent at intake and excluded from the quit rate.

A second important issue is how quitlines serve those quit at intake. Some quitlines provide a separate intervention for those quit at intake. In this case, a separate quit rate for this group is indicated. However, other quitlines treat those quit at intake with similar protocols as those still smoking at intake. No special distinction is made by the quitline. These quitlines view those quit at intake as asking for help and struggling in their efforts to quit. Their quit attempts may not be fully engaged and they are able to benefit from counseling in the same way as smokers who call the quitline. Moreover, those quit at intake receive the same basic protocol as smoking callers. For some quitlines, if separate quit rates were to be reported, the number of completed surveys for both groups (those quit before intake and those smoking at intake) would be too small to have confidence in the results. Therefore, reporting one quit rate with smokers and non-smokers at intake has practical advantages.

Finally, the literature has well established that smokers frequently lapse and relapse, even as they go on to achieve long-term abstinence (Hunt & Bespalec, 1974, Cohen et al, 1989). Hughes and colleagues (1992) followed up self-quitters at 2, 7, 14, 30, and 90 days and six months. They demonstrate that slipping was common and strongly predicted relapse, but that a minority of such slippers went on to become long-term abstainers at six months. Therefore, excluding those who recently quit from quit rates would underestimate the proportion of quitline callers who achieve long-term abstinence.

Based on these considerations, participants who are abstinent 30 days or fewer prior to calling the quitline should be included in calculated quit rates. One exception is for quitlines that offer unique programs for those already quit. In this situation, quit rates should be calculated separately for action/maintenance programs.

- **Include all callers who quit 30 or fewer days prior to calling the quitline in the numerator and denominator used in the abstinence rate calculation. Any caller who started their quit 31 or more days prior to calling should be excluded.***
- *For quitlines that have a program designed specifically to serve those already quit, callers who have already quit should be treated as a special target population and a separate numerator and denominator should be calculated for this group.*

**Please note that this recommendation is consistent with the recommendation for a 30-day point prevalence abstinence measure.*Recommendations*

V. Conduct of the Evaluation

Subject Selection for Evaluation

Consent

Determining procedures for gaining caller consent to participate in an evaluation is an integral part of conducting quitline studies. The nature of the study – whether it is primarily evaluation or primarily research – impacts the consent process used. Evaluation often allows for a simpler participant consent process than research studies, and quitline evaluations will often be determined to be exempt from review by an institutional review board. Federal policy regarding Human Subjects Protection (known as the Common Rule) defines research and evaluation and describes the types of studies that must be reviewed.* Even if quitline evaluations were often judged to be exempt, we feel it is prudent to follow the Common Rule and obtain consent. Evaluations must still provide protection for human subjects and comply with privacy regulations.

It is important to recruit a representative group of quitline callers to participate in the evaluation. Research indicates non-respondents to health and lifestyle questionnaires are more likely to be tobacco users (Hill, Roberts, Ewings, & Gunnell, 1997; Bostrom et al., 1993; Kataniemi et al., 2001). Therefore, seeking consent for evaluation at the time of the six or seven-month follow-up survey call may result in a disproportionate number of callers refusing consent because they have not quit tobacco. Seeking consent at the time of intake, before the result of the quit attempt is known, should produce a more representative sample of survey contacts.

Sampling

Cigarette smoking and cessation behaviors have a strong seasonal component. Cigarette consumption, smoking initiation among youth, quitline call volume (Chandra & Chaloupka, 2003; Wellman & DiFranza, 2003) as well as quitline callers' stages of change (Delnevo, Foulds, Vorbach, & Kazimir, 2006) all exhibit distinct patterns of seasonal variation. In general, consumption and initiation are higher in the summer months, and quitline use and quit attempts are higher in winter months. These elements are also affected by external or environmental factors such as tobacco counter marketing media campaigns and the implementation of excise taxes and clean indoor air restrictions (Delnevo et al., 2006).

**For more information on complying with human subjects protection guidelines, see the Centers for Disease Control and Prevention online brochure: http://www.cdc.gov/epo/ads/HSR%20Brochure%202003.pdf.*

Exhaustive and random sampling represent two approaches to selecting callers to participate in the evaluation. *Exhaustive sampling* attempts to follow-up with all callers who register. This strategy will minimize response bias because no participants are left out. It will also provide follow-up data on the maximum number of callers possible. Depending on quitline call volume, however, this can be an expensive prospect and an inefficient use of resources.

