2020

Behavior Change Techniques to Promote Smoking Cessation During Pregnancy: A Theory-Based Meta-Analysis

Caroline A. Orr
Virginia Commonwealth University

Behavior Change Techniques to Promote Smoking Cessation During Pregnancy: A Theory-Based Meta-Analysis

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

Caroline Amy Orr Bueno

B.A., Psychology, East Carolina University, May 2006

M.A., Health Education and Promotion, East Carolina University, May 2010

M.S., Counseling Psychology, Loyola University, May 2012

Director: Dr. Jessica LaRose

Virginia Commonwealth University

Richmond, Virginia

November 2020

Acknowledgments

This work would not have been possible without the support and guidance of my family and many esteemed researchers along the way, particularly my dissertation committee. I am especially grateful for the help of Dr. Jessica LaRose, who went above and beyond to make sure I stayed focused and made it across the finish line. I also wish to extend special thanks to Dr. Maghboeba Mosavel, Dr. Robert Perera, and Dr. Susan Bodnar-Deren, all of whom generously gave their time and assistance to me as I completed my dissertation and prepared for my final defense. Finally, I owe many thanks to Kate Grant, who has been invaluable since my earliest days at VCU, as well as a long list of faculty and staff who helped me grow from a student into a scholar.

I am eternally grateful for the ongoing support and encouragement from my family, especially my parents, who never doubted me even during those times when I doubted myself. My mother, Dr. Suezanne Orr, is the person who inspired me to pursue a Ph.D in the first place, and continues to inspire me today. I could not have done this without her.

I also want to thank my husband, Avi, for being a constant source of love and support throughout this process. I will always be grateful to him for helping me reach this point and, less than two weeks later, helping me through the birth of our beautiful baby Evangeline, to whom I wish to dedicate this dissertation.

TABLE OF CONTENTS

Abstract

BEHAVIOR CHANGE TECHNIQUES TO PROMOTE SMOKING CESSATION DURING PREGNANCY: A THEORY-BASED META-ANALYSIS

By Caroline A. Orr, M.A., M.S.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2020.

Major Director: Dr. Jessica LaRose

Despite significant progress, smoking during pregnancy remains one of the leading preventable causes of adverse fetal and maternal health outcomes. Using the current best practice standard of psychosocial counseling, only about one out of every 20 pregnant women quits smoking, and relapse rates are very high. Developing more effective interventions to promote smoking cessation during pregnancy is a critical public health priority that requires a thorough understanding of behavior change and its complex pathways and determinants. As such, the purpose of this three-part study was to conduct the first systematic theory-based evidence synthesis of smoking cessation interventions during pregnancy, and to quantify the effectiveness of specific behavior change techniques and behavioral theories used in these interventions, with the long-term goal of informing the development of more effective interventions to reduce smoking during pregnancy.

The **first aim** was to conduct a meta-analysis to produce quantitative estimates of intervention effect sizes and to identify factors that may explain the observed heterogeneity in intervention effectiveness. A search of six major bibliographic databases for prenatal smoking cessation interventions published between 1995 and 2015 yielded 1,223 unique articles, of which 38 met criteria for inclusion and 34 were randomized controlled trials where the primary outcome was late-pregnancy biochemically-validated smoking cessation and the unit of randomization was the individual. The results of a random effects meta-analysis of the 34 randomized controlled trials of prenatal smoking cessation interventions yielded a significant risk ratio for the primary outcome of late-pregnancy smoking cessation, such that women in the treatment groups were 1.53 times as likely to achieve smoking cessation before giving birth than women in the respective control groups (RR = 1.53; 95% CI: 1.30-1.79). Several study-level variables emerged as potential moderators of intervention effectiveness. Treatment-group participants in contingent rewards interventions were 2.82 times as likely to achieve late-pregnancy smoking abstinence than control group participants. In comparison, treatment-group participants in counseling interventions were 1.3 times as likely to achieve late pregnancy smoking abstinence than their control group counterparts. Intensity level was not associated with effectiveness in this sample. Interventions in this review also yielded promising (significant) results for many secondary outcomes of interest, including additional measures of smoking behavior as well as perinatal outcomes. Specifically, treatment group participants were 1.44 times as likely as control group participants to significantly reduce (by at least 50%) their cigarette consumption, 1.54 times as likely to be smoke

free in the early postpartum period, and 1.99 times as likely to be smoke free in the late postpartum period. The results also revealed that smoking cessation interventions reduced the risk of two very common adverse perinatal health outcomes: low birthweight and preterm birth. Specifically, treatment group participants had 73% less risk of delivering a low birthweight or very low birthweight infant and 67% less risk of preterm birth compared to control group participants.

More than two years after the initial completion of the meta-analysis, a subsequent search of the literature for studies published between 2015 and 2020 returned six additional trials that would have met the criteria for inclusion in the original study. Of those, four tested telephone- or text-message-based interventions to encourage quitting among pregnant women, and two used incentives or rewards to promote cessation. Findings across these trials were mixed. There was no clear pattern delineating the studies with significant results from the non-significant results with the exception that incentive-based interventions were more consistently effective than other types of interventions, which is in line with the results of the meta-analysis presented in this dissertation.

The **second aim** was to evaluate the use of the health behavior theory in intervention design, implementation, and evaluation, and to assess whether the use of theory was associated with intervention effectiveness. Of the 26 published trials that explicitly mentioned theory in the introduction or methods, only nine were based on a single theoretical framework. Five of these studies utilized the learning-based theory of operant conditioning, two studies utilized the transtheoretical/stages of change model, one study used social cognitive theory, and one study used social learning theory. Even

among these nine trials, theory was used primarily in a descriptive manner, as opposed to an explanatory or predictive manner. The results of the subgroup analyses and meta-regression models were counter to the hypothesis that use of theory would be positively associated with intervention effectiveness. Scores on two categories of the theory coding scheme ("Was theory tested?" and "Was theory used to tailor or select participants?") were significantly associated with the primary outcome of late-pregnancy smoking abstinence, but both of the associations were negative, indicating that greater use of theory was associated with a lower likelihood of smoking abstinence during the late-pregnancy period. However, this may reflect the limited use of theory in intervention planning and design among trials included in this meta-analysis, rather than the contribution of theory when it is used optimally.

The **third aim** was to isolate the "active" ingredients in prenatal smoking cessation programs by applying a standardized taxonomy of behavior change techniques to identify the techniques, and then quantifying the effectiveness of each individual technique. We first used Abraham and Michie's (2008) 26-item taxonomy to identify theory-derived behavior change techniques in published descriptions of intervention content, then performed a meta-regression analysis to determine whether interventions utilizing more techniques were more likely to be effective, and then used subgroup and moderator analyses in order to quantify the effectiveness of each technique. The results revealed that the total number of behavior change techniques used was not associated with late pregnancy smoking abstinence, indicating that more is not necessarily better. Effect sizes were significantly larger for the treatment group than the control group for subsets of interventions that 1) provided information about the

link between smoking and health (RR = 1.68; 95% CI: 1.26-2.12); 2) provided information about the negative consequences of smoking (RR = 1.38; 95% CI: 1.08-1.77); 3) prompted the formation of intentions to quit smoking (RR = 1.24; 95% CI: 1.00-1.53); 4) provided instructions (RR = 1.51; 95% CI: 1.21-1.89); 5) prompted specific goal setting (RR = 1.48; 95% CI: 1.17-1.88); 6) provided contingent rewards (RR = 2.82; 95% CI: 2.05-3.88); 7) taught participants to use prompts and/or cues (RR = 1.63; 95% CI: 1.03-2.59); and/or 8) had participants agree to a behavioral contract (RR = 2.14; 95% CI: 1.29-3.56).

The results of the review are subject to a number of limitations, particularly stemming from reporting and measurement practices, but several key findings and patterns still emerged. First, behavior change theory is not being utilized to its full capacity in the development and evaluation of prenatal smoking cessation interventions, with only half of the studies in this review (n = 19) reporting an explicit link between at least one behavior change technique and at least one targeted predictor of behavior change. Secondly, many of the most common behavior change techniques used in prenatal smoking cessation interventions were not associated with better intervention outcomes, nor was the quantity of techniques used associated with effectiveness. Third, the current review identified contingent rewards as the most effective behavior change technique for promoting smoking cessation during pregnancy and into the postpartum period when tangible rewards were no longer offered.

While previous meta-analyses have assessed whether or not prenatal smoking cessation interventions were effective, this review expanded on existing findings by using a recently developed taxonomy to identify, isolate, and quantify the effectiveness

of individual behavior change techniques used in interventions, as well as applying a coding scheme to evaluate how theory is being used in the literature and whether the use of theory is associated with the effectiveness of interventions. The results provide a framework for evaluating not only *if* an intervention worked, but also why, how, and under what conditions, marking an important step towards a new set of standards in evidence synthesis and theory-testing in smoking cessation research and beyond.

Vita

Caroline Amy Orr Bueno was born on October 14, 1985, in Lutherville, Maryland and is an American citizen. She graduated from Junius H. Rose High School in Greenville, North Carolina in 2003. She received her Bachelor of Arts in Psychology in 2006 from East Carolina University in Greenville, North Carolina, where she graduated Summa Cum Laude as a member of the Degree in Three program. She subsequently received a Master of Arts in Health Education and Promotion from East Carolina University in 2010, and a Master of Science in Counseling Psychology from Loyola University in Baltimore, Maryland, in 2012.

CHAPTER 1

Despite decades of progress, smoking remains one of the leading preventable causes of poor maternal and fetal/infant outcomes, including preterm birth, low birthweight, and infant mortality (Centers for Disease Control and Prevention [CDC], 2012; Dietz et al., 2010; Hammoud et al., 2005; Rogers, 2009; Salihu, Aliyu, Pierre-Louis, & Alexander, 2003; Vardavas et al., 2010). Reviews of the associated population burden indicate that smoking during pregnancy may account for up to 15% of all miscarriages, 20-30% of all low birthweight deliveries, and may increase overall perinatal mortality by as much as 150% (Andres & Day, 2000). Other adverse health outcomes associated with smoking during pregnancy include an increased risk of birth defects (McDonald, Perkins, Jodouin, & Walker, 2002), fetal growth retardation (Vardavas et al., 2010; US Department of Health and Human Services [HHS], 2001; HHS 2004), placental abruption (HHS, 2001; HHS 2004; Kyrklund-Blomberg, Gennser, & Cnattingius, 2001), sudden infant death syndrome (SIDS) (Anderson & Cook, 1997; CDC, 2013; DiFranza & Lew, 1995), and impaired fetal lung development leading to reduced pulmonary functioning (Upton, Watt, Davey-Smith, McConnachie, & Hart, 1998; Young et al., 2000). Prenatal exposure to smoking can also set the stage for serious long-term health and developmental problems, including psychiatric morbidity and mortality (Ekblad, Gissler, Lehtonen, & Korkeila, 2010), behavioral disorders (Ernst, Moolchan, & Robinson, 2001; Higgins, 2002), and obesity throughout the lifespan (Harris, Willet, & Michels, 2013; Toschke, Koletzko, Slikker, Hermann, & von Kries, 2002; von Kries, Toschke, Koletzko, & Slikker, 2002). Women who continue to smoke during pregnancy are also less likely to attend critical prenatal screenings and more likely to start prenatal

care later in pregnancy, which further compounds the risks associated with smoking (Schneider et al., 2010; Tong et al., 2008).

Reducing the number of women who smoke during pregnancy has the potential to avert many of these negative health outcomes, and could also yield substantial economic savings. Even among women who are still smoking at their first prenatal care visit, those who quit smoking during their pregnancy have better birth outcomes than those who continue to smoke (HHS, 2004). Lightwood, Phibbs, and Glantz (1999) estimate that an annual reduction of smoking prevalence of 1% among pregnant women could prevent 1,300 low birthweight live deliveries and save $21 million in direct medical costs in just the first year. Within seven years, an annual 1% drop in the prevalence of smoking during pregnancy could prevent more than 57,000 low birthweight live deliveries and save over $572 million in direct medical costs (Lightwood et al., 1999). Importantly, research also indicates that the costs of implementing a smoking cessation intervention for pregnant women ($24-$34 per person) are more than made up for by the estimated costs saved ($881) for each woman who quits smoking during pregnancy (Ayadi et al., 2006).

Given the significant short- and long-term health consequences of maternal smoking, and the associated economic burden, reducing the prevalence of smoking among pregnant women in the U.S. is an important public health priority. Highlighting the significance of the problem, Healthy People 2020 devoted three national health objectives to address smoking during pregnancy: 1) reduce the prevalence of women smoking prior to pregnancy to 14% (objective no. MICH-16.3); 2) reduce the prevalence of cigarette smoking among pregnant women to 1% (objective no. MICH-11.3); and 3)

increase the percentage of pregnant smokers who stop smoking during pregnancy to 30% (objective no. TU-6) (HHS, n.d.).

Common approaches to promoting smoking cessation among pregnant women include the provision of psychosocial counseling, peer- and/or partner-support, health education, rewards and incentives, feedback, and pharmacological support (Chamberlain et al., 2013; Lumley et al., 2009). However, using the current best practice standard of psychosocial counseling, only about one out of every 20 pregnant women quits smoking, and relapse rates are very high (Lumley et al., 2009). Furthermore, two recent meta-analytic reviews both concluded that even when positive outcomes are achieved, significant heterogeneity is still present in the data (Chamberlain et al., 2013; Lumley et al., 2009). As a result, there is a lack of clarity about which intervention techniques are responsible for promoting behavior change, and whether technique effectiveness depends on other factors such as participant characteristics, delivery procedures, and/or context.

