



# A Comparative Review of the Effectiveness Trial of N-O-T in Alabama: Guideposts for Future Research

Kimberly Horn and Geri Dino

## ABSTRACT

*The American Lung Association's Not On Tobacco (N-O-T) program is federally recognized as an accessible and effective option for teen smoking cessation. The program is the most widely used teen smoking cessation program in schools and communities across the U.S. Over a decade of research demonstrates the program's solid evidence base, with overwhelmingly positive findings for both smoking cessation and reduction. A recent publication detailing a Not On Tobacco trial in Alabama reported low impact and negative findings. Given the program's widespread use across the country, it is important to understand differential outcomes between the Alabama study and past Not On Tobacco studies, and to explore what we can learn from these differences. The current manuscript: (1) details and compares the methodological differences between the Alabama study and previous Not On Tobacco efficacy and effectiveness studies in other states, (2) explores important issues surrounding the analysis of teen smoking cessation outcomes, and (3) provides guidance for future research for Not On Tobacco and other teen smoking cessation interventions.*

*Horn K, Dino G. A comparative review of the effectiveness trial of N-O-T in Alabama: guideposts for future research. Am J Health Educ. 2009;40(1):47-55. This paper was submitted to the Journal on April 14, 2008, revised and accepted for publication on September 10, 2008.*

## INTRODUCTION

Nearly one-fourth of U.S. high school teens currently smoke and most who smoke now will become adult smokers. Understandably, teen smoking cessation remains a public health priority. The American Lung Association's (ALA) Not On Tobacco (N-O-T) program is acknowledged as an accessible and effective option for teen smoking cessation. N-O-T is consistent with the current Youth Tobacco Cessation Collaborative guidance<sup>1</sup> and recommendations in the teen smoking cessation literature.<sup>2-4</sup> Its evidence base has been federally recognized by the Substance Abuse and Mental Health Services Administration, National Cancer Institute, and the Centers for Disease Control and Prevention, among other organizations.<sup>5,6</sup>

Over the past decade, the National

Office of the ALA and the developers of N-O-T have welcomed, and will continue to welcome, independent replications of past N-O-T efficacy and effectiveness trials. Repeated independent investigations are the vehicle by which intervention programs reach their maximum potential and success. Although independent empirical investigations of N-O-T have been published,<sup>7</sup> no true replications exist. The most recent of the independent investigations is presented in an article by Kohler and colleagues,<sup>8</sup> in the *American Journal of Health Behavior*. Unlike other N-O-T studies, the study by Kohler et al. reported low impact and negative findings among Alabama teens. When studies of the same intervention program achieve discrepant results, methodological comparisons may help understand the discrepancies.

Given that only a handful of empirically tested teen smoking interventions exist, the opportunity for this type of comparative review is rare. Relevant to the N-O-T investigations at hand, two questions arise, "What are the reasons for the differences in study outcomes?" and "What can we learn from them?" As one of only two published independent investigations of N-O-T, the Kohler

---

*Kimberly Horn is an associate professor of Community Medicine and associate director in the Mary Babb Randolph Cancer Center, West Virginia University, Morgantown, WV 26505; E-mail: khorn@hsc.wvu.edu. Geri Dino is an associate professor of Community Medicine and director of the Prevention Research Center, West Virginia University, Morgantown, WV 26505; E-mail: gdino@hsc.wvu.edu.*



et al. study provides an opportunity to seek answers to these questions. The answers may enhance understanding of the N-O-T program, in particular, and of teen smoking cessation interventions, in general.

The current manuscript addresses the above referenced questions by: (1) detailing and comparing the methodological differences between the Kohler et al. study and previous N-O-T efficacy and effectiveness studies; (2) exploring important issues surrounding the analysis of smoking cessation outcomes identified by Kohler et al.; and (3) providing guidance for future research in teen cessation intervention.

### *About N-O-T*

*A brief review.* N-O-T was adopted by the ALA in 1998.<sup>9</sup> N-O-T is designed for 14-to-19 year-old teens who are daily smokers, likely to be addicted, and who volunteer to participate in school or community settings. Influenced by Social Cognitive Theory<sup>10</sup> and the Transtheoretical Model,<sup>11</sup> N-O-T includes 10 hour-long weekly sessions and 4 booster sessions, designed for delivery in same-gender groups by ALA-trained facilitators. Major program goals are to help participants: (1) quit smoking; (2) reduce the number of cigarettes smoked by youth who are unable to quit; (3) increase healthy lifestyle behaviors (e.g., physical activity and nutrition); and (4) improve life skills such as stress management, decision making, coping, and interpersonal skills. A detailed description of the N-O-T program can be found elsewhere.<sup>12,13</sup> The ALA provides N-O-T training nationally using a regional train-the-trainer structure, with 10 Master Trainers. Refer to ALA's website (<http://www.lungusa.org>) for information about program implementation and training. N-O-T has a notable research base and is recognized as an evidence-based program. For instance, a recent Cochrane Review noted the N-O-T program as the only teen smoking program of promise.<sup>14</sup> The Cochrane scientists reviewed the findings from all teen smoking cessation programs ever published, worldwide.

### *N-O-T Evaluation Overview*

*Effectiveness vs. Efficacy.* Public health experts<sup>15,16</sup> and youth tobacco cessation

researchers<sup>17,18</sup> recommend two types of program evaluations for interventions such as N-O-T: effectiveness and efficacy trials. Over the past 10 years, the N-O-T developers, in collaboration with the National Office of the ALA, local ALAs, schools, and communities across the U.S.,<sup>12</sup> conducted both efficacy and effectiveness evaluations with N-O-T.<sup>19</sup> N-O-T research, as approved by the ALA, follows a two-pronged approach. The first is efficacy research, whereby implementation is tightly controlled by scientific investigators. The second is effectiveness evaluation that transports or translates the intervention into real-world school or community settings, with less researcher involvement. The common assertion is that efficacy evaluation is high on internal validity, but may be less generalizable; whereas, effectiveness evaluation is high on external validity at the expense of rigor and control. It is not surprising that researchers traditionally favor efficacy; whereas, practitioners prefer effectiveness. The position of the ALA and N-O-T developers is that this two-pronged approach (that is, using simultaneous efficacy and effectiveness evaluation) fosters program sustainability.