Random Sampling produces a representative sample, allowing survey results to be generalized to the total population of callers. It also reduces evaluation costs by reducing the number of program participants that must be followed up. With proper sampling techniques the potential for response bias is reduced. Comparing callers in the follow-up sample to callers who were not followed-up on participant characteristics at intake or registration can help determine whether bias was introduced in the sampling. If a quitline elects to randomly sample, a strategy must be developed to ensure that a representative sample is selected.

In *Cohort or Time-Limited Sampling,* all callers who register during a limited time period are followed up. An advantage of this strategy is the time periods selected for sampling can be chosen to measure the effect of changes in protocol or the impact of environmental factors. Most importantly, the sample can produce sufficient numbers of completed surveys to calculate a quit rate in a short period of time. However, this method has the potential to introduce bias depending on seasonal and environmental factors.

In a *Rolling Sample***,** a random sample of callers is selected on an ongoing basis, such as weekly or monthly. This method eliminates the influence of seasonal variation and environmental factors on caller characteristics. It allows for examination of protocol changes or environmental factors and keeps costs down. The main limitation of this method is that it may yield a small number of completed follow-up surveys per sampling period (week or month), which may result in a delay before a quit rate can be produced. However, once a rolling sample and follow-up process is underway, both quarterly and cumulative quit rates can be reported.

A common question in sampling is how many people to sample—in general the larger the sample, the more accurate the results. Confidence intervals are a statistical calculation that can be used to help determine how large a sample is needed. A 95% confidence interval allows you to be 95% confident that the quit rate estimate from your sample is an accurate estimate of the "true quit rate" for all program participants within some calculated range, even if the rate can't be narrowed down to a single figure. Table 5 below illustrates a range of sample sizes, and the upper and lower bounds of the confidence interval (CI) for a 25% quit rate.* A more detailed explanation of how to calculate and use confidence intervals when reporting quit rates will be discussed in section 6. The main theme to note from the table below is that in general, the larger the sample size, the narrower the range and the more precise the quit rate estimate. However, diminishing returns are seen with larger and larger samples.

Table 5. Confidence Interval Ranges by Sample Size

Sample Size	25% Quit Rate	
	95% CI Lower Bound*	95% CI Upper Bound
50	15.9%	39.6%
100	17.6%	34.3%
200	19.5%	31.4%
400	21.0%	29.5%
800	22.1%	28.1%
1,000	22.4%	27.8%
1,500	22.7%	27.3%

**Confidence intervals calculated using the freely available Web site: http://faculty.vassar.edu/lowry/VassarStats.html.*

**Note that the upper and lower bounds for the confidence interval are not equidistant from the observed quit rate. When calculating the confidence interval for a proportion, the upper and lower bounds will be equidistant from the observed quit rate only if the proportion is 50%.*

We recommend a minimum sample size of N=400 (number of completed cases) where possible, given the need of the quitline for precision and the resources available. Sample size should be considered for all the groups for which you would like to calculate a quit rate (those who received a standard intervention, those who received an intensive intervention, Latino callers, African American callers, etc.). A sample size that results in a sufficiently precise rate should be selected for each group, if possible.

Recommendations

- *Solicit consent to follow-up early on, such as at the close of intake or registration. Consider consulting a Human Subjects Research or Institutional Review Board regarding consent procedures as appropriate.*
- *Report consent rate with the quit rate.*
- *Conduct follow-up on an ongoing rolling basis, with a random sample of registered callers.*
- *If a rolling follow-up is not feasible, use multiple cohorts or time-limited sampling (following up with callers who register during a limited time period) over the course of the year as a viable second choice.*
- *Where possible, and as decided based on individual program needs for precision and budgetary constraints, complete at least 400 follow-up surveys as part of an outcome evaluation.*

Mode

Common sense suggests that mode consistency is an important criterion in selecting a follow-up survey mode: telephone quitline callers would be best reached with a telephone follow-up survey. However, the changing landscape (increased use of cell phones and Caller ID and a growing refusal to accept calls from unknown sources) suggest that mixed-mode surveys may be a viable solution.

Across the board, quitlines are experiencing declining survey response, which mirrors the declining response to survey research nationwide (Curtin, Presser, & Singer, 2005). The lower the response rate, the less representative the sample may be, and generalization about the quitline population may be more difficult to make.

Mixed-mode surveys ask the same questions and offer the same response choices using two or more survey modes, such as Internet, telephone, interactive voice response or mail. Modes vary in two important ways: The interviewer's role (or lack thereof) and the technology employed.