Developing more effective interventions to reduce smoking during pregnancy requires a thorough understanding of behavior change and its complex determinants. While effectively changing behavior is challenging, evidence strongly suggests that the use of health behavior theory to inform intervention design, research, and evaluation is associated with increased effectiveness (Abraham, Kelly, West, & Michie, 2009; Albarracin, Gillete, Earl, Glasman, Duranti, & Ho, 2005; Noar & Zimmerman, 2005). Behavior change theories provide explicit frameworks for specifying, categorizing, and evaluating interventions, as well as for identifying and understanding the mechanisms through which the observed effects are achieved. However, current reviews and

meta-analyses of smoking cessation interventions for pregnant women rarely use theory to classify intervention components or specify behavioral determinants, which limits our ability to understand the behavior change processes that underlie effective interventions and intervention components, and to use this knowledge to inform and improve upon the design of future interventions (Likis, Andrews, Fonnesbeck, et al., 2014; Michie & Prestwich, 2010). As such, the purpose of this project is to produce the first theory-based quantitative evidence synthesis of behavior change techniques for prenatal smoking cessation, with the ultimate goal of informing the development of more effective interventions to reduce smoking during pregnancy.

CHAPTER 2

Literature Review

Epidemiology of Maternal Smoking Behaviors

Pregnancy is often described as a 'window of opportunity' for addressing health problems and promoting healthy behavior changes including smoking cessation (McBride, 2003). Research indicates that pregnancy increases women's perceptions of risk and susceptibility to health problems, which may increase motivation to quit smoking (Ortendahl & Nasman, 2008; Slade, Laxton-Kane, & Spiby, 2006). Additionally, more women quit smoking during pregnancy than at any other point during their lives, with up to half of women who smoked before pregnancy spontaneously quitting before their first prenatal care first (Woodby, Windsor, Snyder, Kohler, & DiClemente, 1999). The prenatal period is also one of the few times in a woman's life when she has regular, sustained contact with the healthcare system, which in turn gives providers a unique opportunity to offer help with smoking cessation.

In 2008, nearly 13% of pregnant women in the U.S. smoked during the last three months of pregnancy and 17% smoked in the immediate postpartum period (CDC, 2013b). Although nearly half of smokers quit when they decide to become pregnant or upon learning that they are pregnant, only an additional 5%-12% of pregnant smokers quit by the last three months of pregnancy (Tong, 2008; Tong et al., 2013), and an estimated 84% of pre-pregnancy smokers report daily smoking later in pregnancy (Pickett, Rathouz, Kasza, Wakschlag, & Wright, 2005). Furthermore, up to one-third of the women who spontaneously quit early in pregnancy will relapse before the end of

pregnancy (Coleman-Cowger, 2012; Floyd, Rimer, Giovino, Mullen, & Sullivan, 1993), and 70-90% will relapse during the postpartum period (Chamberlain et al., 2013; DiClemente, Dolan-Mullen, & Windsor, 2000; Fang, 2004; McBride et al., 1999).

Evidence suggests that the psychological, behavioral, and biological processes involved in smoking cessation may be different during pregnancy than other life stages. Compared to non-pregnant smokers, women who successfully quit smoking during pregnancy have higher levels of confidence in their ability to abstain from smoking but are less likely to use behavioral strategies that are most common among non-smoking populations, such as reliance on coping skills and distractions (Ruggiero, Tsoh, Everett, Fava, & Guise, 2000). Furthermore, pregnant women who quit smoking display unique patterns of vulnerability to relapse. While ex-smokers are generally most likely to relapse shortly after quitting, when symptoms of nicotine withdrawal are most severe (Killen & Fortmann, 1994; Pomerleau et al., 1986), pregnant women who quit smoking tend to maintain abstinence for the duration of pregnancy but then relapse in the early postpartum period, after symptoms of nicotine withdrawal have largely disappeared (Buja et al., 2011; Stotts, DiClemente, Carbonari, & Mullen, 1996). Many women who quit on their own upon learning of pregnancy may simply enter a period of “suspended smoking” rather than sustained abstinence, as reflected in extremely high postpartum relapse rates (DiClemente, Dolan-Mullen, & Windsor, 2000; McBride et al., 1999). These findings indicate that smoking cessation during pregnancy may often be motivated by a desire to protect the health of the baby, rather than a long-term commitment to quitting (Stotts et al., 1996). Importantly, this also suggests that the

determinants of smoking cessation among pregnant women may be different than in the general (non-pregnant) population.

Demographic Factors

In the U.S. and other high-income countries, smoking is more common among low socioeconomic women and is one of the leading contributors to health disparities (Wanless, 2004). This is reflected in the data on pregnant women, as the prevalence of smoking is markedly higher among lower-SES women. In 2005, just 1.8% of women with a college degree reported smoking during pregnancy, compared to 20.2% of women with less than a high school education (Martin et al., 2007). Compared to non-smokers, women who smoke during pregnancy are significantly more likely to have completed less than 12 years of education, have an annual income of less than $15,000, and be enrolled in Medicaid coverage during pregnancy or at the time of delivery (Tong et al., 2009; Tong et al., 2013). The prevalence of smoking during pregnancy is typically higher in younger age-groups, including adolescents and young women aged 18 to 24 years-old (Martin et al., 2007). In 2005, women aged 18 to 19 years had the highest prevalence of smoking during pregnancy (18.9%), followed by those aged 20 to 24 years (18.6). In comparison, only 11.5% of women aged 25 to 29 years and 7.1% of women aged 30 to 29 years smoked during pregnancy (Martin et al., 2007). Additionally, non-Hispanic white women have a significantly higher prevalence of smoking during pregnancy than black or Hispanic women (Colman & Joyce, 2003; Martin et al., 2007; Tong et al., 2013).

Other socio-demographic factors associated with an increased likelihood of continued smoking during pregnancy include being unemployed, being unmarried,

having an unintended pregnancy, having a criminal history, and living with a smoker (Colman & Joyce, 2003; Kahn, Certain, & Whitaker, 2009; Kaneko et al., 2008; Lu, Tong, & Oldenburg, 2001; Martin, McNamara, Bloch, Hair, & Halle, 2008; Masho, Bishop, Keyser-Marcus, Varner, White, & Svikis, 2013; Tong et al., 2009; Tong et al., 2013). Additionally, multiparous women are more likely than first-time mothers to smoke during pregnancy, possibly because their risk perceptions were skewed by having a previous healthy pregnancy despite concurrent smoking (Schneider et al., 2010).

Predictors of successful smoking cessation during pregnancy include being married, having at least a high school education, being less dependent on nicotine, smoking fewer cigarettes/day prior to pregnancy, starting to smoke at a later age, and having a non-smoking partner (Colman & Joyce, 2003; Ebert & Fahy, 2007; Ingall & Cropley, 2010; Kahn, Certain, & Whitaker, 2009; McBride et al., 1998). Research also indicates that insurance coverage may be an important predictor of smoking cessation among pregnant women. Women with more generous insurance coverage (including coverage for cessation counseling with no copayment and pharmacotherapy with affordable copayment) are up to twice as likely to quit smoking during pregnancy than women with pharmacotherapy-only coverage or no coverage at all (Greene, Sacks, & McMenamin, 2014; Petersen, Garrett, Melvin, & Hartmann, 2006). This is particularly important given that the prevalence of smoking during pregnancy is significantly higher among Medicaid recipients, with estimates indicating that as many as one in four pregnant Medicaid recipients are smokers (CDC, 2000; Martin et al., 2002).

Early enrollment in WIC, the nutrition assistance program for women, infants, and children, is also associated with higher quit rates and, for black women, reduced relapse

rates during pregnancy (Yunzai-Butler, Joyce, & Racne, 2010). Studies also suggest that smokers who enroll in WIC during their first trimester of pregnancy are significantly more likely to reduce smoking than women who enrolled in their third trimester (Brodsky & Viner-Brown, 2006).

Psychosocial, Behavioral, Social and Environmental Factors

In addition to socio-demographic factors, maternal smoking behaviors are also influenced by a variety of complex, often interacting psychosocial, behavioral, and environmental factors (Ahluwalia, Merritt, Beck, & Rogers, 2001; Schneider, Huy, Schuetz, Diehl, 2010). Stress, depression, self-efficacy, perceived control, and social support have been identified as particularly important factors associated with smoking behavior during pregnancy (Blalock, Fouladi, Wetter, & Cinciripini, 2005; Fernander, Moorman, & Azouru, 2010; Holtrop et al., 2010; Ingall & Cropley, 2010; Orr, Blazer, & Orr, 2012). Likewise, environmental factors, such as greater exposure to environmental smoke and living with a smoker have also been linked with a higher prevalence of maternal smoking during pregnancy (Homish, Eiden, Leonard, & Kozlowski, 2012).

In addition to their strong associations with continued smoking during pregnancy, these factors can interact with each other, producing a synergistic effect that may further reduce the likelihood of successful smoking cessation and increase the risk of poor pregnancy outcomes (Ahluwalia et al., 2001; Maxson, Edwards, Ingram, & Miranda, 2012). For example, women who smoke during pregnancy are more likely to report multiple risk factors for both unsuccessful quit attempts and poor birth outcomes, such as illicit drug use, high levels of stress and stressful life events, intimate partner violence, and unplanned pregnancy (Ahluwalia et al., 2001).

Factors associated with quitting. Evidence suggest that the vast majority of pregnant women are aware that smoking during pregnancy poses significant risks to their own health and the health of their baby (Coonrod, Bruce, Malcolm, Drachman, & Frey, 2009; Frey & Files, 2006; Ingall & Cropley, 2010; Orr, Newton, Tarwater, & Weismiller, 2005). However, even women who know about the risks of smoking and have a desire to stop are often unsuccessful (Ingall & Cropley, 2010). Barriers to attending cessation programs include low self-efficacy and fear of failure, concerns about being judged by healthcare providers, low confidence in the effectiveness of cessation programs, reluctance to ask for help, and low or adversarial social support (Ingall & Cropley, 2010; Owens & Penn, 1999; Ussher, Etter, & West, 2006). Logistical constraints (e.g., lack of childcare, work commitments) are also frequently cited as barriers to attending smoking cessation programs (Owens & Penn, 1999; Ussher, Etter, & West, 2006). Finally, concerns about weight gain are another potential barrier to smoking cessation. Many pregnant smokers report that they use smoking as a weight management strategy during pregnancy (Abraham et al., 1994; Pomerleau, Namenek Brouwer, & Jones, 2000) and women who are more concerned about post-cessation weight gain tend to smoke more cigarettes/day, are less likely to make a quit attempt, and more likely to relapse if they do attempt to quit (Berg, Park, Chang, & Rigotti, 2007). Some of these barriers can be counterbalanced if women perceive they may benefit from attending smoking cessation intervention. The most frequently cited benefits to attending smoking cessation programs that women report include being able to deal with cravings more effectively, having someone to discuss their concerns with, and having increased structure and accountability (Ussher, Etter, & West, 2006).

The degree of tobacco/nicotine addiction is another important predictor of smoking cessation, with higher degrees of addiction strongly associated with a reduced likelihood of successful cessation attempts (Schneider et al., 2010). Other smoking-related variables, including age at smoking initiation and current level of nicotine addiction, are also strongly associated with smoking during pregnancy (Curry et al., 2001; Ockene et al., 2002; Schneider et al., 2010; Solomon & Quinn, 2004), such that women who started smoking at a younger age and who are more addicted to nicotine are more likely to continue smoking during pregnancy (Curry et al., 2001; Ockene et al., 2002; Schneider et al., 2010; Solomon & Quinn, 2004).

Social factors. A variety of social influences also impact smoking behavior among pregnant women. For example, limited or negative social support (e.g., interpersonal conflict, pressure not to quit smoking) (Ebert & Fahy, 2007; Ingall & Cropley, 2010; Moiduddin & Massey, 2008; Pickett, Wilkinson, & Wakschlag, 2009; Schneider & Schutz, 2008), unstable living situations (Ingall & Cropley, 2010), neighborhood disadvantage (Elsenbruch et al., 2007), and intimate partner violence and other forms of victimization (Bacchus, Mezey, & Bewley, 2004; Coker, Sanderson, & Dong, 2004; Cheng, Salimi, Terplan, & Chisolm, 2015; Goedhart et al., 2009) are all associated with an increased likelihood of continued smoking during pregnancy. Across the literature, studies consistently identify partner smoking behaviors as a particularly salient influence on women's smoking behaviors during pregnancy (McBride, Curry, Grothaus, Nelson, Lando, & Pirie, 1998; Pollak & Mullen, 1997; Pollak et al., 2001, 2006). Having a partner who smokes can reinforce and strengthen tobacco/nicotine addiction and reduce motivation and attempts to quit, but a partner's efforts to quit

smoking can also be a strong motivator to quit among pregnant women (Koshy, Mackenzi, Tappin, & Bauld, 2010; McBride et al., 1998; Pollak & Mullen, 1997; Pollak et al., 2001, 2006). Similarly, smoking cessation may be reinforced by supportive social networks, or inhibited by the presence of other smokers (Aaronson, 1989).

Lack of social support has been identified as one of the primary reasons for low attendance of smoking cessation interventions among pregnant populations (Klerman, Spivey, & Raykovitch, 2009). Furthermore, evidence suggests that at least some of the positive health outcomes of smoking cessation interventions for pregnant women may be due to the supportive mechanisms by which the intervention components are delivered (Oakley, 1985).

Policy. At the population-level, tobacco policies may also influence maternal smoking and smoking cessation, and in turn, improve birth outcomes (Hawkins & Baum, 2014; Hawkins, Baum, Oken, & Gillman, 2014). Studies indicate that increasing state cigarette taxes significantly reduces the prevalence of maternal smoking, particularly among low-SES women, and also reduces the risk of having low-birth-weight, preterm, and small-for-gestational-age babies (Hawkins & Baum, 2014; Hawkins et al., 2014). Further, environmental cues to smoke (e.g., living and/or working around other smokers) are associated with a greater likelihood of continued smoking during pregnancy (Lu, Tong, & Oldenberg, 2001).