*Replication.* Replication, by definition, involves repeating a study using the same methods, with different participants, different investigators, and in different settings. Generally, the purpose is to: (1) provide assurance that previous results are valid and reliable; (2) determine the generalizability of the methods, techniques, and/or results; (3) apply the previous methods, techniques, and/or results to real-world conditions; and (4) inspire new research. Replication can occur with both efficacy and effectiveness studies.

Before we begin our comparative review and discussion, it is important to underscore that we value Kohler et al.'s interest in conducting N-O-T research. The Alabama N-O-T study approached replication, but a number of the study's methodological features warrant discussion. Ideally, a constructive critique of these issues will aid others who conduct N-O-T trials. Broadly, these include differences between: (1) the

Kohler et al. methodology and that of other published N-O-T evaluation studies; (2) the Kohler et al. protocol and the ALA-approved N-O-T protocol; and (3) Kohler's outcomes measurement and analytic procedures and those of other N-O-T studies as well as the field emergent standards. We discuss each of these differences next.

### **METHODOLOGICAL CONSIDERATIONS**

Until recently, there were limited data to guide teen smoking cessation programming;<sup>1</sup> this was more pronounced ten years ago when N-O-T first began than it is today. Because of research conducted with two federally-recognized teen smoking cessation programs,<sup>20</sup> N-O-T and Project Ex,<sup>21</sup> the field has gained valuable knowledge about cessation programming. Our experiences over the past ten years have shown that N-O-T research implementations require careful consideration as to design, measurement, and other methodological issues. Kohler and colleagues worked with the local Alabama ALA; however, the National Office of the ALA was not involved. This exclusion limited the opportunity for the Alabama study to serve as a true replication of other N-O-T effectiveness studies, and to inform N-O-T program improvements and field advancements across the U.S.

*Matched-pairs design.* Study design is a critical feature of any intervention study. The teen smoking cessation field struggles with design issues, particularly those related to matching versus randomization. Some experts view matching as an alternative to randomization.<sup>19,22</sup> Matching is the process of selecting a predetermined number of matched pairs (i.e., treatment and comparison schools/sites) that are equivalent on several factors potentially influencing the outcomes. Matching is a tedious and systematic process that requires information from multiple sources, including the insights of community and school professionals and local and state-level surveillance data.<sup>22</sup> The process promotes precision and allows control for complex factors that may otherwise be difficult to measure. Similar to



past N-O-T studies, Kohler and colleagues utilized a matched design. Appropriately, the matching criteria included factors such as student enrollment, racial composition, percent of free/reduced lunch, and urban/rural status—factors that may influence anti-tobacco programming. For matching to be optimally effective, it is ideal to have equivalent pairs, in both number and composition, in the treatment and comparison conditions. Kohler et al. included 44 intervention and 27 comparison schools—a respectable number of schools. One of the challenges of matching is recruiting equal and equivalent group sizes. The Kohler et al. sample size was adequate. The large school sample size, however, was compromised by imbalanced intervention and comparison conditions—there were 17 fewer comparison schools. In any matched design, lack of equivalence raises concerns about homogeneity of the study groups prior to intervention.<sup>17,18</sup> Achieving equivalency is challenging. As other researchers employ matched designs for teen cessation trials, it may be useful to analyze and report school-level (or other units of matching) comparisons on the matching factors. This type of analysis aids our understanding of baseline equivalency.

*Use of brief intervention comparison.* The use of brief intervention or “care as usual” comparison groups is an accepted standard in teen smoking cessation trials. Kohler and colleagues used a brief intervention comparison group, as in other N-O-T trials. Unfortunately, the brief intervention materials and procedures were not described. The authors reported only that the comparison group received “... printed smoking cessation pamphlets.” Moreover, it is unclear who delivered the brief intervention, for how long, and at what point in the study timeline the comparison group received the intervention. More specifically, an important consideration related to the brief intervention is the timing of delivery. For example, past N-O-T studies attempted to approximate the brief intervention delivery with the intended “quit week” for N-O-T participants, giving teens in both groups

equal opportunity to quit prior to the initial post-baseline follow up. Detailed information on comparison group treatment, especially in seminal publications, may be useful to others who attempt N-O-T replications. Additionally, descriptions of materials and methods used in a brief intervention can inform program differentiation—that is, the extent to which the comparison group is different from or similar to the N-O-T intervention group. Ideally, researchers that conduct N-O-T trials should strive toward the use of the same or comparable brief intervention materials.

### PROTOCOL AND FIDELITY CONSIDERATIONS

*N-O-T implementation fidelity.* By definition, implementation fidelity is how well an intervention or program is implemented compared to how it is intended or prescribed. Experts<sup>23</sup> posit five criteria for measuring fidelity of implementation: (1) adherence—whether the intervention is being delivered as designed; (2) duration—the number, length, or frequency of sessions implemented; (3) quality of delivery—the manner in which the facilitator delivers the program related to methods prescribed; (4) participant responsiveness—the extent to which participants are engaged; and (5) program differentiation—whether critical features that distinguish the program from the comparison condition are present or absent during implementation.