The greatest benefit of mixed mode surveys is the potential to increase response rates (Dillman & Tarnai, 1988). The CDC's Pregnancy Risk Assessment Monitoring System (PRAMS) survey has had success using a mailed survey first, and then non-respondents are followed up by telephone. In 2002, the PRAMS unweighted response rates for individual states ranged from 58.5% - 86.1% (Williams et al., 2002).* Mixed modes can also help to reduce non-response error, which occurs when a significant number of people in the sample do not respond to the survey and have different characteristics (such as tobacco use status) than those who do respond. Providing more than one mode has potential to better reach all types of participants.

Despite these advantages, mixed mode surveys can also pose problems. One major concern is the increased variation in elapsed time from quit date to follow-up. Designs in which non-respondents to the first survey mode are attempted with a second mode necessarily take more time to implement than a single mode design. A second concern is mode effect: people's answers to any particular question may vary depending on the survey mode (Schwarz, Strack, Hippler, & Biship, 1991). Mode effects are caused by the presence or absence of an interviewer, differences between visual vs. aural communication and whether the interviewer or the respondent controls delivery of the stimulus.Common problems include cognitive processing of scales, primacy – recency

**http://www.cdc.gov/prams/2002PRAMSSurvReport/Appendix/AppB.htm.*

effects and response bias due to social desirability. For the most part, mode effects are controllable. Dillman (2007) recommends using a "unimode" survey construction in which survey items are designed to be received by the respondent in a similar way regardless of the mode of delivery.

Cost is also a concern. Usually adding a second survey mode will increase response rates *and* costs. Even when adding a relatively inexpensive mode, such as a Web survey, an economy of scale sufficient to reduce costs will likely be achieved only with a very large number of completed surveys. Finally, additional modes may require the collection of additional contact information. Participants may be reluctant to provide a mailing or e-mail address in order to receive a telephone-based intervention.

Recommendations

- **Use telephone as the mode of follow-up, as long as the target goal of a 50% response rate can be achieved.**
- *Depending on available budget, use one or more strategies to help reach a target 50% response rate include mailing pre-notification letters, providing monetary incentives and making at least 15 attempts to reach each participant.*
- *If the minimum response rate cannot be achieved, explore using a mixed-mode survey.*

The Evaluation Team

In the preceding sections we have made detailed recommendations regarding identifying the numerator and denominator for calculation of quitline quit rates. We have also made detailed recommendations regarding the appropriate methods for conducting follow-up surveys and highlighted the importance of survey response rates in interpreting quit rate findings. Because of these issues we believe it is important that a team conduct the evaluation with experience in quitline evaluation.

Quitline contract managers may elect to rely upon external evaluators or the quitline's internal evaluation staff. It is also possible for the evaluation to be conducted by the quitline provider. In either case, evaluators should abide by the Guiding Principles for Evaluators as set forth by the American Evaluation Association, stating that "evaluators should disclose any roles or relationships they have that might pose a conflict of interest (or appearance of a conflict) with their role as an evaluator," and that any actual or perceived conflicts "should be clearly articulated in reports of the evaluation results." In cases where outcomes evaluation is performed by the quitline provider, quitlines should follow CDC recommendations that evaluation staff must be entirely separate and independent of the counseling staff (CDC, 2004). Potential evaluators should be judged upon the transparency of their reporting (e.g. clearly defined sample selection, survey methods employed, number lost to follow-up and causes of loss to follow-up) and demonstrated ability to achieve adequate response rates on follow-up evaluation surveys. *Please note that the authors of this NAQC Issue Paper wish to acknowledge that they are from a university or an independent evaluation firm.*

Recommendations (Please note that the authors of this NAQC Issue Paper wish to acknowledge that they are from a university or an independent evaluation firm.)

- **Use an evaluator with experience in quitline evaluation.**
- **Evaluation may be conducted by an external evaluator or internally by quitline service providers as long as those individuals conducting the evaluation are entirely separate from and independent of the counseling staff.**
- **Select an evaluation team based upon transparency of reporting and demonstrated ability to achieve adequate response rates on follow-up evaluation surveys.**

VI. Calculating & Statistically Representing Quit Rates

Reporting abstinence rates is fairly straightforward, but requires several considerations. This section discusses reporting abstinence rates by service type, confidence intervals and reporting logistic regression results and odds ratios.