Systems and interacting factors. As described above, the relationships between these factors are dynamic and interactive, creating a complex web of risk factors. The interaction of socioeconomic status with other demographic and psychosocial risk factors is of particular importance, as socioeconomic status is such a

strong determinant of maternal smoking behaviors. For example, lower-SES women are twice as likely as higher-SES women to have a partner who smokes (Schneider et al., 2010). Lower-SES women also tend to report higher levels of stress and lower levels of social support, which are all associated with an increased risk of continuing to smoke during pregnancy (Crittenden, Manfredi, Cho, & Colecek, 2007; Ebert & Fahy, 2007; Holtrop et al., 2010; Pickett, Wilkinson, & Wakschlag, 2009; Schneider & Schutz, 2008). Social disadvantage may also contribute to smoking behaviors, such that disadvantaged people are more likely to smoke as a coping mechanism to deal with their life circumstances (World Health Organization [WHO], 2008).

Additionally, smoking during pregnancy often clusters with other detrimental health behaviors, including inadequate prenatal care (Moore, Blatt, Chen, Van Hook, & DeFranco, 2016), alcohol use (Masho, Bishop, Keyser-Marcus, Varner, White, & Svikis, 2013), illicit drug use (Masho et al., 2013), poor dietary habits and failure to adhere to guidelines for dietary supplements (such as folate) (Moore et al., 2016). As such, women who smoke during pregnancy often face a complex web of risks that increase the likelihood of adverse pregnancy outcomes and reduce the likelihood of achieving and maintaining smoke-free status.

Gender and stigma. Smoking-related attitudes, beliefs, and behaviors are further complicated by their interaction with gender, which is increasingly recognized as a critical factor in the study of prenatal smoking cessation (Bottorff et al., 2012, 2014; Burgess, Fu, & van Ryn, 2009). As Heaton (2009) and Greaves and colleagues (Greaves, Kalaw, & Bottorff, 2007; Greaves & Tungohan, 2007) note, issues of gender and power play a key role in smoking and smoking cessation, yet they are rarely

considered in the design and evaluation of smoking cessation interventions. For example, evidence suggests that approaches that play heavily on women's roles as mothers and nurturers, and messages that emphasize the need to "protect" women from their own behavior or that emphasize the needs and rights of the unborn fetus to the point of minimizing the needs and rights of the woman could disempower and marginalize women (Greaves & Tungohan, 2007). This is particularly relevant for prenatal smoking cessation programs, as smoking during pregnancy disproportionately affects socially disadvantaged women who are already marginalized. Failure to consider the intersection of gender and smoking behavior when designing and evaluating prenatal smoking cessation interventions could result in harmful unintended consequences, including stigmatization of already marginalized women, adverse psychological outcomes (resulting from feelings of guilt, shame, and loss of control), and delay in or avoidance of seeking healthcare (Burgess et al., 2009).

Given the complex, dynamic factors underlying prenatal smoking behaviors, effective smoking cessation interventions must consider these multiple risk factors, including those that are modifiable and non-modifiable. Additionally, the design of smoking cessation programs must weigh the potential risks against expected benefits while also taking into account practical constraints (e.g., staffing, budgets) and cost/benefit ratios. Furthermore, it may be necessary to vary intervention techniques and intensity to match the needs of pregnant women, different stages of quitting, and different stages of pregnancy. It is also necessary to define the appropriate target population(s), which in some cases may involve intervention providers, family members, social support systems, and/or others.

Smoking Cessation Interventions

The U.S. Public Health Service (USPHS) Clinical Practice Guidelines call for all pregnant smokers to be offered psychosocial interventions, such as cognitive-behavioral therapy, in addition to self-help materials for smoking cessation (DHHS, 2008; Tobacco Use and Dependence Guideline Panel, 2008). These recommendations were further affirmed by the American College of Obstetricians and Gynecologists (ACOG), which released an updated committee opinion in 2010 that closely aligned with the USPHS guidelines (ACOG, 2010). The current best practice for prenatal smoking cessation involves psychosocial counseling delivered in the prenatal care setting (DHHS, 2008), which is recommended as a first-line approach before any pharmacological treatments are considered. Both the ACOG and the USPHS are based on empirical evidence indicating that brief (5-15 minutes) counseling interventions using the "5 A's" (ask, advise, assess, assist, arrange), in combination with pregnancy-specific educational materials, can increase quit rates by 30% to 70% among pregnant smokers (Melvin, Dolan-Mullen, Windsor, Whiteside, & Goldenberg, 2000). For women who need additional assistance, the guidelines call for referral to specialty services such as telephone quitlines or tobacco dependence treatment specialists.

In its most recent recommendation statement, the U.S. Preventive Services Task Force concluded that the existing evidence does not allow for a sufficient assessment of the balance of benefits and harms of nicotine replacement products or other pharmacological treatments for smoking cessation aids during pregnancy (Siu, 2015). Therefore, it is recommended that nicotine replacement therapy should be used only

under close supervision and after careful consultation about the (known) risks of continued smoking and the (potential) risks of nicotine replacement therapy. Furthermore, if nicotine replacement is used, it should only be used with patients who have made a clear commitment to quit smoking.

Coverage for pregnancy-specific smoking cessation services increased greatly in the early to mid-1990's due to changes in the public and private insurance market (Ibraham, Schauffler, Barker, & Orleans, 2002); however, coverage expansions occurred at the state level and often did not reflect clinical practice guidelines. In the late 1990's, Medicaid programs in over 30 states covered medication-assisted smoking cessation services for pregnant women, while just 20 covered non-medication-based smoking cessation counseling (Schauffler, Mordavsky, Barker, & Orleans, 2001). Even where these services are covered, failure to refer pregnant women to smoking cessation counseling has been identified as a significant barrier (Thorndike, Rigotti, Stafford, & Singer, 1998). Over 95% of health care providers report routinely asking pregnant patients about their smoking habits, which is a higher rate of inquiry than among any other patient population (Thorndike, Rigotti, Stafford, & Singer, 1998). However, referrals to smoking cessation counseling and follow-up services are no higher among pregnant women than among patient groups, indicating a wide gap between assessment of smoking status and implementation of clinical guidelines for pregnant smokers (Thorndike et al., 1998). One potential explanation for this gap is a lack of funding for research on effective dissemination of evidence-based smoking cessation interventions during the prenatal period (Orleans, Barker, Kaufman, & Marx, 2000).

Currently, reviews of smoking cessation interventions for pregnant women reveal modest success (Chamberlain et al., 2013; Lumley, Oliver, Chamberlain, & Oakley, 2004; Lumley et al., 2009). Using the current best practice standard of psychosocial counseling, only about one out of every 20 pregnant women quits smoking during pregnancy (Lumley et al., 2009). Furthermore, there is significant heterogeneity in the effectiveness of prenatal smoking cessation programs, and, when positive outcomes are observed, it remains unclear which intervention techniques or combination of techniques are responsible for the change (Chamberlain et al., 2013; Lumley et al., 2009). There is also a problematic gap between theory, research, and practice. For example, despite the strong association and theorized pathway between social support and maternal smoking behaviors, interventions designed to enhance social support yield mixed outcomes overall, and are no more effective than other types of smoking cessation interventions for pregnant women (Chamberlain et al., 2013; Lumley et al., 2009). In fact, there is insufficient evidence at this time to recommend one approach over the others (Lumley et al., 2009).

Specific Aims

The overarching aim of this project is to produce the first theory-based quantitative evidence synthesis of behavior change techniques *as reported in published trials* of prenatal smoking cessation interventions, with the ultimate goal of informing the development of more effective interventions to reduce smoking during pregnancy. Drawing upon recent advances in theory-building and program evaluation, this study will add to the literature by using standardized, theory-based definitions of behavior change techniques to identify intervention components and quantify their unique contributions to the effectiveness of prenatal smoking cessation interventions. In doing so, the results of this study will help facilitate effective prenatal smoking cessation interventions and allow for the accumulation of evidence on key outcomes, such as which techniques are most effective and which factors may moderate their effectiveness. The results also have the potential to contribute to the refinement of health behavior theories. Given that these theories form the conceptual basis of smoking cessation interventions, improving the theory itself could lead to more effective intervention designs and better inform practice (Noar & Mehrota, 2011).

To achieve the overarching aim, this project is broken down into three primary aims, starting with a meta-analysis as the foundation off of which the next two steps build:

Aim 1: To conduct a meta-analysis to determine the effect size of smoking cessation interventions on the primary outcome of smoking cessation during pregnancy, and on

the secondary outcomes of a) postpartum smoking abstinence, b) smoking reduction, and c) maternal and fetal health outcomes.

1) **Sub-aim 1a:** To quantify heterogeneity (inconsistency) of effect sizes.
2) **Sub-aim 1b**: To identify sources of heterogeneity in effect size estimates using subgroup analyses (for categorical variables) and univariate meta-regression models (for continuous variables) examining how intervention effectiveness differs according to characteristics of the intervention, study design, and participants.

Aim 2: To evaluate the use of behavior change theory in prenatal smoking cessation interventions based on the results of the meta-analysis in Aim 1.

1) **Sub-aim 2a:** To assess the use of theory as a guiding framework in prenatal smoking cessation interventions, using Michie & Prestwich's (2010) coding scheme for evaluating the extent to which an intervention is theory-based.
2) **Sub-aim 2b:** To determine whether the use of theory explains variation in intervention effects by conducting subgroup analyses on categorical theory variables and using univariate meta-regression models for continuous theory-related variables.

Aim 3: To identify the potential "active ingredients" in prenatal smoking cessation interventions.

1) **Sub-aim 3a:** To identify standardized, theory-linked behavior change techniques used in published randomized controlled trials of prenatal smoking cessation interventions, using a coding process described by Michie and colleagues' (Abraham & Michie, 2008; Michie et al., 2008; Michie et al., 2009a).

2) **Sub-aim 3b:** To evaluate the effectiveness of each technique using subgroup analyses to calculate the effect size of interventions that used the technique compared to those that didn't use the technique, and to determine whether the total number of active BCTs used in an intervention is associated with effectiveness using a univariate meta-regression model.
3) **Sub-aim 3c:** To explore whether the effect size estimates of BCTs identified as effective in sub-aim 3b differ according to characteristics of the study design, intervention, or participants.

CHAPTER 3

Specific Aim 1

Smoking cessation interventions encompass a wide variety of approaches and techniques that seek to address the problem at different levels of intervention and intensity. These include population-level interventions such as smoking taxes and mass media campaigns, organizational-level interventions such as workplace and healthcare system policies, interpersonal-level interventions such as partner- and family-based support programs, and individual-level interventions such as telephone counseling, hypnotherapy, motivational interviewing, contingency management, incentives, health education, and pharmacotherapy.

In a recent meta-analysis of prenatal smoking cessation trials, Lumley and colleagues (2014) found that the majority of interventions were multimodal, or consisted of more than one intervention strategy. The most common intervention strategies used to promote prenatal smoking cessation are individual-based techniques including the provision of advice and counseling, motivational interviewing, tailored counseling based on the stages of change, feedback, incentives, social support, and pharmacological therapy (Lumley et al, 2014).

Types of Interventions

Incentives/rewards-based interventions, which involve the provision of material rewards to precipitate or reinforce behavior, are a promising approach to smoking cessation in the general population (Cahill, Hartmann-Boyce, & Perera, 2015). However, while incentive-based interventions have shown more promise than other

behavioral interventions, the evidence on their use among pregnant populations, specifically, is mixed, and methodological problems limit the quality of evidence produced by many evaluations of incentive-based interventions (Higgins, et al., 2012). A review of six controlled trials found that financial incentives were associated with higher levels of smoking cessation during and after pregnancy among low-income women, though not among the wider population of pregnant smokers (Higgins, et al., 2012). Additionally, evidence suggests that the provision of incentives may be effective when combined with other intervention techniques such as peer support, but not when provided as an isolated intervention technique (Chamberlain et al., 2013). *Contingent rewards* are a type of incentive-based intervention that involve providing positive reinforcement (via financial or other material rewards) when behavioral goals are met and withholding that reinforcement when goals are not met (Higgins & Petry, 1999). This approach is based largely on the principles of operant conditioning, and has shown promising results with pregnant women (Donatelle et al., 2004).

Counseling interventions encompass a variety of approaches, ranging from brief (1-3 minute) smoking-specific counseling provided at prenatal care visits to structured cognitive behavioral therapy delivered by trained mental health professionals (Chamberlain et al., 2013; Lumley et al., 2014). There is mixed evidence on the effectiveness of counseling as a primary intervention technique, possibly because of the wide variation encompassed within counseling interventions. In a 2013 meta-analysis of smoking cessation interventions for pregnant women, Chamberlain and colleagues found that the provision of counseling was associated with increased effectiveness when it was combined with other intervention strategies or when it was tailored to the

specific needs of individual women, but not when it was provided as a single component intervention.

Health education interventions focus on increasing knowledge and raising awareness of the risks of smoking and benefits of quitting, as well as providing educational materials, giving instructions, and building skills to promote successful smoking cessation and maintenance (Windsor et al., 1993; Windsor, Boyd, & Orleans, 1998). Educational interventions are often delivered at the individual level, but they can also be delivered at the organizational level (e.g., health system initiatives), community-level (e.g., community-wide awareness campaigns), and population-level (e.g., nationwide media campaigns). Health education is included as a common component in many interventions, and is often supplemented by other intervention techniques. However, a meta-analysis by Chamberlain and colleagues (2013) found that health education was not associated with increased effectiveness when provided alone *or* in combination with other intervention techniques. The lack of effectiveness associated with health education interventions may be explained by the fact that most pregnant women are already aware of the risks of smoking (Coonrod, Bruce, Malcolm, Drachman, & Frey, 2009; Frey & Files, 2006; Orr, Newton, Tarwater, & Weismiller, 2005). Additionally, since health-related information and education are common components of standard prenatal care and smoking cessation interventions, it may be difficult to isolate the effects of these techniques when delivered as the active ingredient(s) of an intervention.