Fidelity of implementation is one of the reasons that the ALA utilizes a national train-the-trainer structure and a standardized curriculum. To that end, facilitators use a highly prescribed N-O-T curriculum when they implement the program. To increase adherence, all N-O-T facilitators must be trained and certified to use the program. The authors of the Alabama study report that they allowed facilitators to “... vary program activities as necessary for their situations and settings...” They indicate that this resulted in two major changes from other published N-O-T studies—mixed gender groups and duration of sessions. They also report “other variations occurred” but do not describe

those variations. These issues cited by Dr. Kohler bring to the forefront the important and often challenging issue of program fidelity. Although implementation fidelity may have been tracked in the Kohler et al. Alabama study, it was not documented in the article. We appreciate the importance of flexibility in real-world implementations and commend Kohler et al. for being responsive to school needs. For the purposes of this manuscript, we address those fidelity concerns related to the key N-O-T features.<sup>9</sup> Specifically, some of the fidelity violations in the Alabama study raise questions about how much of the program participants actually received and the extent to which program variations impacted the outcomes.

One of the unique features of the N-O-T program is that it is gender-tailored. In particular, the curriculum provides gender-specific content and recommends same gender groups using same gender facilitators.<sup>9</sup> We recognize that some schools across the U.S. use mixed gender N-O-T groups and, in many cases, these are the only conditions under which the program can be offered. However, over the past ten years, facilitators across the country have indicated that they obtain higher quit rates with same-gender than with mixed-gender groups. Additional rationale for these recommendations is documented elsewhere.<sup>13</sup> Pertinent to program adherence—a feature of fidelity—according to Kohler et al.,<sup>8</sup> most of the groups in their study did not use gender-separate groups. In contrast, all of the matched-group design studies reported by Horn and Dino in past N-O-T trials used same-gender groups, per ALA recommendations. Most facilitators in the Alabama study were female; whereas, most participants were male. In fact, of 44 schools, Kohler et al. reported that there were only 12 male facilitators. Of note, N-O-T participants in past studies report that they prefer same-gender groups.<sup>24</sup> Although we encourage using the program as intended, we prefer that schools use mixed gender groups rather than not use the program at all. We acknowledge facilitator recruitment challenges, particularly among males. One of the advantages to communicating and



collaborating with the National Office of the ALA is access to knowledge and resources related to facilitator and youth recruitment. Going beyond pure theoretical and developmental justification, we need more research to determine if gender separation and same gender facilitators are associated with higher sustained quit rates among N-O-T participants.

Program duration is another important feature of fidelity. N-O-T includes 10 sessions (4 booster sessions are also available), that are theoretically and sequentially ordered to address the critical processes of quitting. We understand that in real-world conditions, events and situations occur that prohibit or limit sessions. Past assessments show that N-O-T participants complete about 70% of the program (i.e., 7/10 sessions).<sup>5</sup> Certainly, dosage or duration is associated with cessation outcomes.<sup>3</sup> Kohler and colleagues report that in their study only 23 of 44 schools turned in attendance/session logs. Without the session logs, attendance cannot be accurately determined. Low attendance implies low program dosage. Kohler et al. also reported that the N-O-T group sizes ranged from 1 to 20 teens. The ALA's national N-O-T training protocol recommends a maximum group size of 12 teens. Grounded in social cognitive theory, a key element of N-O-T is social support, including peer and adult modeling. The level of social support and peer dynamics can vary significantly between groups with one versus twenty teens. Such features are intended to facilitate participant engagement—also a measurable feature of fidelity. Fidelity issues were significant in the Alabama study. Fidelity of implementation should not be minimized by researchers investigating N-O-T or any other cessation intervention. As the lead N-O-T researchers, we also have failed to meet adequate standards in fidelity measurement. Intervention researchers conducting efficacy and effectiveness studies using N-O-T should routinely measure fidelity. Fidelity measures are essential to explain the degree of variation in implementation across settings and how it might mediate or moderate smoking cessation outcomes.

## MEASUREMENT AND ANALYSES CONSIDERATIONS

*Baseline differences.* The Alabama study examined individual baseline differences between the N-O-T and comparison teens. They compared teens on 7 variables (past N-O-T studies used 11 or more). Previous N-O-T studies included baseline variables of comparison that could potentially influence outcomes, as demonstrated in relevant literature. In the Kohler et al. study, N-O-T and comparison group teens were significantly different on several individual baseline factors. Basic research methods underscore that the more similar the groups, the more credible the results. When significant baseline differences emerge, as in the Kohler et al. study, it may suggest that the matching process was not optimal. If matching is performed correctly, like randomization, minimal observed differences should be found at the study outset. Past N-O-T research using a matched design reported baseline comparisons at the level of both the school and the individual, with few differences observed.<sup>22,25</sup> Future N-O-T research using matched design could be greatly improved by standardized matching procedures and matching variables widely accepted and validated by the field.

Imbalances between the N-O-T and comparison participants or schools may also influence results. Differences at baseline could show a treatment effect where none exists; alternatively, imbalances could mask a real effect. As an example, Kohler et al. argue that the baseline differences found in their study would be in favor of N-O-T because the comparison group had lower expectancies for quitting, negating any treatment significance for N-O-T. An alternative perspective is possible. Close examination of the baseline characteristics suggests that the overall Alabama sample may have begun the study in a low state of readiness. Specifically, over half of the sample was somewhat-to-low motivated. It also appears that most were not daily smokers. Kohler et al. did not report daily smoking, and instead used classifications of light, moderate, or heavy smoking. The “heavy” classification required

smoking on  $\geq 10$  days in past 30 days, averaging 6 cpd. Past N-O-T studies report means of up to  $\geq 20$  cpd at baseline. Even with the minimum criteria used by Kohler et al., only 61% of the sample was considered “heavy” smokers. The Kohler et al. study is one of many teen smoking studies challenged by smoking status classification. Researchers can make important contributions to the field by developing and testing classifications for smoking status. Certainly, the Kohler et al. method is one option. Nonetheless, until we work together to consistently and accurately classify smokers (and quitters) for research purposes, we are delaying progress. Also of note, the mean age of the Alabama teens was almost 18 years old, compared to a mean age of 16 in past N-O-T trials. The Alabama study did not permit teens younger than 16 to enroll in the study. Taken together, these comparisons suggest important differences between the types of smokers who enrolled in Alabama and those who participated in other N-O-T trials and in real-world implementations. Our points are not intended to dismiss or negate the Kohler et al.'s findings; rather, we emphasize that between group baseline differences should be carefully examined for clinical relevance as well as statistical significance, and that field-driven standards are necessary to further the empirical basis of the field.