Confidence Intervals

As discussed in Section 5 above, in order to conserve resources we recommend using a sample of callers in a quitline evaluation. While evaluations based on a sample of callers result in abstinence prevalence rates that are only an estimate of all those that call the quitline, a confidence interval produces an accurate estimate within some calculated range, even if the quit rate can't be narrowed down to a single figure.

Confidence intervals for proportions may be calculated for free at the VasserStats Web site (http://faculty.vassar.edu/lowry/VassarStats.html). Many excellent Web sites provide this service, each of which may use a slightly different formula, so calculated confidence intervals may be slightly different. Once calculated, a confidence interval will be a range within which we can be sure to a certain degree of certainty (e.g. 95% sure) that the "true" quit rate falls somewhere within that range.

Many quitlines would like to compare their abstinence rates to those of other quitlines. Confidence intervals may be used as a crude measure to determine if two abstinence rates are different from another. If the confidence intervals for two rates do not overlap at all, the two rates are significantly different. However, if the confidence intervals for the two rates do overlap, they may be similar, or in some cases, significantly different. Therefore, examining the overlap in confidence intervals will identify most, but not all, significantly different abstinence rates.

Recommendations

- **Report confidence intervals with all abstinence rates and include the number of subjects used in the calculation.**

Logistical Regression & Odds Ratio

As stated earlier, each quitline serves a unique mix of people in terms of demographics and tobacco use history characteristics. Some of these characteristics are known to be related to higher or lower quit rates; therefore, presenting simple "absolute abstinence rates" can be "problematic" (Hughes et al., 2003). In addition to these intake characteristics, the type and degree of service used by the individual caller is also known to be related to quitting (dose-response).

Logistic regression is a statistical methodology that can show the relationship between an outcome and other relevant factors simultaneously such as program use, caller characteristics or other factors. Comparing logistic regression results can explain the impact of these factors on abstinence rates and make comparison of abstinence rates across quitlines more meaningful. In sum, logistic regression can help illustrate the context of a quitline by helping readers see abstinence rates in the context of the populations quitlines serve.

When applied to quitlines, we suggest the outcome (dependent variable) in logistic regression be 30-day abstinence for a particular quitline program. The controlling factors could be the number of sessions attended, program satisfaction and demographic and tobacco use history variables. Such a model would help the quitline staff and stakeholders understand and describe how service utilization, demographics and tobacco use history are related to the prevalence rate for their quitline.

A logistic regression produces odds ratios, which may be transformed into relative risks and differential probabilities of quitting for key variables in the model. We recommend all three statistics be reported. To see

how logistic regression results and accompanying odds ratios, relative risks and differential probabilities, see Liberman (2005) for a relatively easy to understand discussion. Conducting logistic regressions requires a certain level of statistical expertise, which some quitlines may not have access to. Therefore, these are suggestions, but not strong recommendations.

Recommendation

- *Explore the possibility of reporting abstinence rate findings as odds ratios via logistic regression. If conducted, relative risk and differential probabilities should also be reported.*

VII. Interpreting Quit Rates

Once an abstinence rate (i.e. the proportion of callers that are abstinent) is calculated for each intervention and reported with confidence intervals, it may appear easy to interpret. The number is specific and bounded by an interval. However, how does a quitline manager know their quit rate of 15.6% is good enough, for example?

One way of making an abstinence rate more meaningful is to view rates over time. If a rolling abstinence rate is calculated quarterly, fluctuations from one quarter to another can be assessed by comparing the point prevalences with their confidence intervals (see Section 6 for a fuller discussion of confidence intervals). Such fluctuations may reflect either a problem or a success for a quitline or a change in the 'type' of tobacco users that are calling the quitline. The literature has shown that quitline callers with different characteristics are more or less likely to quit using tobacco (specific examples will be discussed in *"Understanding Variation in Quit Rates"* below).

Another way of better understanding the success of a program through quit rates is to compare one quitline's quit rate to those of other quitlines. However, such comparisons must be made with caution. A state or province that specifically targets their quitline's services to a population that has greater difficulty quitting (e.g. uninsured callers) would expect a lower quit rate than a quitline that does not target such a group. The following section describes how understanding and reporting participant characteristics can help interpret abstinence rates, especially when they are compared to other quitlines.

Understanding Variation in Quit Rates

Even after the measurement of quit rates has been standardized, different programs may have different results. Having a lower quit rate is not necessarily a sign of lower quality services. Both participant characteristics and program characteristics can contribute to variation in quit rates.