Social support-based interventions are among the most common types of interventions during the prenatal period (Fiore et al., 2002; May & West, 2000). Included

within this category of intervention are structured provider-delivered social support programs, as well as “buddy systems” and other programs aimed at mobilizing social support within a woman’s existing support network (Carlson et al., 2002; May & West, 2000). These may be delivered within the setting of prenatal care, in the community, at home, or via telephone or computer. Evidence suggests that the provision of social support is associated with improved quit outcomes among pregnant women when delivered in higher intensity intervention contexts (McBride et al., 1998), though reviews of social support interventions suggest that poor research methodology may limit the quality of evidence (May & West, 2000).

Nicotine replacement therapy (NRT) involves the use of nicotine gum, patches, lozenges or other delivery-systems to help patients quit smoking by treating the underlying nicotine addiction (Henningfield, Fant, Buchhalter, & Stitzer, 2005). The aim of NRT is to alleviate cravings and other symptoms of nicotine withdrawal to increase the likelihood of successful cessation and maintenance. Clinical practice guidelines call for NRT to be offered to “all smokers trying to quit, except in the presence of special circumstances,” including pregnancy and breastfeeding (Fiore, Jaen, Baker, et al., 2000). There are still significant safety concerns regarding the potential for adverse effects on the fetus (Lumley et al., 2014; Slotkin, 2008). Therefore, it is recommended that nicotine replacement therapy should only be offered to pregnant women as a last resort, and even then, only after careful consideration of whether the risks of continued smoking outweigh the potential risks of nicotine replacement therapy (Siu, 2015).

Moderating factors

Participant characteristics. Reviews of smoking cessation interventions and outcomes indicate that different techniques and intensities may be necessary for different subgroups of pregnant women (Floyd, Rimer, Giovino, Mullen, & Sullivan, 1993). This is especially true for low-SES pregnant women, who appear to reap the most benefit from more intensive interventions (Floyd et al., 1993). Additionally, evidence suggests that women with mental health problems such as depression are less likely to achieve and maintain smoking abstinence than women without such mental health problems (Cinciripini et al., 2000; Rigotti et al., 2006). As stated previously, women who live with partners who smoke are also less likely to successfully quit smoking, suggesting that partner smoking status may moderate the effects of intervention techniques (McLeod 2004; Polanska 2004). These findings highlight the importance of considering the individual characteristics of pregnant women who smoke.

Intervention delivery. In addition to characteristics of the participants and their social contexts, characteristics of intervention delivery may also play an important role in moderating intervention effectiveness. Intervention delivery characteristics include the provider, format, setting and intensity of the intervention (Davidson et al., 2003). Research in this area is largely inconclusive. For example, evidence suggests that smoking cessation services delivered by medical providers are often viewed negatively by pregnant women (Ingall & Cropley, 2010), while nurse-delivered interventions typically have high acceptability among pregnant smokers (Bullock et al., 2009;

Published reports and reviews of smoking cessation programs and other interventions tend to conflate intervention delivery characteristics and intervention techniques, making it difficult to determine whether certain intervention techniques are

more effective when delivered in a certain setting or dose, or by a specific type of provider. Coding and reporting on intervention delivery characteristics would reduce uncategorized intervention content, and therefore facilitate the investigation of how intervention content relates to effectiveness (Michie & Abraham, 2008).

The following section reviews the reliability of these methods.

Methods of Identifying Pregnant Smokers

When evaluating smoking cessation interventions, another characteristic that must be considered is the method of assessing smoking status. Methods for assessing smoking during pregnancy can be broken down into two basic categories: self-report and objectively-validated measures. Although self-reported smoking status is used widely throughout the literature, there is substantial evidence that this method may be unreliable because of the social stigma attached to smoking during pregnancy (Britton, Brinthaupt, Stehle, & James, 2004; Rebagliato, 2002). Studies comparing self-report and biochemical measures of smoking status have found deception rates ranging from 24% (Windsor, Woodby, Miller et al., 2000) to 50% (Kendrick et al., 1995) among pregnant populations. Because of the high potential for bias in self-report measures of smoking status, biochemical markers are the preferred method of assessing smoking among pregnant women.

The most widely used biochemical marker of smoking status is cotinine, a metabolite of nicotine that is considered to be the best indicator of nicotine consumption (Rebagliato, 2002). With a half-life of about 20-hours, cotinine accumulates in bodily fluids such as blood, saliva, and urine, making it a stable marker of recent (past 2-3 days) exposure to nicotine (Rebagliato, 2002). Researchers and medical professionals

have used cotinine cutoff values ranging from 10-25 ng/mL for saliva, 10-20 ng/mL for serum, and 50-200 ng/mL for urine samples to differentiate smokers from non-smokers (Kim, 2016). However, these cut-points were established in studies of non-pregnant populations and may lead to inaccurate assessments of smoking status among pregnant women due to accelerated cotinine metabolism during pregnancy (Dempsey, Jacob, & Benowitz, 2002; Hegaard et al., 2007).

Methods

Search Strategy

Databases & Search Engines. Randomized controlled trials assessing the effectiveness of smoking cessation interventions for pregnant women were identified from six major bibliographic databases: Cochrane Library, PubMed/MEDLINE, Google Scholar, Science Direct, ProQuest, Web of Science, and CINAHL Complete. The search strategy will include combinations of the following keywords: “pregnancy/ OR pregnant/ OR prenatal/ OR antenatal/ OR maternal,” “smoking cessation/ OR tobacco cessation/ OR quit smoking/ OR stop smoking,” “trial/ OR intervention/ OR program/,” “RCT/ OR randomized controlled trial.” We will also include appropriate MeSH terms associated with the keywords.

Given that cultural, organizational, and policy-level factors have been shown to influence the effectiveness of smoking cessation interventions, the current review is limited to interventions delivered within the United States and published in English language journals. Considering major reforms to Medicaid in the 1980s and early 1990s that overhauled funding mechanisms and expanded pregnancy-related coverage to

women with incomes at or below 133 percent of FPL (Boben, 2000), the search was limited to studies published between 1995 and 2015. The search and subsequent meta-analysis were initially completed in 2017, but was updated upon final review of the dissertation in 2020. Although we did not include studies published after 2015 in the meta-analysis, we added a brief narrative review summarizing the results, and also discussed emerging trends and implications in the overall discussion.

Reference List Search. The reference lists of all included articles were reviewed for additional trials.

Journal Search. The ten journals for which the greatest number of articles were returned in the initial search were identified and the content tables inspected to identify any additional trials.

Inclusion/Exclusion Criteria

Criteria for considering studies for this review included characteristics of the study and study design, participant characteristics, intervention design and purpose, nature of comparison group(s), and outcome measure(s). Inclusion/exclusion criteria are described below.

Types of studies. All randomized controlled trials where a primary aim of the study was smoking cessation in pregnant women will be considered. (To be retained, studies must include a measure of smoking abstinence in the second or third trimester.)

Types of participants. Pregnant smokers (18 years and older) in any trimester of pregnancy. Smokers are defined as women who: Smoke an average of at least 1 cigarette/day; and/or self-identify as a current smoker.

Types of interventions. All interventions with the stated purpose of helping pregnant women quit smoking during the prenatal period will be considered. These include: 1) Self-help interventions; peer-led interventions; professional-led interventions; individual interventions; and group interventions; 2) Counseling interventions; educational interventions; incentive/reward-based interventions; social support-based interventions; and other types of interventions targeting psychosocial variables; and 3) Any of the previously-mentioned interventions with or without medication-assisted cessation.

For the first stage of the review, interventions were classified based on the primary strategy (e.g., counseling, education, incentives/rewards, etc). However, since many interventions involve multiple strategies, we coded for all active behavior change techniques. This is described further under Aim 3.

Pharmacological-only interventions were excluded, but interventions were included if they used nicotine replacement therapy as an adjuvant technique (in addition to behavior change techniques).

Types of outcome measures

Primary Outcomes (required for inclusion): The primary outcome of interest is late pregnancy smoking abstinence, defined as point prevalence abstinence (biochemically validated or self-reported) and using the latest smoking status measure taken in pregnancy.

Secondary Outcomes (not required for inclusion): Additional outcomes of interest include: 1) **Other measures of smoking behavior** (including reduction in smoking, as measured by daily cigarette consumption or biochemical markers of tobacco consumption; secondary smoke exposure; continued abstinence in the postpartum

period, etc); 2) **Perinatal outcomes** (including birthweight; low birthweight [proportion of births at less than 2500 g] and very low birthweight [less than 1500 g]; preterm birth [births at less than 37 weeks]; other adverse perinatal outcomes [e.g., fetal growth restriction]; and perinatal deaths); 3) **Maternal outcomes** (including measures of psychological health [such as anxiety, depression, and stress] and physical health [such as pregnancy-related complications and self-reported health status); 4) **Measures of theoretical determinants of behavior change**[1], including: *knowledge* ("an awareness of the existence of something"[2]), *skills* ("an ability or proficiency acquired through practice"), *social role/identity* ("a coherent set of behaviors and displayed personal qualities of an individual in a social setting"), *beliefs about capabilities* ("acceptance of the truth, reality, or validity about an ability, talent, or facility that a person can put to constructive use"), *optimism* ("the confidence that things will happen for the best or that desired goals will be attained"), *beliefs about consequences* ("acceptance of the truth, reality, or validity about outcomes of a behavior in a given situation"), *reinforcement* ("increasing the probability of a response by arranging a dependent relationship, or contingency, between the response and a given stimulus"), *intentions* ("a conscious decision to perform a behavior or a resolve to act in a certain way"), *goals* ("mental representations of outcomes or end states that an individual wants to achieve"),

[1] Theoretical determinants were specified a priori using the Theoretical Domains Framework (Cane, O'Connor, & Michie, 2012).
See: Cane, J., O'Connor, D., & Michie, S. (2012). Validation of the theoretical domains framework for use in behavior change and implementation research. Implementation Science 7(37), 1-17.

[2] Definitions were derived from the American Psychological Association's Dictionary of Psychology (2007), as used in Cane, O'Connor, & Michie's (2012) Theoretical Domains Framework.
See: American Psychological Association (APA): APA Dictionary of Psychology. Washington, DC: American Psychological Association; 2007.

memory, attention, & decision processes (“the ability to retain information, focus selectively on aspects of the environment, and choose between two or more alternatives”), *environmental context and resources* (“any circumstance of a person’s situation or environment that discourages or encourages the development of skills and abilities, independence, social competence, and adaptive behavior”), *social Influences* (“those interpersonal processes that can cause individuals to change their thoughts, feelings, or behaviors”), *emotion* (“a complex reaction pattern, involving experiential, behavioral, and physiological elements, by which the individual attempts to deal with a personally significant matter or event”), *behavioral regulation* (“anything aimed at managing or changing objectively observed or measured actions”), and other[3] relevant constructs included in the published studies. Please see Appendix B for expanded operational definitions of each of the theoretical constructs.

Data Collection and Analysis

Selection of studies

Abstracts of search results were reviewed for relevance. Those that clearly did not meet inclusion criteria (e.g., studies conducted outside of the U.S.; interventions that used only pharmacological treatment) were removed after a cursory review of the title, abstract, and/or publication information. The primary reviewer then examined and applied the inclusion/exclusion criteria to the full text of all remaining articles returned in the search, while a second independent reviewer examined the full text of a random selection of 20% of the returned articles, and the results were reviewed for agreement.

[3] Additional theoretical determinants may be derived inductively for constructs not identified a priori.

Where there was disagreement, consensus was reached by discussion and review of the criteria for inclusion.

Data extraction and management

Data from selected studies were extracted by two independent reviewers using a structured form. We first pilot-tested the data extraction form on a subset of studies not included in the current review to identify any problems or sources of confusion, and made revisions where necessary. Using the revised forms, the primary reviewer performed data extraction on 100% of the sample, while the second reviewer independently performed data extraction on 20% of the sample, and the results were compared. Interrater reliability was calculated to measure agreement between the two reviewers. Where there was disagreement, consensus was reached by discussion (Rosenthal, 1987).