*Missing Data.* Retention is a common problem in teen cessation studies. The Kohler et al. report acknowledges that the loss-to-follow-up was a “major problem” in their study. Despite this acknowledgement, the authors posit that an analysis of missing data patterns revealed that differences between groups did not change for most variables over time. Reasons for participants' non-attendance at follow up and the resulting missing data should be closely examined to determine whether reasons are potentially related to smoking behaviors or cessation outcomes. Specifically, a critical issue is whether or not missing participant data occur by choice, or by chance.<sup>26</sup> There are three possible scenarios involving missing data:<sup>26</sup> (1) data are completely missing at random (MCAR)—unrelated to any study



variable—such as, a youth's parent was relocated to another school district during the course of the study, school rezoning occurs, or a youth was taken out of class to attend another extra curricular activity; (2) data are missing at random (MAR)—absence is not definitively but could be related to a study variable—such as, a youth is not available at follow up because he or she is in in-school suspension during the time of data collection; and (3) data are not missing at random (NMAR)—absence is correlated with key study outcomes—such as, when a participant is absent from follow up because he or she does not want to admit smoking, or because he or she discontinued the program because of lack of interest in quitting.

The analyses used in the Kohler et al. study (i.e., generalized estimating equations or GEE) make stringent assumptions of missing data. Specifically, GEE assumes that data are MCAR. GEE has become a standard method for analyzing non-normal longitudinal data and other correlated response data, but it is not commonly used by addiction intervention researchers. GEE is a more stringent assumption than similar analyses that assume that data are missing at random, or MAR. Importantly, the analyses of missingness for MCAR have distinct assumptions.<sup>27</sup> The analyses conducted by Kohler et al. suggest that a substantial amount of missing data were *not* MCAR, and therefore, may dilute the appropriateness of GEE for these analyses. Traditionally, data are considered MCAR when the probability that an observation is missing is unrelated to the value of any other variables. For example, Kohler et al. found that unavailability for follow up was significantly related to gender; as such, these data are not MCAR because missingness is correlated with gender. An alternative to GEE is maximum likelihood (ML). Both GEE and ML are relevant when there are data from completers and non-completers of an intervention program, such as N-O-T. Research shows that for completers-only data, GEE and ML analyses show similar results. However, GEE and ML analyses of a full data set may reveal marked differences. It could be concluded that both

the analyses of the completers only and the GEE analysis of the full data set produce misleading conclusions about the relationships between the response and covariates. The Kohler et al. study brings to light the importance of considering various types of missing data and suitable analyses. Critically, there are alternative ways to assess missing data patterns and relevant outcomes in teen smoking cessation programs. The Kohler et al. study presented an option not typically used in our field. Their method is worth further examination. An equally important issue, however, is to determine the nature of the missing data. Future teen smoking cessation studies, including our own N-O-T investigations, should consider tracking and documenting the extent to which data are NMAR, MAR, or MCAR. As a field, we have fallen short in our examination and understanding missing data.

*Quit rates.* The analysis of quit rates is probably the single most important issue in teen smoking cessation. The Kohler et al. paper criticizes past N-O-T research for reporting 24-hour quit rates at end-of-program. We acknowledge that there are multiple ways to assess smoking cessation outcomes. Teen smoking cessation interventions have been developed by researchers from a variety of disciplines. Thus, outcome measures have been explored in different ways<sup>28</sup> over the years. To date, our N-O-T research has reported past 24-hour point prevalence, while documenting mean days of continuous abstinence<sup>19</sup> at end of program, 6 months, and 15 months—an approach used by other tobacco cessation researchers.<sup>29-31</sup>

Undoubtedly, the teen smoking cessation field continues to struggle with consistent standards of outcomes measurement.<sup>2,29</sup> The examination of quit rates is a hotly debated topic. The field of smoking cessation for both adults and youth traditionally examines quit rates using one or more of the following conditions:<sup>28,32</sup> (1) 24-hour point prevalence abstinence; (2) 7-day point prevalence abstinence; (3) 30-day prolonged abstinence; and (4) 6-month prolonged abstinence. Point prevalence abstinence is the proportion of participants

not smoking at a certain point in time.<sup>28</sup> Continuous abstinence is the proportion of participants not smoking at all since the critical “quit date.” Finally, prolonged complete abstinence is absence of smoking for a long period of time.<sup>28</sup> Tobacco cessation experts maintain that the cutpoints for these smoking cessation outcome measures are historically arbitrary. To address this issue, in a study published in *Addictive Behaviors* in 2004, Velicer and Prochaska<sup>28</sup> compared four different outcome measures from three population studies. They made the following conclusion (p.51):

“The first three measures (24-hour point prevalence, 7-day point prevalence, and 30-day prolonged abstinence) all correlated in excess of .98 with each other. The only measure that did not demonstrate the same degree of almost perfect equivalence was 6-month prolonged abstinence, but even here, the lowest correlation with the other three measures was .82. For practical purposes, the first three measures will result in the same conclusions when used as outcome measures in smoking cessation studies.”