Nicotine dependence has been found to be a very strong predictor of quitting behavior with those smoking fewer cigarettes per day and/or having a larger time period from waking to first cigarette, having a higher rate of quitting (Hyland et al., 2004; Hyland et al., 2006; Hymowitz et al., 1997; Velicer, Redding, Sun, & Prochaska, 2007). Therefore, it is also important to report nicotine dependence characteristics with quit rates and to consider them when interpreting quit rates.

Indicators of social disadvantage have also been shown to influence quit rates. Many studies have reported a strong correlation between level of quitting and one but usually not all of the following variables: income level, education level or socio-economic status (SES) (Hyland et al., 2004; Hymowitz et al., 1997; Lee & Kahende 2007; Velicer, Redding, Sun, & Prochaska, 2007). In some cases only one of these was found to have an effect on quitting, while in other cases only one of these variables was investigated. Co-morbidities such as mental health and alcohol use have also been associated with lower quit rates. In addition, several participant characteristics have been shown to influence outcomes. One of the main predictors of quitting across studies is age, with older smokers more likely to successfully quit than younger smokers (Hyland et al., 2004; Hymowitz et al., 1997; Lee & Kahende, 2007; Velicer, Redding, Sun, & Prochaska, 2007). Gender has also been found to be associated with quitting success. In general, studies have found that males have had

more success quitting than females (Hyland et al., 2004; Hymowitz et al., 1997). While some studies have found an increased chance of quitting for Non-Hispanic Whites (Lee & Kahende, 2007), this finding is not consistent across the literature. Other studies have found no difference for race or ethnicity (Hyland et al., 2006; Hymowitz et al., 1997; Velicer, Redding, Sun, & Prochaska, 2007).

Many of the indicators listed above can be reported using the MDS Intake Questions including age, gender, race/ethnicity, education level, cigarettes per day and time from waking to first cigarette. Income level and insurance status are not required items on the MDS, however, if available these should also be reported and considered. In addition, any other characteristics that distinguish a quitline's population served from other quitlines should be reported.

It is important to note that many of the studies regarding participant characteristics that affect quitting behavior used a general population of smokers, not specifically treatment-seeking smokers. It is possible that the relationship of predictors to outcomes found for the general population of smokers may not hold for smokers who proactively seek out cessation treatment. Also, it is worth noting none of the findings listed above are definitive, as the predictors varied across studies. In addition, a study of demographic differences in smoking cessation in the U.S. from 1950 to 1990 found trends changed across time (Gilpin & Pierce, 2002). Therefore, it is likely that indicators noted above may change or disappear and new indicators may appear.

Finally, a number of program characteristics are associated with variability in quit rates. As discussed in Sections 4 and 6, the evidence for a dose-response effect is large, so that the number of calls received is positively associated with higher quit rates (Rabius, 2007) and Stead, Perera and Lancaster (2008) found an association between intended program intensity and effect size. A wealth a literature finds a strong positive relationship between the availability of pharmacological therapy and its use (Stead, Perera, Bullen, Mant, & Lancaster, 2008). Other services, such as integrated Web-Internet interventions are emerging. However, it is too early to say how these services will influence abstinence.

Recommendations

- **Report information about the demographic and clinical characteristics of callers and program characteristics with quit rates.**
- **Use caution when comparing your quit rate to those of other quitlines. Consider the similarity of the quitline programs as well as the demographic and tobacco use characteristics of respondents.**

ACKNOWLEDGEMENTS

Authors:
NAQC would like to acknowledge the authors of this issue paper, Lawrence An, MD, University of Minnesota and Anne Betzner, PhD, Michael G. Luxenberg, PhD, Julie Rainey,BA, Traci Capesius, MPH, and Emily Subialka, BA, Professional Data Analysts, Inc. The authors were responsible for conceptualizing the paper, obtaining permission to use original empirical data, performing all analyses in the paper, drafting the original paper, and incorporating comments from NAQC staff, Advisory Council members, and NAQC general membership.

Contributors:
NAQC acknowledges Christopher Anderson, BA, Sharon Cummins, PhD and Shu-Hong Zhu, PhD for analyzing and contributing data from the California Smokers' Helpline to the paper.

For managing the feedback and review process, support of the authors and editing NAQC would like to acknowledge Tamatha Thomas-Haase, MPA and Jessie Saul, PhD. For final editing, layout, and design of the paper, NAQC would like to acknowledge Brenda Bryan, MAdmin.