Assessment of risk of bias in included studies

The methodological quality of the included studies was assessed using the guidelines recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins & Green, 2011). An overall risk of bias assessment (high, low, or unclear) was made based on the following criteria:

1) Sequence generation (checking for selection bias): For each included study we reviewed the methods used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups. Methods were categorized as:
    - Low risk of bias (any truly random allocation process, e.g., random number table, random number generator);

- High risk of bias (any non-random process, e.g., odd or even birthdate, hospital or clinic record number); or
- Unclear risk of bias

2) Allocation concealment (checking for selection bias): For each included study, we reviewed the methods used to conceal the allocation in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrollment. Methods were categorized as:
    - Low risk of bias (e.g., telephone, web-based, or other central randomization; sequentially numbered sealed opaque envelopes);
    - High risk of bias (e.g., open random allocation, such as a list of random numbers; assignment envelopes used without appropriate safeguards; medical record number; date of birth; any other explicitly unconcealed procedure); or
    - Unclear risk of bias
3) Masking (checking for performance bias): For each included study, we reviewed the methods used (if any) to mask study participants and key study personnel (e.g., intervention providers and outcome assessors) from knowledge of which intervention arm a participant received. However, masking is often not feasible (particularly for providers) in the context of psychosocial and educational interventions. Methods were categorized as:
    - Low risk of bias (either [1] masking of participants and key study personnel ensured, and unlikely that the masking could have been compromised, or [2] partial or no masking, but the reviewers judge that the

outcome and the outcome measurement are not likely to be influenced by lack of masking);

- High risk of bias (either [1] likely that masking of participants and key study personnel could have been compromised, or [2] partial or no masking, and the reviewers judge that the outcome and the outcome measurement are likely to be influenced by lack of masking); or
- Unclear risk of bias

4) Incomplete Outcome Data (checking for attrition bias): For each included study, we reviewed the completeness of outcome data for the primary outcome, including attrition and exclusions from the analysis. We noted whether attrition and exclusions were reported, the numbers in each intervention group (compared with the total number of randomized participants), reasons for attrition or exclusions where reported, and whether missing data were balanced across groups. Methods were categorized as:
    - Low risk of bias (any one of the following: [1] no missing data; [2] reasons for missing data unlikely to be related to true outcome; [3] missing data balanced in numbers across intervention groups, with similar reasons for missing data across groups; [4] missing data not enough to have a clinically relevant impact on observed effect size; or [5] missing data imputed using appropriate methods);
    - High risk of bias (any one of the following: [1] reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; [2] enough

missing data to have a clinically relevant impact on observed effect size; [3] “as-treated” analysis with significant departure of intervention received from assigned at randomization; or [4] potentially inappropriate application of simply imputation); or

- Unclear risk of bias

5) Outcome Reporting (checking for selective reporting bias): For each included study, we reviewed how the possibility of selective outcome reporting bias was examined and what was found. Methods were categorized as:
    - Low risk of bias (where it is clear that all of the studies’ pre-specified outcomes and expected outcomes have been fully reported);
    - High risk of bias (where outcomes were not fully reported because of one or more of the following: [a] one or more reported primary outcomes were not pre-specified; [b] not all pre-specified outcomes were reported; [c]outcomes of interest are reported incompletely and cannot be used; or [d] study fails to report results of a key outcome that would be expected to have been reported in such a study); or
    - Unclear risk of bias
6) Reliability of outcome measures (checking for detection bias): Because of the inconsistency of self-reported measures of smoking status, biochemical validation of smoking abstinence is considered the standard for smoking cessation trials (West, 2005; Shipton, 2009). Therefore, biochemical measures (e.g., cotinine levels, expired air carbon dioxide) are the preferred method for assessing smoking outcomes. For each study we noted whether the smoking

outcome was biochemically validated (including specification of the measure[s] used) or measured via self-report only. Measures were categorized as:

- Low risk of bias (biochemically validated);
- High risk of bias (not biochemically validated); or
- Unclear risk of bias

Where possible, we also reported the reliability (e.g., internal, test-retest) of the instruments used to assess any outcome measures.

7) Implementation of intervention: Three common types of implementation problems (Walsh, 2000) were assessed:
    - Not all participants in intervention group received the intervention;
    - Intervention group participants did not receive all components of the intervention; and/or
    - Control group participants receiving some or all of the intervention.

    Where possible, we reviewed the results of any process measures or evaluation(s) reported. Implementation of the intervention was categorized as:
    - Low risk of bias (process evaluation indicates that most participants received the intervention as planned);
    - High risk of bias (process evaluation indicates that a significant proportion of participants did not receive the intervention as planned); or
    - Unclear risk of bias (no process evaluation reported)
8) Other bias: We also considered any additional sources of bias in the study, such as conflicts of interest, which were categorized as:
    - Low risk of bias (study appears to be free of other sources of bias)

- High risk of bias (there is at least one important additional risk of bias)
- Unclear risk of bias

The overall risk of bias score was coded as low risk if no significant sources of bias were present and/or if sources of bias, when present, were unlikely to seriously alter the results. Studies were coded as high risk if sources of bias were present and posed a substantial risk of affecting the interpretation of results. Studies were coded as unclear risk sources of bias were present and could raise doubt about the validity of results, but did not clearly influence the study results or interpretation.

Measures of Treatment Effect

Dichotomous data

For dichotomous data, the results were calculated as risk ratios (RR) with 95% confidence intervals, where RR = $\frac{a/(a+b)}{c/(c+d)}$, as seen in the table below.

	Outcome 1 (e.g., smoking)	Outcome 2 (e.g., not smoking)
Intervention	A	b
Control	C	d

In line with the standards of the Cochrane Tobacco Group, smoking cessation outcomes were converted from an odds ratio for continued smoking, to a RR for quitting. Therefore, an average RR > 1 indicates a positive outcome. For secondary outcomes where fewer events are desired (e.g., preterm birth; depression, mean # of cigarettes/day), an average RR < 1 is interpreted as a positive outcome. For tests

involving cell frequencies of zero, 0.5 was added to each cell in order to have defined odds ratios prior to computing the risk ratio.

Ordinal data

When possible, we treated data reported on an ordinal scale as a continuous outcome, as recommended in Section 9.2 of the Cochrane Handbook. Data were dichotomized if the original analysis did not allow it to be summarized using methods for continuous data, or if there was a conceptually logical cut-point (e.g., if smoking is measured by quantity of cigarettes/data cut-point could be introduced to dichotomize smokers [≥ 1 cigarette/day] and non-smokers [0 cigarettes/day]).

Unit of analysis issues

While the effects of clustering can be adjusted using an intra-cluster correlation coefficient (ICC), the current review only included two cluster randomized trials. Although there is not a hard rule for the number of studies required to perform an analysis on a subgroup of studies, it is generally recommended that at least 3 studies are needed to form a unique subgroup of any kind (Borenstein, Hedges, Higgins, & Rothstein, 2009). Therefore, we excluded these two studies from the primary meta-analysis due to the potential for bias; however, we included these studies when coding for behavior change techniques, use of theory, and other descriptive statistics.

Comparison Groups

Studies with multiple intervention arms (e.g., a control group plus treatment 1 group plus treatment 2 group) present unit of analysis issues if multiple comparisons are made against a single control group. To address this issue, we used an approach put forth by the Cochrane Handbook and described in section 16.5.4. This approach, which

has been employed in prior meta-analyses (see Lumley et al., 2014), involves selecting one intervention arm and excluding the other(s) to create a single pairwise comparison. While this results in a loss of information due to the exclusion of one (or more) intervention arm, it was the most appropriate choice given the specific goals of this meta-analysis. When undertaken, we selected the intervention arm that was specified by the authors as the primary focus of the study. If the authors did not provide such specification, we selected the intervention arm with the greatest number of "active" behavior change techniques (please see Aim 3 for a detailed description of "active" behavior change techniques).

Prior to deciding on this approach, we considered several options for dealing with studies with multiple intervention arms. The most commonly-used approach is to combine the intervention groups into a single group to create a single pairwise comparison. However, this is problematic when the same study compares multiple, conceptually different intervention techniques (e.g., counseling vs. incentives vs. control group). Given that the purpose of this meta-analysis was to isolate specific components of interventions, combining two or more conceptually different intervention arms into one group would make it impossible to identify the unique active ingredients, and thus was considered to be inconsistent with our goals.

We also considered an alternative approach, which involves creating somewhat independent, artificial comparisons by dividing up the shared intervention group evenly among the comparisons, as described in section 16.5.4 of the Cochrane Handbook (for example, if a study compared 100 patients receiving incentives to 200 patients receiving counseling to 300 patients in a control group, this approach would create comparisons

of 50 incentive participants vs. 200 counseling participants and 50 incentive participants vs. 300 control participants). However, because of the small intervention-group sample sizes in several of the multiple-arm studies, the loss of power created by cutting the sample size in half would likely influence the results in a significant and conceptually meaningful way, particularly when assessing heterogeneity statistics (for more information, please see the section below entitled “Heterogeneity of Effect Sizes”). Additionally, as discussed in section 16.5.5 of the Cochrane Handbook, constructing multiple comparisons from the same study conflicts with the assumptions of a random effects model. According to the Handbook: “A random-effects meta-analysis allows for variation by assuming that the effects underlying the studies in the meta-analysis follow a distribution across studies. The intention is to allow for study-to-study variation. However, if two or more estimates come from the same study then the same variation is assumed across comparisons within the study and across studies.”

Statistical Analyses

Data Synthesis

Meta-analysis is a statistical method for systematically synthesizing data from multiple, independent studies assessing similar outcomes (Brockwell & Gordon, 2001; Cooper, Hedges, & Valentine, 2009; Lispsey & Wilson, 2001). The findings from meta-analyses are reported in the form of effect sizes, which provide an indication of the magnitude of change evident across all studies included in the analysis, as well as selected subsets of studies. While some meta-analyses seek to re-test the original hypotheses tested in the individual studies, the current review seeks to extract relevant

data for the purpose of testing new hypotheses that were not addressed in the primary studies, thus making a unique contribution to the literature.

Analyses were conducted using Comprehensive Meta Analysis Software (Borenstein, Hedges, Higgins, & Rothstein, 2005). To combine the effects of the studies, we used random-effects meta-analysis, which assumes that the effects in the studies are *not* all the same and thus accounts for this additional source of variation (DerSimonian & Laird, 1986; Hedges & Vevea, 1998). Compared to the alternative fixed effects model, the random effects model is more conservative and considered a more appropriate model for combining the results of studies that may differ clinically (e.g., characteristics of participants or intervention design) and/or methodologically (e.g., differences in measurement). Additionally, when compared to fixed effects models, random effects models provide a more accurate estimate of the degree of precision in meta-analytic findings, and demonstrate a much lower risk of Type 1 bias in significance tests (Hunter & Schmidt, 2000).

Random effects models were used to calculate an overall risk ratio (for all trials), as well as to calculate risk ratios for subgroup analyses.

Heterogeneity of Effect Sizes

The goal of evidence synthesis is not simply to produce a summary effect size, but rather to make sense of the pattern of effects. When effect sizes are not consistent across studies, it is important to identify this and attempt to explain the sources of heterogeneity. Statistical heterogeneity refers to variation in true effect sizes. However, the total observed variation in effects reflects both true heterogeneity *and* random

within-study error, so it is necessary to use multiple indicators of heterogeneity in order to identify how much of the total variance reflects real differences in effect sizes.

Statistical heterogeneity was assessed using both the *Q* statistic (Cochran, 1954) and I^2 statistic (Higgins & Thompson, 2002), in addition to visual inspection of the forest plots. A significant Q statistic and a high I^2 value indicate the presence of variance that is not due to sampling error and that may be accounted for by moderators. The Q statistic reflects the ratio of observed variability in effect sizes to expected variability in effect sizes. It tests the null hypothesis that all studies share a common effect size. Therefore, a statistically significant *Q* statistic is interpreted as an indicator of true heterogeneity of effect sizes, and indicates that moderator analyses are appropriate. However, the inverse is not necessarily true, as a non-significant *Q*-value may reflect low power (rather than the absence of heterogeneity). It has also been suggested that the Q-statistic may have excessive power to detect negligible heterogeneity when the sample size is large, which may lead to the erroneous conclusion that there is significant true heterogeneity when none exists. Furthermore, while the Q statistic can be used to evaluate the statistical significance of true heterogeneity, it cannot be used to quantify the extent of true heterogeneity.

The I^2 statistic describes the proportion of total variation across studies that is due to non-random heterogeneity, and as such, provides an indicator of the proportion of observed variance that reflects true differences in effect size (Higgins, Thompson, Deeks, & Altman, 2003). Using criteria specified by Higgins and colleagues (2003), I^2 values of 25% or less were interpreted as an indicator of low heterogeneity, 50-74% as moderate, and 75% or greater as high. Higher values indicate that a larger proportion of

variance is non-random, and thus point to the need for further exploration of heterogeneity through moderator analyses and/or meta-regression.

Subgroup Analyses

In accordance with the goals of this meta-analysis, subgroup analyses were conducted on categorical variables to investigate whether effect sizes were influenced by 1) characteristics of the study/intervention; 2) characteristics of the study sample; and 3) use of theory. These analyses were conducted using the subgroup method for moderator estimation explicated by Lipsey and Wilson (2001). This method, which is described as a meta-analytic analogue to the analysis of variance (ANOVA) test, allows for the regrouping of effect sizes into mutually exclusive groups based on scores on the moderator variable. The subgroup test assesses the level of heterogeneity of effect sizes using the Q-statistic (which represents the sum of squares) at between-group (Q_b) and within-group (Q_w) levels, where Q_b is analogous to the ANOVA's F-test (Borenstein, Hedges, Higgins, & Rothstein, 2009). A significant between-group heterogeneity statistic (Q_b) is indicative of effect sizes that vary across groups by more than would be expected due to sampling error alone. A significant within-group heterogeneity statistic (Q_w) indicates that heterogeneity exists within the group, beyond what can be explained by the moderator (i.e., that the effect sizes within a group are statistically different from each other) (Lipsey & Wilson, 2001). However, when using a random effects model, Q_w is typically considered not meaningful due to the assumptions of the model (Borenstein et al., 2009). In summary, if the between-group heterogeneity statistic is significant (i.e., if the p-value for Q_b is < 0.05), this indicates that there are significant differences

between levels of the subgroup (e.g., low SES vs. not low SES samples; high psychosocial risk vs. low psychosocial risk; high vs. low vs. unclear risk of bias, etc).

In addition to the Q-statistic, we also evaluated confidence intervals surrounding the effect size of each subgroup as an additional indicator of statistical significance, using the approach outlined in section 9.6.3.1 of the Cochrane Handbook. As stated in the Cochrane Handbook, "[n]on-overlap of the confidence intervals indicates statistical significance, but […] the confidence intervals can overlap to a small degree and the difference still be statistically significant." This approach has been successfully employed in methodologically-similar meta-analyses (e.g., Hysong, 2009).

Categorical variables of interest related to study/intervention characteristics included: Intervention type (contingent rewards/incentives; counseling; social support; NRT + counseling; educational); intervention delivery type (counselor; medical provider; trained peer educator or peer counselor; trained study staff; technology-delivered; volunteer); risk of bias (high/low/unclear); intensity of intervention contact (reflecting frequency and duration of participant contact with intervention deliverer and/or materials, where 1=low; 2=moderate; and 3=high); use of cultural tailoring (yes/no); use of organizational- or provider-level change strategies (yes/no); referral to community resources (yes/no); and assessment of smoking in the woman's social network (yes/no).