A larger issue relates to the clinical relevance of our cessation outcomes measures. What does each of these measures tell us about why one measure is more important, or better, than another one? Do we know that a teen who reports quitting for 30 days versus 24 hours at the end of an intervention is more likely to show prolonged abstinence six months later? The types of outcome measures we employ are important, to the extent that they predict some future outcome, including sustained abstinence. For programs, such as N-O-T, that employ the facets of the Transtheoretical Model of behavior change (i.e., the stages of change<sup>11,33,34</sup>) 24-hour point prevalence is consistent with the staging algorithm. For example, according to Prochaska, “...a 24-hour point prevalence abstinence measure represents the percentage of participants who are taking action at follow-up.”<sup>28</sup> The N-O-T program compels teens to act to quit smoking. If they have moved into action by the end of the pro-



gram, it is relevant and meaningful, even if they do so at the very end of the program, thereby meeting an expectation to quit. For long-term follow up purposes, consistent with stages of change, a 6-month prolonged abstinence measure would reflect the maintenance stage. Undoubtedly, the percentage means of quitting will decrease over time, as our long-term follow up studies have demonstrated.<sup>28</sup> At the very least, 24-hour end-of-program cessation is an intermediate outcome not to be minimized.<sup>28</sup>

An emerging opinion is that 30-day abstinence is the optimal criterion for a successful cessation among teen smokers.<sup>29</sup> We support Kohler et al.'s efforts to push the field on this issue and agree that a 30-day sustained abstinence measure may be appropriate at 6- or 12-month follow-up. However, we assert that a 30-day point prevalence measure taken at the end of the N-O-T program is inappropriate, as it might also be for other multi-session cessation programs where the quit date occurs in the middle of the program. First, if N-O-T is delivered according to protocol, 30 days prior to the end of the program tracks back to session five, which is "quit week" - the week that teens set their goal to quit. It would be ideal if teens (or humans, in general) could fulfill behavior change goals without waver. We know that is not the case with teen smoking; quitting is a fluctuating, non-linear process,<sup>35-37</sup> which is why the program holds 10 sessions. N-O-T teens have reported that they use the full length of the program to quit.<sup>38</sup> A 30-day point prevalence period must provide the opportunity for teens to quit smoking. Second, Kohler et al. reported that facilitators were allowed to change the program to meet their school needs, including varying session offerings. If the program were not implemented with fidelity, many teens may not have had the opportunity set their quit date or to attempt quitting 30 days prior to the end-of-program data collection. If taken at face value, it is important to reiterate that despite the restrictive criteria used by Kohler et al., the study found that over four times more N-O-T than comparison teens quit smoking. The time x treatment interaction

was significant at the end of the program. Kohler et al. did not find significant effects using the 30-day point prevalence measure at 6- or 12-month follow up. As supported by Kohler et al., this absence of statistical significance is likely due to the high attrition. The low quit rates for Alabama teens across conditions may also be a factor. We discuss this issue later in this article.

*ITT (Intent to Treat) vs. Compliant (or "Completers") analyses.* The teen smoking cessation field continues to debate the application of ITT versus Compliant analyses<sup>29</sup> for calculating quit rates. ITT analysis takes into account all participants lost to follow-up—those who do not attend follow-up observations are considered to be program failures (i.e., it is assumed that they did not quit smoking) and remain in the denominator (total number of quitters/total enrolled at baseline). Compliant analyses considers the quit rates of those who complied or completed an intervention, acknowledging that with teens, in particular, reasons for loss to follow up may not be related to smoking (total number of quitters/total available at follow up). This is relevant to our earlier discussion of the missingness of data. For example, past N-O-T research found that reasons for not attending follow up may be due to extracurricular activities, work, graduation, transfer, home schooling, sickness, and so on. These factors would be considered missing at random or missing completely at random and not necessarily related to smoking status.<sup>27, 39</sup> Alternatively, reasons such as refusal, suspension, and drop out would be considered not missing at random and likely related to smoking status.<sup>39</sup> Simply missing a follow-up assessment may be very different than leaving the program.

Nonetheless, many researchers<sup>1,40,41</sup> recommend using ITT as the primary analysis for teen smoking cessation. The Kohler et al. study followed these recommendations and used ITT as the primary analysis. Notably, if attrition is large, as was the case in the Alabama study, applying ITT analysis can significantly under-power a study.<sup>4,42</sup> The Alabama study reports a loss to follow up higher than reported in previous matched-

design studies of the N-O-T program. Interestingly, a N-O-T study by Mermelstein and colleagues reports a follow-up rate consistent with rates reported by Horn and Dino.<sup>7</sup> Kohler et al. report that the change in composition of the intervention participants was significant; 65% of the intervention group was available at end of program follow up and only 23% was available at 6-month follow up. This low number of participants at 6-month follow up may be one of the reasons that significant differences did not hold over time. Some experts assert that ITT analysis is not useful or appropriate in population or community-level interventions such as N-O-T. McDonald et al.<sup>4</sup> and McDonald<sup>42</sup> posit that compliant analysis is a more reasonable approach as it allows researchers to better understand those participants who are most likely to enroll, participate, and complete the intervention. It is critical to understand the teens who are willing and compliant participants. We can compare the completers' characteristics with the characteristics of those who are lost to follow-up<sup>43</sup>—understanding both types of participants will further the field's knowledge on teen smokers willing to enroll in cessation. The bottom line is that both ITT and Compliant sample analyses yield meaningful information. Researchers may consider reporting both in their studies.

## OTHER ISSUES

The body of published N-O-T research includes 18 articles, to date. The Kohler et al. article highlighted one N-O-T study with a negative finding for females. Kohler et al. presented a second study with positive findings, but criticized the researchers for not reporting ITT analyses. It is important to point out that the N-O-T review paper includes 16 evaluations (both efficacy and effectiveness studies) from five different states, spanning 1998-2003. Using 24-hour point prevalence and ITT analyses, the effectiveness studies found an overall end-of-program quit rate of 26%. The efficacy trials showed a quit rate of 15% and 8% for N-O-T and comparison youth, respectively (OR=1.9, p=.003). The mean number of days of continuous abstinence



nence among N-O-T teens was 21 days.<sup>19</sup> Last, Kohler et al. reported that the N-O-T researchers failed to report longer term follow up (“> 10 weeks”). To clarify, a N-O-T study reporting 15-month post-baseline follow up was published in the *American Journal of Public Health* in 2004; 6-month follow up was published in the *Journal of School Nursing* in 2001.