NAQC would also like to acknowledge its Advisory Council members for their role in reviewing and approving this Issue Paper, most notably the three members who served as primary reviewers: Christopher Anderson, Mary Kate Salley and Scott Leischow, PhD. The feedback from NAQC members during the review and comment phases of the process was also invaluable. All member feedback ensured the paper's relevancy to the field.

Funders:
NAQC's Quality Improvement Initiative is made possible with funds from The Centers for Disease Control and Prevention, The American Cancer Society and The American Legacy Foundation. The contents of this publication are under the editorial control of NAQC and do not necessarily represent the official views of the funding organizations.

Recommended Citation:
NAQC. (2009). *Measuring Quit Rates. Quality Improvement Initiative* (L. An, MD, A. Betzner, PhD, M.L. Luxenberg, PhD, J. Rainey, BA, T. Capesius, MPH, & E. Subialka, BA). Phoenix, AZ.

References

American Evaluation Association. (July 2004). *AEA guiding principles for evaluators.* Retrieved from http://www.eval.org/Publications/GuidingPrinciplesPrintable.asp.

Benowitz, N., Ahijevych, K., Hall, S., Hansson, A., Henningfield, J.E., Hurt, R. D.,

Jacob, P., Jarvis, M. J., LeHouezec, J., Lichtenstein, E., Tsoh, J., &Velicer, W. (2002). Biochemical verification of tobacco use and cessation. *Nicotine and Tobacco Research, 4,* 149-159.

Bloomberg, S., & Luke, J. (2007). Wireless substitution: Early release of estimates based on data from the National Health Interview Survey, July – December 2006. National Center for Health Statistics. Retrieved from http://www.cdc.gov/nchs/nhis.htm.

Boström, G., Hallqvist, J., Haglund, B., Romelsjö, A., Svanström, L., & Diderichsen, F. (1993). Socioeconomic differences is smoking in an urban Swedish population. The bias introduced by non-participation in a mailed questionnaire. *Scand J Soc Med, 21 (2),* 77-82.

Cantor, D., O'Hare, B., & O'Connor, K. (2007). The use of monetary incentives to reduce nonresponse in random digit dial telephone surveys. Advances in Telephone Survey Methodology. In J. Lepkowski, C. Tucker, J. Brick, E. Leeuw, L. Japec, P. Lavrakas, et. al., *Advances in Telephone Survey Methodology* (pp. 471-498). John Wiley & Sons, Inc.

Centers for Disease Control and Prevention (2004). Telephone Quitlines: A resource for development, implementation, and evaluation. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Office on Smoking and Health.

Chandra, S., Chaloupka, F. (2003). Seasonality in cigarette sales: patterns and implications for tobacco control. *Tobacco Control, 12,* 105-107.

Cohen, S. J., Lichtenstein, E., Prochaska, J. O., Rossi, J. S., Gritz, E. F., Carr, C. R., Orleans, C. T., Schoenbach, V. J., Biener, L., Abrams, D., DiClemente, C., Curry, S., Marlatt, G. A., Cummings, K. M., Emont, S. L., Giovino, G., & Ossip-Klein, D. (1989). Debunking myths about self-quitting: Evidence from 10 prospective studies of persons quitting smoking by themselves. *American Psychologist, 44,* 1355-1365.

Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly, 69 (1),* 87-98.

De Leeuw, E., Callegaro, M., Hox, J., Korendijk, E., Lensvelt-Mulders, G. (2007). The Influence of Advance Letters on Response in Telephone Surveys: A Meta-Analysis. *Public Opinion Quarterly, 71 (3),* 413–443.

Delnevo, C., Foulds, J., Vorbach, U., & Kazimir, E. (2006). Seasonal variations in stage of change among Quitline clients. *Tobacco Control, 15,* 70-71.

Dew, Jr., D. E., Graber, J. E., Goble, L. , Liu, K. , Parsons, A. & Yager, N. (2004). Sponsorship and Selling: Telephone Interview Greetings and Respondent Cooperation at First Contact. *Paper presented at the annual meeting of the American Association for Public Opinion Research, Pointe Hilton Tapatio Cliffs, Phoenix, Arizona.*

Dillman, D. (1991). Mail surveys: a comprehensive bibliography, 1974-1989. Chicago, IL: Council of Planning Librarians.

Dillman, D. (2007) Mail and Internet Surveys: The Tailored Design Method, Second Edition. Hoboken, N.J.: John Wiley & Sons, Inc.