Categorical variables of interest related to the study sample included: Low SES sample (yes/no, where yes indicated that the study authors explicitly described the sample as low-SES **and/or** more than 50% of participants had less than a high school education, were on Medicaid, and/or were receiving WIC benefits); majority minority sample (yes/no, where yes indicated that at least 50% of the sample was comprised of

racial or ethnic minority participants); health of the study sample (yes/no, where yes indicated that the study specifically focused on a population with mental health or substance use disorders); high psychosocial risk (yes/no, where yes indicated that that study participants were explicitly described as high risk **and/or** > 50% of participants reported low social support and/or high perceived stress).

Subgroup analyses were also used to evaluate the use of theory as a potential source of variability of effect sizes. Categorical variables of interest related to the use of theory included: Intervention based on a single theory (yes/no); explicit mention of any behavior change theory (yes/no); and the specific theory that was mentioned (Operant conditioning; Transtheoretical Model; Social Cognitive/Social Learning Theory). This is described in more detail under Aim 2.

The use of each of the 27 BCTs was also investigated through subgroup analyses, with yes indicating that the intervention used the specified BCT. In addition to the 27 BCTs, we also coded for the use of guidance on dealing with specific smoking-related triggers and the use of strategies aimed at emotional regulation (e.g., strategies to alleviate depressive symptoms). This is described in more detail under Aim 3.

Meta-Regression

Since subgroup analyses can only be used to examine heterogeneity among different levels of categorical variables, we used random effects univariate meta-regression models to examine heterogeneity explained by continuous variables. Similar to simple regression, meta-regression examines how an outcome variable (i.e., effect size estimate) is predicted by one or more explanatory variables (or covariates).

As recommended in Section 9.6.4 of the Cochrane Handbook, the log-transformed value of each risk-ratio will be used when conducting the meta-regressions on our primary outcome. As such, the exponential of the regression coefficient is interpreted as an estimate of the relative change in intervention effect with a unit increase in the explanatory variable. The proportion of between-study variance explained by the covariate will be calculated using the adjusted R^2 statistic, which compares the estimated between-study variance when covariates are included in the model with the value of the between-study variance when covariates are *not* included in the model (Kelley & Kelley, 2012). The Q-statistic derived from the goodness-of-fit test provides an indicator of whether the model accounted for significant heterogeneity. In meta-regression, the goodness-of-fit test assumes a null hypothesis that unexplained variance is zero. Thus, a non-significant Q-statistic indicates that the model explained significant variation in the distribution of effect sizes, while a significant Q-statistic indicates the presence of significant variance that was *not* explained by the model.

Continuous variables of interest included: Gestational age at baseline; cigarettes smoked per day at baseline; Theory Coding Scheme subscale scores and total score; and total number of BCTs used.

Results

Search Results

The database searches yielded a total of 1,223 unique articles. Of these, 805 were excluded based on the title or abstract not containing data suggesting the study was a randomized trial or otherwise did not meet inclusion criteria. After the abstract and title review, 419 articles were retrieved for further (full text) review. The reference

list and journal search yielded 7 additional, non-duplicate articles that met all inclusion criteria and were included in the review. Of these articles, 386 were removed because they did not meet one or more of the criteria for inclusion, including 229 that did not meet study design criteria, 88 that were conducted outside of the U.S., 27 that did not meet study population criteria, 31 that did not meet reporting or outcome requirements, and 11 for other reasons. After the full-text review, a total of 31 articles were retained for the review. Combined with the 7 additional articles identified through the reference list and journal search, this resulted in a study sample of 38 articles representing 38 independent trials. Relevant articles associated with the trials were used to retrieve additional study characteristics as needed.

Study Characteristics

Table 1.0.1 presents an overview of the characteristics of included studies. Of the 38 trials included in the review, 36 were randomized controlled trials where the unit of randomization was the individual and 2 were randomized at a cluster. The majority of interventions (n=30) were traditional two-armed trials (with a control group compared to a treatment group), while eight trials included three or more arms, adding up to a total of 87 study arms across the 38 trials.

Table 1.0.1: Study Characteristics

Study Characteristics	**No. (K)**	**%**
Design		
RCT	36	95%
Cluster R	2	5%
Number of arms in trial		
Two	30	79%
More than two	8	21%

Risk of Bias		
Low	13	34%
High	12	32%
Unclear	13	34%
Sample size		
0-50 participants	3	8%
51-100 participants	7	18%
101-500 participants	23	61%
> 501 participants	5	13%

Study quality and adequacy of reporting also varied significantly. Using Cochrane's Risk of Bias tool (Higgins & Green, 2011), 34% of trials (n=13) were categorized as low risk of bias, indicating that sources of bias, if present, are unlikely to seriously alter the results. Thirty-two percent of trials (n=12) were coded as high risk, indicating that sources of bias are present and pose a substantial risk of affecting the interpretation of results. The most common sources of potential bias included high attrition rates, incomplete implementation, and incomplete outcome reporting. The remaining 34% of trials (n=13) were coded as unclear risk, indicating that sources of bias could raise doubt about the validity of results. In most cases, studies coded as 'unclear risk' were categorized as such due to inadequate specification of randomization procedures, blinding, and/or allocation concealment.

Participant Characteristics

Most participants were low-SES women in their mid-twenties (mean age 25.5 years of age; Range = 22 to 30.5 years), with at least one previous pregnancy. The mean gestational age of participants at baseline was 15 weeks (Range = 9.2 to 28 weeks). Most study samples were described as generally healthy, while 4 trials specifically focused on populations with mental health or substance use disorders. In 19 studies, more than 50% of participants had less than a high school diploma, and 28

studies were coded as “low SES” (study authors described the sample as low-SES **and/or** more than 50% of participants had less than a high school education, were on Medicaid, and/or were receiving WIC benefits). Ten studies were coded as “majority minority” (at least 50% of the sample non-white). Twenty-two studies were coded as “high psychosocial risk” (study participants were explicitly described as high risk **and/or** >50% of participants reported low social support and/or high perceived stress).

Smoking habits among participants varied significantly. The average participant began smoking between the ages of 14-16 and reported smoking an average of 19.2 cigarettes per day prior to learning of their pregnancy (Range = 13-25.6). Average cigarette consumption declined after women learned they were pregnant. At baseline, participants reported smoking an average of 9.7 cigarettes per day (Range = 5-18), a reduction of 10.5 cigarettes/day from the pre-pregnancy period. In addition to personal cigarette smoking, environmental tobacco exposure was common: In the 14 studies that reported on the presence of other smokers in the household, at least 50% of participants said they lived with a smoker; in 9 of these studies, at least 70% of women reported living with a smoker.

Intervention Characteristics

As seen in **Table 1.0.2**, sixteen interventions were categorized as ‘single’ interventions, indicating that the main intervention strategy accurately and comprehensively reflected all intervention content. The other 22 interventions were coded as ‘multiple’ interventions, meaning that additional, distinct intervention strategies (on top of the main intervention strategy) were offered to all treatment-group participants (e.g., incentives as a main strategy, but supplemented by social support). When

categorized by the main intervention strategy, 19 of the trials were coded as ‘counseling’, nine as ‘vouchers/incentives’, six as ‘social support’, three as ‘nicotine replacement therapy’ (supplemented by behavioral and/or psychosocial counseling), and one as ‘education’. Twenty-four trials included intervention content that was tailored or personalized according to participants’ smoking-related attitudes, beliefs, or behaviors, while four studies included content that was tailored specifically to participants’ racial, ethnic, or cultural background. Additionally, ten trials specifically elicited participants’ feedback about the helpfulness and/or acceptability of the intervention.

Table 1.0.2: Intervention Characteristics

Intervention Characteristic	No. (K)	%
Type of Intervention		
Single	**16**	**42%**
Multiple	**22**	**58%**
Main Intervention Strategy		
Counseling	**19**	**50%**
Vouchers/Incentives	**9**	**24%**
Social Support	**6**	**16%**
NRT (+supplement)	**3**	**8%**
Educational	**1**	**2%**
Deliverer		
Study personnel	**13**	**34%**
Mental health counselors	**8**	**21%**
Medical providers	**7**	**18%**
Peer educators	**4**	**11%**
Other	**3**	**8%**
Primary mode of Delivery		
Face-to-face	**19**	**50%**
Telephone, video, or computer	**16**	**42%**
Equal mix of both	**3**	**8%**
Setting (of trial)		
Community clinics	**12**	**32%**
Hospital-based clinics	**10**	**26%**
Medicaid/WIC clinics	**9**	**24%**
Other	**7**	**18%**

Setting (of delivery)		
Primarily within clinic	**24**	**64%**
Primarily outside of clinic	**14**	**36%**
Tailored		
For culture or ethnicity	**24**	**64%**
For smoking habits or beliefs	**4**	**11%**
Low SES sample		
Yes	**28**	**74%**
No	**10**	**26%**
Majority Minority sample		
Yes	**10**	**26%**
No	**28**	**74%**
High Psychosocial Risk		
Yes	**22**	**58%**
No	**16**	**42%**

Most interventions were delivered by trained study staff (n=13), mental health counselors (n=8), or medical providers, which included doctors and nurses (n=7). Four interventions were categorized as “peer-delivered”, which included trained peer educators, peer counselors, and peer supporters. Two interventions were delivered via technological resources (one by computer, and one by video), and one was delivered by trained volunteers. Related to the source of delivery is the mode of delivery. Most trials involved multiple modes of delivery (e.g., written materials, face-to-face contact, telephone calls), but for the purposes of this analysis, we coded for the primary mode of delivery for the main intervention strategy. Of the 38 trials, half (n=19) were primarily delivered via face-to-face contact and 16 were delivered primarily by telephone, video, and/or computerized systems (n=11 were delivered by telephone; n=5 by video or computer). The remaining 3 trials involved an equal mix of face-to-face and technologically-delivered intervention strategies.

Most interventions were based within community prenatal care clinics (n=12), while 10 took place in hospital-based prenatal care clinics and another 9 took place in Medicaid/WIC-specific prenatal care clinics. The remaining 7 intervention sites were categorized as 'other', with two taking place in managed care/HMO's, two in clinics on military bases, one in an OB-clinic, one in a 'multispecialty clinic', one in an addiction center. While all of the interventions involved at least some contact with participants within a clinic setting, the main intervention strategy was often delivered in a setting outside of the clinic. Out of the 38 total interventions, 24 were delivered primarily within the setting of a clinic or other medical center, while 14 were delivered outside of a medical setting (e.g., via telephone or contact with social supporters at home and in the community).

Primary Outcome: Smoking abstinence in late pregnancy

As seen in **Table 1.1.0**, the results of a random effects model using data from 34 randomized controlled trials revealed a significantly larger effect size for smoking abstinence in late pregnancy (28 weeks through birth) in the treatment groups compared to control groups (RR = 1.53; 95% CI: 1.30-1.79). The heterogeneity statistic *Q* was statistically significant (Q[33]=63.04; p=0.01), and the corresponding I^2 statistic indicated that approximately 47.7% of the heterogeneity reflected true differences in effect size. Based on the presence of significant heterogeneity, subgroup analyses were deemed to be appropriate.

Subgroup Analysis: Intervention Type

As seen in **Table 1.1.1,** the subgroup analysis by intervention type (contingent vouchers; counseling; social support; NRT + counseling supplement; educational), two

groups had effect sizes that were significantly larger for the treatment group compared to intervention group: voucher/contingent rewards-based interventions (n=9) (RR = 2.82; 95% CI: 2.05-3.88) and counseling interventions (n=16) (RR=1.30; 95% CI: 1.10-1.54). Risk ratios were not significantly different between the treatment and control groups for interventions classified as NRT+ counseling supplement (n=3) (RR = 2.81; 95%CI: 0.74-10.70) or social support (n=6) (RR=1.18; 95% CI: 0.91-1.53). Only one study was classified as an educational intervention; as such, there was insufficient data to perform subgroup analyses on this type intervention. Between-group heterogeneity was significant ($Q_b[3]=21.61$, $p<0.001$), indicating that effect sizes across groups differed by more than sampling error.

The proportion of true heterogeneity was reduced to zero ($I^2=0\%$) for the contingent reward-based subgroup, indicating that intervention type accounted for all of the within-group variance in this subgroup. For the counseling subgroup, the proportion of true within-group heterogeneity was low to moderate ($I^2 = 38.2\%$), indicating the presence of some unexplained within-group variance due to other factors such as characteristics of intervention delivery and/or participants.

Removing two studies that were identified as outliers based on the forest plot (Pollak, 2007; Tuten, 2012) reduced the heterogeneity statistic Q from Q[3]=21.61 ($p<0.001$) to Q[3]=18.73 ($p<0.001$), but the results of the subgroup analysis did not change.

Taken together, the results indicate that contingent voucher-based interventions and counseling interventions were the only two categories of interventions that significantly increased the likelihood of achieving late-pregnancy smoking abstinence

compared to their respective control groups. Contingent-voucher based interventions appear to be the most effective, as this subgroup of interventions had the largest effect size. Classifying the studies by intervention-type reduced overall between-study heterogeneity from Q[33]=63.04 to Q[3]=21.61, indicating that intervention type accounted for about 65.7% of the between-study variance.

Note: In a meta-regression model, intervention-type accounted for 66% of the between-study variance. Using contingent-voucher-based interventions as the reference group, the regression coefficients for counseling-based interventions (b= -0.782; 95%CI: -1.16- -0.401; p=0.0001) and social support-based interventions (b= -0.835; 95%CI: -1.30- -0.366; p=0.0005) indicated that these two types of intervention were associated with a significantly reduced likelihood of late-pregnancy smoking abstinence. The regression coefficient for interventions classified as NRT+ counseling supplement was negative but not significant, indicating that the effect size did not significantly differ from the contingent-voucher-based intervention reference group.