The Kohler et al. manuscript also urges future study of the cost effectiveness of the N-O-T program. We concur fully that this is part of the ultimate worth of a program. Soon after the release of the Alabama study, Dino, Horn and colleagues published a formal cost effectiveness analysis of N-O-T in *Prevention Science*.<sup>44</sup> This study utilized a Markov transition model of decision analysis to explain stage progression of smoking cessation from the age of 17 years to 25 years among participants who received N-O-T or a 20-minute brief intervention. The primary outcome measure was 7-month post-baseline follow up using an ITT analysis. Study findings predicted that out of a cohort of 100 N-O-T students, 10 will quit smoking and remain smoke-free at the age of 25 years and 14 will reduce smoking, resulting in 102.22 life years saved and a total of 20.11 years discounted life years (DLY) saved. Among brief intervention comparison youth, six will quit smoking and nine will reduce, indicating 64.31 life years saved and a total 12.65 DLY saved. Results indicate that N-O-T is a cost-effective school-based smoking cessation option, as cost effective as school-based primary tobacco prevention, and potentially more cost effective than adult tobacco use cessation.

Most importantly, the contextual factors of the Alabama N-O-T study were not addressed in the Kohler et al. study. If, in fact, the quit rates among Alabama N-O-T participants were lower than national N-O-T averages (which we cannot compare to past N-O-T data), an important question might be “Why were the Alabama quit rates lower than those found in other states?” Kohler and colleagues do not address this question, but we assert that the socio-cultural context of any trial should be considered. Upon

review of state tobacco profiles provided by the Campaign for Tobacco-free Kids,<sup>45</sup> we found that Alabama faces a challenging tobacco culture. Approximately 12,400 Alabama youth under age 18 become new daily smokers each year. The CDC recommends that Alabama allocate between \$26.7 million and \$71.2 million per year for effective and comprehensive tobacco prevention programming. In 2007, Alabama spent about \$680,000—2.6% of the CDC’s minimum recommendation. As such, Alabama ranks 49th among 50 U.S. states for tobacco prevention funding. This fact alone suggests that there are important contextual factors that may influence Alabama teen smokers as well as their quit attempts. Also of note, the tobacco industry markets heavily in Alabama, with annual marketing expenditures estimated at \$265 million. Our research has demonstrated lower quit rates among N-O-T teens in states with this type of tobacco culture, such as West Virginia and North Carolina.<sup>46</sup> Moreover, like West Virginia and North Carolina, Alabama is affected by rural Appalachian culture. Data indicate that rural teens have more difficulty quitting than their non-rural peers.<sup>47</sup> As reported in our previous studies with Appalachian teens,<sup>47,48</sup> cessation programs tailored for rural youth may need to consider topics such as tobacco growing economies, favorable tobacco environments, favorable attitudes about use, geographic isolation and lack of access to services, cultural and traditional values and customs, poverty, and stress and coping. Our past research with rural youth suggests that all of these factors can influence recruitment, engagement, and cessation success.<sup>49</sup> It is essential that both researchers and practitioners develop and test strategies that will be responsive to communities and states, such as Alabama, that face exceptional challenges in their teen smoking cessation missions. Future studies of N-O-T or other teen cessation programs should consider potential socio environmental and political factors as mediators or moderators of cessation outcomes.

In their conclusion, Kohler and colleagues state that “...the hour or so per

week that teachers spent on this program may have been more productively spent on other activities with little loss in terms of new quitters.” This is a strong and, perhaps, unfounded statement which essentially instructs teachers and other practitioners not to use the N-O-T program but to engage in some other unnamed activity. It is important to underscore the fact that some Alabama teens were helped by the N-O-T program, like thousands of other teens across the U.S. This statement minimizes the importance of those Alabama schools, facilitators, and teens that succeeded with N-O-T. For N-O-T or other intervention programs, one negative study should not lead to a recommendation of non-use. To that end, it is critical to understand why the negative findings occurred.

We acknowledge that our past research is not without flaw. We have learned a lot over the years. Few methodological guideposts for teen smoking intervention existed when N-O-T, one of the first teen smoking cessation programs in the U.S., was launched 10 years ago. Per YTCC<sup>1,20</sup> reports, researchers continue to grapple with fundamental issues such as definitions and validity of measurements for youth smoking cessation. Although many research questions remain, N-O-T has provided a vehicle to help shape and understand some of the key constructs in the field (e.g., reduction as an outcome, gender tailored programming, weekend/weekday smoking). No doubt, intervention trials are fraught with challenges. Our purpose is not to negate or dilute Kohler’s findings. Certainly, some of our own studies, especially those in rural states, found lower than average quit rates. We continually seek an understanding of these differences. We appreciate and acknowledge Dr. Kohler’s efforts to apply some of the field’s emerging recommendations—we all gain from new knowledge. In fact, the findings of this study facilitated an opportunity to address several necessary guideposts for future teen smoking cessation research. Before now, there were limited comparative benchmarks for N-O-T investigations or other programs. Refer to Summary Table 1.



In summary, Kohler and colleagues found that 6% of the N-O-T teens quit smoking at the end of program, compared to 2% of the comparison teens. Although this difference was significant, the N-O-T program outcomes reported in Alabama were notably lower than average N-O-T end-of-program rates reported elsewhere in the literature. Significant group differences did not hold at 6- and 12-month follow up. As described above, the low rates of quitting may be attributed to several factors. First, the Alabama rates reflect a 30-day point prevalence observation, using ITT—the most restrictive of all cessation outcomes measurement. Past N-O-T studies reported 24-hour point prevalence, ITT and compliant, with mean number of days for continuous abstinence. Future N-O-T studies should address standard phraseology for quitting, including examination of quit rates under various conditions (24 hour and 7-day point prevalence and 30-day prolonged abstinence at longer-term follow up). Examining these definitions of quitting may lead to greater understanding about the relationship between the short-term processes of stopping on the longer-term state of abstinence.