Dillman, Don A., Jon R. Clark, and Michael A. Sinclair. 1995. How Prenotice Letters, Stamped Return Envelopes, and Reminder Postcards Affect Mailback Response Rates for Census Questionnaires. *Survey Methodology, 21* , 1–7.

Dillman, D., & Tarnai, J. (1988). Mode effects of cognitively-designed recall questions: A comparison of answers to telephone and mail surveys. In P. Bremer, R. Groves, L. Lyberg, N. Mathiowetz, & S. Sudman, *Measurement errors in surveys* (pp.73-93). New York: Wiley.

References cont...

Fiore, M., Jaen, C., Baker, T. et al. (2008). *A clinical practice guideline Treating tobacco use and dependence: 2008 Update.* Rockville MD: USDHHS, U.S. Public Health Service. Retrieved from http://www.surgeongeneral.gov/tobacco/treating_tobacco_use08.pdf

Gilpin, E., & Pierce, J. (2002). Demographics differences in patterns in incidence of smoking cessation: United States 1950-1990. *AEP, 12 (3),* 141-150.

Glasgow, RE, Vogt, TM, & Boles, SM. (1999). Evaluating the Public Health Impact of Health Promotion Interventions: The RE-AIM Framework. *American Journal of Public Health. 89,* 1322-1327.

Groves, R., Biemer, P., Lyberg, L. (Eds, 1988). Telephone Survey Methodology. NY: Wiley.

Groves, R., Dillman, D., Eltinge, J., Little, R.(Eds, 2001). Survey Nonresponse. NY: Wiley.

Hall, S., Delucchi, K., Velicer, W., Kahler, C., Ranger-Moore, J., Hedeker, D., Tsoh, J., & Niaura, R. (2001). Statistical analysis of randomized trials in tobacco treatment: longitudinal designs with dichotomous outcome. *Nicotine & Tobacco Research, 3,* 193-202.

Hembroff. L., Rusz. D., Rafferty, A., McGee, H., & Ehrlich, N. (2005). The Cost-Effectiveness of Alternative Advance Mailings in a Telephone Survey. *Public Opinion Quarterly, 69 (2),* 232-245.

Hill, A., Roberts, J., Ewings, P., & Gunnell, D. (1997). Non-response bias in a lifestyle survey. *J Public Health Med., 19 (2),* 203-207.

Hughes JR, Gulliver SB, Fenwick JW, Valliere WA, Cruser K, Pepper S, Shea P, Solomon LJ, Flynn BS. (1992). Smoking cessation among self-quitters. *Health Psychology 11*(5), 331-334.

Hughes, J., Keely, J., Niaura, R., Ossip-Klein, D., Richmond, R., & Swan, G. (2003). Measures of abstinence in clinical trials: Issues and recommendations. *Nicotine & Tobacco Research, 5,* 13-25.

Hunt, W. A. & Bespalec, D. A. (1974). An evaluation of current methods of modifying smoking behavior. *Journal of Clinical Psychology, 30,* 431-438.

Hyland, A., Li, Q., Bauer, J., Giovino, G., Steger, C., & Cummings, K. (2004). Predictors of cessation in a cohort of current and former smokers followed over 13 years. *Nicotine & Tobacco Research, 6 (Suppl 3),* 363-369.

Hyland, A., Borland, R., Li, Q., Yong, H-H., McNeill, A., Fong, G., O'Connor, R., & Cummings, K. (2006). Individual-level predictors of cessation behaviours among participants in the International Tobacco Control (ITC) Four Country Survey. *Tobacco Control, 15 (Supplement 3),* 83-94.

Hymowitz, N., Cummings, K., Hyland, A., Lynn, W., Pechacek, T., & Hartwell, T. (1997). Predictors of smoking cessation in a cohort of adult smokers followed for five years. *Tobacco Control, 6 (Suppl 2),* 57-62.

Kotaniemi, J., Hassi, J., Kataja, M., Jönsson, E., Laitinen, L., Sovijärvi, A., & Lundbäck, B. (2001). Does non-responder bias have a significant effect on the results in a postal questionnaire study? *European Journal of Epidemiology, 17 (9),* 809-817.

Lee, C., & Kahende, J. (2007). Factors associated with successful smoking cessation in the United States, 2000. *American Journal of Public Health, 97 (8),* 1503-1509.

Lepkowski, J. M., Tucker, C., & Brick, J. M. (2008). Advances in Telephone Survey Methodology. Hoboken, NJ : Wiley-Interscience.

Liberman, A. (2005). How much more likely? The implications of odds ratios for probabilities. *American Journal of Evaluation, 26 (2),* 253-266.