Subgroup Analysis: Risk of Bias

When grouped by risk of bias, the effect size for late pregnancy smoking abstinence remained significant across all three levels (see **Table 1.1.2** for full results) and all three sets of confidence intervals overlapped, indicating that intervention effectiveness did not differ according to risk of bias classification. This was further confirmed by a non-significant between-group heterogeneity statistic (Q_b[2]=1.14; p=0.565). Within-group heterogeneity remained moderate to high for ‘high risk’ (I^2=60.38) and ‘unclear risk’ (I^2=59.81) groups, while it was much lower in the ‘low-risk’

group (I^2=13.90), indicating that the proportion of true heterogeneity of effect sizes was lower for studies categorized as low risk of bias.

In a meta-regression model, risk of bias did not explain any of the between-study variance (R^2=0.0), which confirms the findings of the subgroup analysis.

Subgroup Analysis: Intervention Deliverer

As seen in **Table 1.1.3**, when grouped by intervention deliverer (counselor, medical, peer, trained study staff, technology-delivered, or volunteer), the effect size for late pregnancy smoking abstinence was significantly larger for the treatment group than the control group only for interventions delivered by counselors (n=11; RR=1.42; 95% CI: 1.08-1.85) and trained study staff (n=12; RR=1.84; 95% CI: 1.31-2.57). Overlapping confidence intervals indicated that the difference *between* these subgroups was not significant. This was further confirmed by a non-significant overall between-group heterogeneity statistic (Q_b[5]=2.72; p=0.743).

The effect sizes for interventions delivered by medical providers (n=5; RR=1.35; 95% CI: 0.89-2.04) and peer educators or peer counselors (n=4; RR=1.38; 95% CI=0.93-2.04) were not significantly different when comparing treatment to control groups. Insufficient data prohibited subgroup analyses of interventions classified as volunteer-delivered (n=1) and technology-delivered (n=1).

Even when grouped by intervention deliverer, the proportion of true within-group heterogeneity remained moderately high for interventions delivered by counselors (I^2=54.26) and trained study staff (I^2=60.52), as well as for interventions delivered by medical providers (I^2=45.83), indicating that additional variables were contributing to the

observed heterogeneity in late-pregnancy smoking abstinence. Peer-delivered interventions were characterized by within-group homogeneity (I^2=0.0).

The limited sample size within subgroups limits power for evaluating the statistical significance of the differences in effect sizes between subgroups. Thus, it is possible that the non-significant findings in this subgroup analysis were due to low power, rather than homogeneity of effect sizes. With that in mind, the results suggest that although interventions delivered by counselors and trained study staff were most likely to promote smoking cessation, intervention deliverer was not a significant source of heterogeneity. This is further supported by the results of a meta-regression model, which showed that intervention deliverer did not account for any of the between-study variance (R^2=0.00).

Subgroup Analysis: Contact Intensity

See **Table 1.1.4** for full results. When grouped by the intensity of contact (reflecting both the duration and frequency of contact), risk ratios for late pregnancy smoking cessation remained significant for all three levels of the variable, and a non-significant between-groups heterogeneity statistic indicated that there were no significant differences between levels (Q_b[2]=1.72; p=0.422).

In a random-effects meta-regression model using level 1 (the lowest intensity) as a reference group, contact intensity did not account for any of the between study heterogeneity (R^2 analog = 0.00), which supports the findings of the subgroup analysis. Hence, contact intensity was not a significant source of heterogeneity.

Subgroup Analysis: Context (within routine prenatal care or not)

See **Table 1.1.5** for full results. When grouped by whether the intervention was delivered as part of routine prenatal care or not, the effect size for late pregnancy smoking abstinence was only significantly larger for the treatment group compared to the control group for interventions that were delivered within the context of routine prenatal care (n=21; RR= 1.84; 95% CI: 1.47-2.23). A significant heterogeneity statistic (Q_b[1] =6.99; p=- 0.008) indicated that the difference between groups was significant, and non-overlapping confidence intervals supported this conclusion. Thus, it can be concluded that intervention effectiveness varied significantly depending on the context of delivery, such that the effect size for interventions delivered within the context of prenatal care was greater than the effect size for interventions delivered outside of routine prenatal care. Heterogeneity was nearly identical for both levels of this variable, with approximate I^2 values of 41.5, indicating the presence of a moderate degree of within-group heterogeneity.

In a univariate random effects meta-regression model, the context of the intervention accounted for 19% of the between-study variance. The regression coefficient for interventions delivered within the context of prenatal care was significant in a positive direction (b= 0.388; 95% CI: 0.0804-0.696; p=0.014), indicating that the likelihood of late pregnancy smoking cessation was significantly greater for participants in interventions delivered within the context of prenatal care compared to the reference group of interventions delivered outside the context of routine prenatal care. These results are in line with the findings of the subgroup analysis.

Subgroup Analysis: Cultural Tailoring

As seen in **Table 1.1.6,** when grouped by the presence of cultural tailoring (or not), the effect size for late pregnancy smoking abstinence was only significantly greater for treatment group participants compared to control group participants for interventions that were not tailored specifically to the culture of intervention participants (n=31; RR=1.59; 95%CI: 1.33-1.90 for non-culturally tailored studies versus n=3; RR=1.33, 95%CI: 0.78-1.64 for culturally tailored studies). However, this may be due to the small sample size in the non-culturally-tailored subgroup, which only included three studies. Furthermore, there was no evidence of a significant moderating effect of cultural tailoring, as evidenced by a non-significant between-group heterogeneity statistic ($Q_b[1]=2.61$; $p=0.106$). Within-group heterogeneity remained moderate for interventions that were not culturally-tailored ($I^2=51.70$) and low ($I^2=0.0$) for culturally-tailored interventions, though this difference was likely a reflection of the difference in sample size between subgroups.

In a meta-regression model, cultural tailoring did not explain any of the between-study variance ($R^2=0.00$), indicating that it was not a significant source of heterogeneity.

Subgroup Analysis: Organizational/provider-level intervention strategies

When grouped by the presence of organizational- and/or provider-level strategies, the effect size for late pregnancy smoking abstinence remained significant for both groups, though it was larger for those interventions that did not include organizational/provider-level strategies strategies (n=12; RR=2.65; 95%CI: 1.92-3.65) than for those that did include such strategies (n=22; RR=1.29; 95%CI: 1.12-1.48). A significant between-group heterogeneity statistic ($Q_b[1]=16.34$; $p<0.001$) and

non-overlapping confidence intervals indicated that the difference between subgroups was significant. True heterogeneity of effect sizes was moderate for interventions that did include organizational/provider-level strategies (I^2=29.07) and low for interventions that did not include such strategies (I^2=6.66) (See **Table 1.1.7** for full results). This may be due to the fact that interventions including organizational/provider-level strategies typically employed more intervention strategies overall, and thus represented a more heterogeneous group of interventions.

In a meta-regression model, the use of organizational- or provider-level strategies accounted for 68% of the between-study variance in late-pregnancy smoking abstinence. The regression coefficient (b= -0..722; 95%CI: -1.08- -0.368; p=0.0001) indicated that the use of such strategies was associated with a significantly reduced likelihood of achieving late-pregnancy smoking abstinence compared to the reference group of interventions that did not use such strategies.

Subgroup Analysis: Low-SES

As seen in **Table 1.1.8,** when grouped by the socioeconomic status of participants (low-SES vs not low-SES), the effect size for late pregnancy smoking abstinence was significantly different between treatment and control groups only for low-SES samples (n=26; RR= 1.74; 95%CI: 1.43-2.13 for low-SES; n=xx; RR…..for high SES). Non-overlapping confidence intervals and a significant between-group Q-statistic (Q_b[1]=8.2; p=0.004) indicated that socioeconomic status was a significant moderating factor. Specifically, smoking cessation interventions appear to be more effective when delivered to low-SES samples. However, due to differences in the number of studies in each subgroup, these results should be interpreted with caution. Within-group

heterogeneity remained moderate for both low-SES (I^2=36.66) and non-low-SES (I^2=36.58) subgroups, indicating that other variables were contributing to the observed heterogeneity.

A meta-regression model revealed that 27% of the between-group variance was explained by socioeconomic status. Using non-low-SES as the reference group, interventions delivered to samples classified as low-SES were associated with a significantly increased likelihood of late-pregnancy smoking abstinence (b=0.387; 95%CI: 0.073-0.702; p=0.016). These results are in line with the results of the subgroup analysis.

While acknowledging that the small sample size for the non-low-SES subgroup (n=8) is a limitation, the results of this subgroup analysis are notable, given that previous research suggests that low-SES women are less likely to quit smoking during pregnancy.

Subgroup Analysis: High Psychosocial Risk

When grouped by psychosocial risk (high risk vs not high risk), the effect size for late pregnancy smoking abstinence remained significant for both high risk (n=20; RR=1.48; 95%CI: 1.19-1.84) and low-risk (n=14; RR=1.60; 95%CI: 1.24-2.06) samples. Overlapping confidence intervals and a non-significant between-group heterogeneity statistic (Q_b[1]=0.20, p=0.656) indicate that psychosocial risk was not a significant moderating factor. Within-group heterogeneity was low to moderate for the low-risk subgroup (I^2=36.00) and moderate for the high-risk (I^2=55.28) subgroup, indicating that additional variables were contributing to the observed variability in late-pregnancy smoking abstinence in both subgroups. (See **Table 1.1.9** for full results).

In a meta-regression model, psychosocial risk status did not account for any of the between-study variance (R^2=0.0), confirming the findings of the subgroup analysis.

Subgroup Analysis: Majority Minority Sample

When grouped by racial/ethnic composition of study samples, the effect size for late pregnancy smoking cessation remained significant for interventions delivered to majority minority samples (n=10; RR=1.46; 95% CI: 1.12-1.92) and those classified as non-majority-minority samples (n=24; RR=1.56; 95%CI: 1.27-1.92). Overlapping confidence intervals and a non-significant between-group heterogeneity statistic (Q_b[1]=0.147; p=-0.702) indicated that there was no significant moderating effect according to racial/ethnic composition. Within-group heterogeneity remained moderate for both subgroups, though it was lower for the majority-minority subgroup (I^2=31.69 compared to I^2=53.854). (See **Table 1.1.10** for full results).

In a meta-regression model, majority-minority status did not account for any of the between-study variance (R^2=0.0), confirming the findings of the subgroup analysis.

Subgroup Analysis: Health Status (Mental health or substance abuse disorder)

When grouped by the health status of participants, the effect size for late pregnancy smoking cessation was significantly different between control and treatment groups only for those interventions delivered to generally healthy samples (n=30; RR=1.56; 95%CI: 1.30-1.86). The effect size for interventions delivered specifically to samples with mental health or substance use disorders was not significant (n=4; RR=1.37; 95%CI: 0.861-2.17). However, the between-group heterogeneity statistic was non-significant (Q_b[1]=0.269; p=0.604) and the confidence intervals of the subgroups overlapped, indicating that the effectiveness of smoking cessation interventions did not

vary significantly according to the health status of participants. However, due to the small sample size in the 'unhealthy' subgroup, these results should be interpreted with caution. Within-group heterogeneity remained moderately high for interventions delivered to generally healthy samples (I^2=48.9) and those delivered to women with mental and/or substance abuse disorders (I^2=44.4), indicating that other variables were contributing to the observed heterogeneity in effect sizes. See **Table 1.1.11** for full results.

In a meta-regression model, health status did not account for any of the between-study variance (R^2=0.0), confirming the findings of the subgroup analysis.

Meta-Regression: Baseline Characteristics

In a univariate meta-regression model, baseline smoking (cigarettes/day) cessation explained 7% of the between-study variance in late pregnancy smoking abstinence. The regression coefficient for baseline smoking (b=0.03; 95%CI: -0.030-0.091; p=0.328) indicated a non-significant, positive association between cigarettes smoked per day at baseline and late pregnancy smoking abstinence. See **FIgure 1.1.12** for full results.

A second univariate model revealed that gestational age (in weeks) at baseline did not account for any of the between-study variance in late pregnancy smoking outcomes (R^2 analog = 0.0). The regression coefficient for gestational age (b= -0.017; 95%CI: -0.053-0.019; p=0.353) indicated a non-significant, negative association between gestational age and late-pregnancy smoking abstinence. See **Table 1.1.13** for full results.

Secondary Outcomes

Significant Reduction in Smoking

Four studies included a measure of significant reduction in smoking, defined as reducing baseline cigarette consumption by at least 50% by the last assessment before delivery (late pregnancy). This outcome was measured dichotomously. Data were based on self-reported smoking status. A random effects model revealed a significant difference between control and intervention groups (RR = 1.44; 95%CI = 1.21-1.70), such that intervention group participants were more likely to report significantly reducing cigarette consumption. (See **Table 1.2.0** for full results).

Early (0-6 months) Postpartum Smoking Abstinence

Eleven studies included measures of early postpartum smoking abstinence, defined as biochemically validated smoking abstinence measured after birth but before 6 months postpartum. The effect size for this time-point was significantly different between the control and intervention arms (RR=1.54; 95% CI: 1.16-2.03), such that intervention group participants were more likely to achieve smoking abstinence in the early postpartum period. (See **Table 1.2.1** for full results).

Late postpartum (past 6 months) smoking abstinence

Four studies included measures of late postpartum abstinence, defined as biochemically validated smoking abstinence at 6 months postpartum or later. A random effects model revealed a significant difference between treatment and control groups (RR=1.99; 95%CI: 1.07-3.69), with a greater likelihood of smoking abstinence among treatment group participants. (See **Table 1.2.2** for full results).

Low birthweight or very low birthweight delivery

Four studies included measures of low birthweight or very low birthweight deliveries among treatment and control group participants. A random effects model revealed a significant difference in the risk of low birthweight or very low birthweight between the treatment and control group, with a significantly lower risk among treatment group participants (RR= 0.377; 95%CI: 0.219-0.649). (See **Table 1.2.3** for full results).