Second, the attrition was very high at all follow up points, leaving a large number of teens counted as “failures” even though the nature of missingness was unknown. The rate of attrition in the N-O-T group was significantly higher than the comparison group and attrition analyses showed that those available for follow up were significantly different from those who were not. Third, it is possible that N-O-T teens did not receive a sufficient dosage of the program. The authors reported substantial variation in program delivery and session attendance could not be fully tracked. Fourth, it is also possible that the N-O-T program was not effective for Alabama teens as delivered.

Certainly, there is room for improvement for reaching, retaining, and successfully aiding Alabama teen smokers. All of these factors open the door for future research possibilities that may inform N-O-T program improvements and advance our knowledge of teen smoking cessation in general. The impact of N-O-T was not optimal in Alabama; this should not lead to the conclusion that N-O-T should be dismissed as an option for teen smokers in Alabama or elsewhere. We call on the field to work together for programmatic improvements,

to be collegial, and to build on the progress we have made thus far with programs such as N-O-T and Project Ex. We have come too far to start over. It is our intent that the guideposts that emerged from this review will serve future investigations of N-O-T and other teen cessation programs.

## REFERENCES

1. Orleans C, Arkin E, Backinger C, et al. Youth Tobacco Cessation Collaborative and National Blueprint for Action. *Am J Health Behav* 2003;27(Suppl 2):S103-S119.
2. Clayton R, Ries Merikangas K, Abrams D. Introduction to Tobacco, Nicotine, and Youth: The Tobacco Etiology Research Network. *Drug Alcohol Depend* 2000;1(59):Suppl 1:S1-4.
3. Sussman S, Dent C. A meta-analysis of teen cigarette smoking cessation. *Health Psych* 2006;Sep;25(5):549-57.
4. McDonald P, Colwell B, Backinger C, Husten C, Maule C. Better Practices for Youth Tobacco Cessation: Evidence of Review Panel. *Am J Health Behav* 2003; 27 (Suppl 2): S144-S158.
5. Dino G, Horn K, Goldcamp J, Kemp-Rye L, Westrate S, Monaco K. Teen Smoking Cessation: Making it work through school and community partnerships. *J Pub Health Managmt Prac* 2001;7:71-80.
6. Kelder KH, Gortmaker SL, Wiecha JL, Dino GA, Horn KA, Franks AL, Simoes E. School-based programs: lessons learned from CATCH, Planet Health, and Not-On-Tobacco. *Prev Chron Disease* 2007 Apr;4(2):A33. Epub 2007 Mar 15.
7. Mermelstein R. Web-based support as an adjunct to group-based smoking cessation for adolescents. *Nic Tob Res* December 2006, 8 (S1): S69 - S76.
8. Kohler CL, Schoenberger YM, Beasley TM, Phillips MM. Effectiveness evaluation of the N-O-T smoking cessation program for adolescents. *Am J Health Behav* 2008 Jul-Aug;32(4):368-79.
9. American Lung Association. Not on Tobacco (N-O-T): A total approach to helping teens stop smoking. New York: American Lung Association, 1999.
10. Bandura A. Social cognitive theory of self-regulation. *Organztl Beh Human Decision Proc* 1991;50:248-287.

**Table 1. Summary of Guideposts for Future Research**

1. Develop open, collaborative partnerships among investigators conducting research on the same smoking cessation programs.
2. Be consistent in research design features, regardless of the types of design used (e.g., matched, randomized, etc.), when conducting single program evaluations on programs such as N-O-T.
3. Strive toward and measure implementation fidelity when using prescribed program curricula, and report/publish fidelity outcomes.
4. Use consistent comparison group intervention materials as used by other teen cessation researchers and report/publish descriptions of those materials.
5. Use consistent inclusion criteria for and classification of teen smokers and inclusion criteria.
6. Track and document reasons for missing data so that patterns (e.g., NMAR, MAR, MCAR) can be explored.
7. Explore and compare results of alternative methods for attrition analyses.
8. Strive toward standards for quit rate analyses; explore and compare analyses of outcomes under various conditions and cessation criteria.
9. Apply cost effectiveness analyses where data permit.
10. Investigate the influences of socio environmental factors on recruitment, implementation, and cessation outcomes.