Link, M., Mokdad, A., Kulp, D., Hyon, A. (2006). Has the National Do Not Call Registery Helped or Hurt State-level Response Rates? A Time Series Analysis. *Public Opinion Quarterly.* Vol. 70, No. 5, Special Issue, pp. 794-809.

References cont...

North American Quitline Consortium. (2008). *Measuring Reach of Quitline Programs.* Cummins, S, Editor. Phoenix, AZ: North American Quitline Consortium.

North American Quitline Consortium Minimal Data Set Intake & Follow-Up Questions, July 19, 2005.

North American Quitline Consortium. (2008). Realizing Opportunities: Implementation Assessment of the Minimal Data Set in North America. Minimal Data Set Workgroup, Editors, Phoenix, AZ: North American Quitline Consortium.

Patrick, D., Cheadle, A. Thompson, D., Diehr, P., Koepsell, T., & Kinne, S. (1994) The validity of self-reported smoking: a review and meta-analysis. *American Journal of Public Health, 84,* 1086-1093.

Rabius, V., Pike, K., Hunter, J., Wiatrek, D., & McAlister, A. (2007). Effects of frequency and duration in telephone counseling for smoking cessation. *Tobacco Control, 16 (Suppl 1)*, i72-i74.

Rainey, J. & Huggins, M. (2007, October). Strategies for Evaluating Cessation Programs – Improving Response Rates. Paper presented at the National Conference on Tobacco or Health, Minneapolis, MN.

Schwarz, N., Strack, F., Hippler, H., Bishop, G. (1991). The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology, 5,* 193-212.

Singer, E., Groves, R., & Corning, A. (1999). Differential incentives: Beliefs about practices, perceptions of equity, and effects on survey participation. *Public Opinion Quarterly,* 63 (2), 251-260.

Singer, Eleanor. (2002). The use of incentives to reduce nonresponse in household surveys." In R. Groves, D. Dillman, J. Eltinge, & R. Little, *Survey Nonresponse* (pp. 163–78). New York: Wiley.

Singer, E., Hoewyk, J., & Maher M. (2000). Experiments with Incentives in Telephone Surveys. *Public Opinion Quarterly,* 64, 171-188.

Stead, L., Perera, R., & Lancaster, T. (2008). Telephone counselling for smoking cessation (Review). *Cochrane Database of Systematic Reviews 2008*, Issue 3.

Stead, L., Perera, R., Bullen, C., Mant, D., & Lancaster, T. (2008). Nicotine replacement therapy for smoking cessation. *Cochrane Database of Systematic Reviews 2008*, Issue 1. Art. No.: CD000146. DOI: 10.1002/14651858.CD000146.pub3.

Tuckel, PS; O'Neill, H. (2002). The vanishing respondent in telephone surveys. *Advertising Res. 42*:26–48.

Velicer, W. F., Prochaska, J. O., Rossi, J. S., & Snow, M. G. (1992). Assessing outcome in smoking cessation studies. *Psychological Bulletin, 111,* 23-41.

Velicer, W., Redding, C., Sun, S., & Prochaska, J. (2007). Demographic Variables, smoking variables, and outcome across five studies. *Health Psychology, 26 (3)*, 278-287.

Velicer, W., & Prochaska, J. (2004). A comparison of four self-report smoking cessation outcome measures. *Addictive Behaviors, 29*, 51-60.

Wellman, R., & DiFranza, J. (2003). Seasonality in onset of youth smoking parallels seasonality in cigarette sales. *Tobacco Control, 12,* 339-340.

West, R., Hajek, P., Stead, L., & Stapleton, J. (2005). Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction, 100*, 299-303.

Williams, L., Morrow, B., Shulman, H., Stephens, R., D'Angelo, D., Fowler, CI. (2002). *PRAMS 2002 Surveillance Report.* Atlanta, GA: Division of Reproductive Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention. Retrieved from http://www.cdc.gov/prams/2002PRAMSSurvReport/.

Important Note About the Recommendations Contained in this Issue Paper:

Recommendations that are necessary in order to implement the standard are found in **bold-print** and recommendations that are viewed as important but not critical to implementation are italicized.

Learn more about NAQC: www.naquitline.org
3030 N. Central Avenue, Suite 602, Phoenix, AZ 85012-2713
Ph: 602.279.2719. Fax: 602.279.2740. Email: naqc@naquitline.org.

NORTH AMERICAN QUITLINE CONSORTIUM