11. Prochaska JO, Norcross JC, DiClemente CC. *Changing for Good*. New York, NY: William Morrow, 1994.
12. Dino GA, Horn KA, Goldcamp J, Kemp-Rye L, Westrate S, Monaco K. Teen smoking cessation: making it work through school and community partnerships. *J Public Health Manag Prac* 2001;7(2):71-80.
13. Dino GA, Horn K, Zedosky LK, Monaco K. A Positive Response to Teen Smoking: Why N-O-T? *Sch Health Issues* 1998;46-58.
14. Grimshaw GS, A. Tobacco Cessation Interventions for Young People. *Cochrane Database of Sys Rev* 2006(Issue 4).
15. Green LW. From Research to Best Practices in Other Settings and Populations. *Am J Health Beh* 2001;25(3):165-178.
16. Green L, Glasgow, R. Evaluating the Relevance, Generalization, and Applicability of Research. *Eval Health Profess* 2006; 29(1):126-153.
17. Sussman S, Dent CW, Burton D, Stacy AW, Flay BR. *Developing school-based tobacco use prevention and cessation programs*. Thousand Oaks, CA: Sage, 1995.
18. Sussman S, ed. *Handbook of Program Development for Health Behavior Research & Practice*. Thousand Oaks, CA: Sage Publications, Inc., 2001.
19. Horn K, Dino G, Goldcamp J, Kalsekar I, Mody R. The Impact of Not On Tobacco on Teen Smoking Cessation: End-of-Program Evaluation Results, 1998 to 2003. *J Adol Res* 2005;20(6):640-661.
20. Youth Tobacco Cessation Collaborative. Available at: <http://www.youthtobacco cessation.org>. Accessed April 8, 2007.
21. Sussman S, Dent C, Lichtman KL. Project EX: outcomes of a teen smoking cessation program. *Addict Behs* 2001;May-Jun;26(3):425-38.
22. Dino GA, Horn KA, Goldcamp J, Maniar SD, Fernandes A, Massey CJ. Statewide demonstration of not on tobacco: a gender-sensitive teen smoking cessation program. *J Sch Nurs* 2001;17(2):90-7.
23. O'Donnell C. Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K-12 Curriculum Intervention Research. *Rev Educ Res* 2008;78(1):33-84.
24. Horn K, Dino GA, Gao X, Momani A. Feasibility evaluation of Not On Tobacco: the American Lung Association's new stop smoking programme for adolescents. *Health Educ J* 1999;5:192-206.
25. Dino G, Horn K, Goldcamp J, Fernandes A, Kalsekar I, Massey C. A 2-year efficacy study of Not On Tobacco in Florida: an overview of program successes in changing teen smoking behavior. *Prev Med* 2001;33(6):600-5.
26. Sloboda Z, Bukoski WJ (eds). *Handbook of Drug Abuse Prevention: Theory, Science, and Practice*. New York: Kluwer Academic Publishers 2003.
27. Schafer JL, Grahamn JW. Missing data: Our view of the state of the art. *Psych Methods* 2002(7):147-177.
28. Velicer F, Prochaska JO. A comparison of four self-report smoking cessation outcomes. *Addict Behs* 2004;29:51-60.
29. Copley TT, Lovato CY, Manske S. Identifying Outcome Indicators and Measurement Items for Evaluating Youth Smoking Cessation Interventions: Background Paper. *Canadian Tobacco Control Research Initiative* 2005. Available at: <http://www.ctcri.ca/>. Accessed September 9, 2008.
30. Stanton WR, Smith KN. A critique of evaluated smoking cessation programmes. *J Adolesc* 2002;4:427-438.
31. Ossip-Klein DJ, Bigelow G, Parker SR, Curry S, Hall S, Kirkland S. Classification and assessment of smoking behavior. *Health Psych* 1986;5(Suppl):3-11.
32. Velicer WF, Prochaska JO, Rossi JS, Snow MG. Assessing outcome in smoking cessation studies. *Psychol Bull* 1992;111(1):23-41.
33. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. Applications to addictive behaviors. *Am Psychol* 1992;47(9):1102-14.
34. Prochaska JO, & DiClemente, C. C. Stages and processes of self change of smoking: Toward an integrative model. *J Consult Clin Psych* 1996;51:390-395.
35. Sussman S, Dent CW, Severson H, Burton D, Flay BR. Self-initiated quitting among adolescent smokers. *Prev Med* 1998;27:A19.
36. Curry SJ GL, McBride C. Reasons for quitting: intrinsic and extrinsic motivation for smoking cessation in a population-based sample of smokers. *Addict Behav* 1997;Nov-Dec;22(6):727-39.
37. Turner L, Mermelstein R, Flay B. Individual and Contextual Influences on Adolescent Smoking. *Ann NY Acad Sci* 2004(1021):175-197.
38. Dino G, Horn K, Meit H. A pilot study of Not-On-Tobacco: A smoking program for adolescents. *Health Educ* 1998;6:230-241.
39. Flay BR, Phil D, Hu FB, Richardson J. Psychosocial predictors of different stages of cigarette smoking among high school students. *Prev Med* 1998;27:A9-18.
40. Garrison MM, Christakis DA, Ebel BE, Wiehe SE, Rivara FP. Smoking cessation interventions for adolescents: A systematic review. *Am J Prev Med* 2003; 25(4):363-7.
41. Backinger CL, Fagan P, Matthews E, and Grana, RA, adolescent and young adult tobacco prevention and cessation: Current status and future directions. *Tob Control* 2003;12 (Suppl 4): IV46-53.
42. McDonald C. Intentions to quit smoking in substance-abusing teens exposed to a tobacco program. *J Sub Abuse Treat* 2000;18(3):291-308.
43. Milton MH, Maule CO, Backinger CL, and Gregory DM. Recommendations and guidance for practice in youth tobacco cessation. *Am J Health Behav* 2003;27(Suppl 2):S159-69.
44. Dino G, Horn K, Abdulkadri A, Kalsekar I, Branstetter S. Cost-effectiveness analysis of the not on tobacco program for adolescent smoking cessation. *Prev Sci* 2008 Mar;9(1):38-46. Epub 2008 Feb 20.
45. Campaign for Tobacco-free Kids. Available at: <http://www.tobaccofreekids.org>. Accessed April 8, 2008.
46. Horn KA, Dino GA, Kalsekar ID, Fernandes AW. Appalachian teen smokers: not on tobacco 15 months later. *Am J Public Health* 2004 Feb;94(2):181-4.
47. Horn K, Dino G, Kalsekar I, Massey C. Exploring the relationship between mental health and smoking cessation: A study of rural teens. *Prev Sci* 2004;5:113-126.
48. Horn K, Dino G, Momanani A. Smoking and Stress among rural adolescents: The Gender Factor. *Am J Health Stud* 1998;14:183-193.
49. Massey CJ, Dino GA, Horn KA, Lacey-McCracken A, Goldcamp J, Kalsekar I. Recruitment barriers and successes of the American Lung Association's Not-On-Tobacco Program. *J Sch Health* 2003;73(2):58-63.