NICU Admissions

Two studies included measures of NICU admissions among treatment and control group participants. A random effects model revealed a non-significant difference in the risk of NICU admissions between groups (RR=0.749; 95%CI: 0.469-1.20). (See **Table 1.2.4** for full results).

Preterm birth

Two studies included measures of preterm birth among treatment and control group participants. A random effects model revealed a significant difference in the risk of preterm birth between groups, with a significantly lower risk of preterm birth among treatment group participants (RR=0.434; 95%CI: 0.244-0.774). (See **Table 1.2.5** for full results).

Preterm birth or low birthweight/very low birthweight

A random effects model examining the risk of preterm birth *or* low birthweight/very low birthweight deliveries revealed a significant difference between treatment and control groups, with a significantly lower risk of experiencing either adverse outcome among treatment group participants (n=4; RR=0.401; 95%CI: 0.238-0.674). (See **Table 1.2.6** for full results).

Any serious adverse event

Two studies included measures of ‘any adverse event’, which was defined as any adverse perinatal outcome, including low birthweight, preterm birth, NICU admission, or fetal demise. The risk ratio was not significant for this comparison (RR=1.039; 95%CI: 0.396-2.72). (See **Table 1.2.7** for full results).

Discussion

The purpose of this chapter was to conduct a meta-analysis of high-quality published trials of prenatal smoking cessation interventions conducted among U.S. women, and to explore the study-level factors that influenced the effectiveness of interventions. The primary outcome of interest was late pregnancy smoking cessation, but additional outcomes including smoking reduction and perinatal health were also assessed when possible. The results of the meta-analysis also served as the basis for further analyses, as described in the next two chapters. The following section presents a summary of the main results, followed by a discussion of the limitations and implications of the review.

Summary of main results

Overall, 38 studies met the strict criteria for inclusion in the review, though the quality of intervention, evaluation, and reporting practices still varied significantly within the sample. About two-thirds of the sample was classified as ‘unclear risk of bias’ or ‘high risk of bias’, while one-third was classified as ‘low risk of bias’. However, risk of bias did not appear to significantly influence the primary intervention outcome of late-pregnancy smoking cessation, possibly because the strict inclusion criteria limited the sample to rigorous trials. Thus, while there was significant variation in study quality,

the baseline was high and much of the variability stemmed from reporting and specification of intervention content, rather than study design or implementation.[4]

Description of interventions

Most interventions included in the review targeted psychosocial factors at the individual- and interpersonal-levels of influence. Two trials included techniques specifically designed to address intervention implementation and dissemination (El-Mohandes et al., 2008; Pbert et al., 2004), one trial included a component to increase the sustainability of the intervention (Donatelle et al., 2000a), and one trial included extensive formative research in the community to increase intervention uptake and acceptability (Patten et al., 2012). While many important determinants of smoking are found at the individual- and interpersonal-levels of influence, the limited focus on higher levels of influence is notable and provides a potential avenue to increase effectiveness through more comprehensive, multi-level interventions.

Interventions in the review were characterized by significant within- and between-study variation in content, delivery, setting, and intensity. Intervention content was typically delivered through multiple modes of delivery, such as face-to-face counseling plus telephone calls and written materials. Similarly, most interventions included content delivered in multiple settings; most commonly, the primary intervention content was delivered within a clinic or hospital setting, with supplemental material delivered at home. Intervention intensity varied significantly, ranging from brief, single-session interventions to weekly sessions lasting 10 or more weeks. The design of many trials allowed for significant variation in intensity within the same intervention. For

[4] Upon final review of the dissertation in 2020, a secondary literature search was conducted to identify studies meeting our inclusion criteria, published after 2015. These studies are summarized in the overall discussion section of the paper.

example, there was often no limit (upper or lower) on the length of time spent on telephone calls to participants, so calls could range from a few minutes to 30 minutes or longer. While variation was present across all types of interventions, it was particularly notable in counseling- and social support-based interventions.

Counseling interventions were characterized by significant variation in delivery, content, and intensity. Within this category, interventions ranged from a brief, unstructured, one-time counseling session to structured cognitive behavioral therapy and motivational interviewing delivered across a span of several weeks or more. Even within trials, the intensity of counseling varied significantly. For example, in the trial conducted by Hennrikus and colleagues (2010), the length of the single counseling session ranged from 15 minutes to 75 minutes. As a result, there was significant variation encompassed within this subgroup, making it difficult to accurately summarize the results with one effect size. Additionally, counseling interventions were supplemented by other intervention content, including written material, videos, feedback, and/or follow-up calls or mailers. Typically, multiple types of supplemental materials and/or techniques were delivered alongside counseling. This variation makes it difficult to determine whether counseling was effective on its own, and which supplemental materials/techniques (if any) were associated with increased effectiveness. Counseling interventions were generally well accepted by participants, though dropouts and implementation fidelity were significant problems, especially for multi-session interventions. In one study, participants explicitly stated that they would have liked more face-to-face contact (Cinciripini, 2000). Counseling appeared to be more effective when delivered earlier in pregnancy, as noted by Donatelle and

colleagues (2006). This may be because motivation is higher earlier in pregnancy, or because women who have not experienced any known health consequences of smoking late in pregnancy may feel that it is safe to continue smoking for the duration of pregnancy.

Social support interventions were similar to counseling interventions in many ways. First, they tended to be less structured and were characterized by significant variation in intervention content, delivery, and intensity. Secondly, they were accompanied by multiple types of supplemental materials and techniques, including educational materials (e.g., brochures, booklets, and videos), counseling, instruction, and in one case, scrapbooking (Hennrikus, 2010). Social support interventions were often delivered via multiple modalities, such as face-to-face plus telephone. Similarly, they were often delivered in multiple settings, with some intervention content delivered at a clinic or other health setting, and other intervention content delivered at home. This resulted in significant between-study variability. In one trial, participants were even given the choice of in-home or clinic-based social support (Malchodi, 2003), resulting in significant within-study variation, as well. Other interventions in this category were very unstructured, allowing participants to control the dose of intervention. For example, Hartmann and colleagues (1995) provided additional support to participants who asked for it, including those in the control group. This was in addition to the services provided during routine prenatal care. However, the authors did not quantify this additional support, thus making it difficult to determine what type and how much additional support was provided and how this influenced intervention outcomes. Notably, most social support-based interventions did not measure perceived social support as an outcome

nor as a mediating variable. When it was measured, the results were mixed. Participants in the trial conducted by Hennrikus and colleagues (2010) reported increases in both positive *and* negative support behaviors. In that intervention, pregnant women reported significant increases in several negative social support behaviors including expressing doubt about the woman's ability to quit or stay smoke-free, expressing anger about the woman's smoking, criticizing the woman for smoking, and trying to evoke guilt about smoking. In the postpartum period, women reported increases in an even greater number of negative social support behaviors. In the trial conducted by McBride and colleagues (2004), perceived social support actually decreased over the study period.

Contingent rewards-based interventions were far more structured than any other type of intervention, with a predetermined schedule of check-ins and rewards. They also included fewer supplemental materials and techniques, though all included some form of written or verbal instructional, supportive, and/or educational component. All trials in this subgroup provided tangible rewards, typically in the form of retail gift cards or vouchers for diapers, baby food, or related supplies. Importantly, in the series of studies by Donatelle and colleagues (2000a; 2000b; 2000c), decreasing the value of the incentive was associated with smaller treatment effects, suggesting that the dose of incentive may influence effectiveness. In the same series of studies, treatment effects were similar when incentives were given to women-only compared to when they were given to women *and* a designated supporter. Thus, it may not be worth expending resources to give incentives to supporters, as incentives given to the pregnant woman appear to account for most or all of the treatment effect. In studies that compared

contingent rewards to non-contingent rewards (as opposed to a control group), the results suggested that the contingency component of the intervention contributed significantly to its effectiveness, which is consistent with the theorized mechanisms of action (Heil, 2008; Higgins et al., 2004; Higgins et al., unpublished; Higgins, 2014; Tuten, 2012). Contingent rewards-based interventions also included more face-to-face contact with deliverers, typically in a clinic or hospital setting. Additionally, while non-completion was still a problem in this subgroup, it was less problematic than in other categories of interventions, suggesting that the rewards and/or the structured format promoted compliance. This may have also contributed to the effectiveness of this group of interventions.

Meta-Analysis Results

The results of a random effects meta-analysis yielded a significant risk ratio for the primary outcome of late-pregnancy smoking cessation, such that women in the treatment groups were 1.53 times as likely to achieve smoking cessation before giving birth than women in the respective control groups. Several study-level variables emerged as potential moderators of intervention effectiveness. Not surprisingly, effect sizes varied according to the type of intervention (e.g., counseling, contingent rewards, health education, etc.). In subgroup analyses, the effect size for late pregnancy smoking cessation remained significant for two categories of intervention: contingent rewards and counseling. Interventions categorized as “contingent rewards” had a significantly larger effect size than any other category of intervention. Treatment-group participants in *contingent rewards* interventions were 2.82 times as likely to achieve late-pregnancy smoking abstinence than control group participants. In comparison, treatment-group

participants in *counseling* interventions were 1.3 times as likely to achieve late pregnancy smoking abstinence than their control group counterparts.

Effect sizes also varied according to the setting of the intervention, such that interventions delivered within the context of prenatal care appeared to be more effective than interventions delivered outside of routine prenatal care. One potential explanation for this finding is that women may be more likely to attend and adhere to smoking cessation programs when they do not require additional time or planning. As with most types of interventions, greater attendance and adherence during smoking cessation interventions is associated with better quit outcomes (Barker et al., 2004), so maximizing adherence and minimizing low-attendance and dropouts is an important consideration when designing interventions. It is also possible that the structured setting of routine prenatal care enhanced implementation fidelity, so women were more likely to receive the intervention as intended when it was delivered within the context of prenatal care. Additionally, since women tend to form trusting relationships with their prenatal care providers, it is possible that interventions delivered in the context of routine prenatal care are associated with positive interpersonal factors such as trust, open and honest communication, and social support. Furthermore, many interventions that were delivered within the context of routine prenatal care also included some component(s) delivered outside of the clinical setting. For example, many interventions provided women with written materials or scheduled follow-up phone calls and/or mailers to be delivered after the primary intervention and outside of the context of prenatal care. Thus, while the main intervention was delivered within a routine prenatal care setting, additional intervention components delivered at home or in the community may have

contributed to the observed effects. While further research is needed to isolate the context of the intervention from other factors, intervention planners may wish to consider implementing smoking cessation programs within routine prenatal care when possible. Additionally, intervention planners and evaluators should consider using qualitative research to explore pregnant women's views on attending and adhering to smoking cessation programs. For example, it may be beneficial to ask women if and why they would prefer interventions delivered within routine prenatal care, and if additional intervention content (delivered outside of routine prenatal care) is useful to them. Eliciting provider views may also yield important findings that could enhance implementation fidelity. While interventions that were delivered within the setting of routine care appeared to be more effective than those delivered in other settings, relying on existing staff to deliver the intervention may overburden prenatal care providers and clinic staff, potentially leading to lower implementation fidelity. In the study conducted by Kendrick and colleagues (1995), project staff reported that the use of existing clinic staff to deliver the intervention and collect data negatively impacted the intervention.

Intervention effects also varied by participant socioeconomic status, such that the effect size for interventions delivered to low-SES women was significantly larger than the effect size for non-low-SES participants. It is possible that this finding reflects differences in intervention type/content, as certain types of intervention may be more likely to be delivered to low-SES women. For example, of the nine studies that provided contingent rewards, eight were delivered to low-SES samples. Given that contingent rewards-based interventions were found to be the most effective category of intervention, the larger effect size for the low-SES group could reflect more effective

intervention content rather than variation in participant characteristics. Nevertheless, it is notable that the likelihood of achieving late-pregnancy smoking abstinence was greater among low-SES women, as previous studies suggest that low-SES women often have lower quit rates during pregnancy and are more likely to continue smoking throughout pregnancy (McLeod, 2004; Tong et al., 2009; Tong et al., 2013).

Intervention effects did not vary significantly by other intervention characteristics including level of contact intensity, type of intervention deliverer, use of cultural tailoring, or use of organizational/provider-level strategies. Similarly, intervention effects did not vary according to participants' psychosocial risk status or health status, baseline smoking habits (cigarettes/day), or gestational age, nor by the racial/ethnic composition of the study sample.

Baseline smoking habits, as measured by the number of cigarettes smoked per day at the first assessment, accounted for only 7% of the between-study variance. It is possible that other smoking-related characteristics, such as the number of previous quit attempts or length of time as a smoker, may account for additional variance in effect sizes for late pregnancy smoking cessation. However, in the current sample, lack of reporting and inconsistent reporting practices prohibited us from exploring these factors as sources of heterogeneity. Somewhat surprisingly, psychosocial risk status (high/low) did not account for any of the between-study variance in late pregnancy smoking cessation. However, this may be explained by the fact that pregnant smokers comprise a high psychosocial risk group of pregnant women in general, and thus most pregnant smokers could be considered high risk on this measure. Future studies should explore specific psychosocial risk factors as moderators of intervention effectiveness.

Additionally, further research could explore whether certain behavior change techniques are more effective for pregnant smokers with specific psychosocial risk factors. Similarly, it was unexpected to find that health status did not account for any of the between-study variance in late pregnancy smoking cessation. However, this may be explained by the way health status was measured. Due to sample size limitations, we measured health status by creating a variable to represent any mental or physical health disorder. Thus, we were unable to explore whether specific mental or physical health conditions were associated with the effectiveness of the intervention.

The finding that higher intensity interventions were no more effective than lower intensity interventions has important implications for reporting practices, intervention design, resource allocation, and even research ethics. Prior research on the association between intervention intensity and effectiveness has yielded mixed conclusions. While some evidence indicates that higher intensity interventions are more likely to be effective, other reviews have found no relationship between intervention intensity and smoking cessation outcomes among pregnant women (Chamberlain et al., 2014; Chapman & Wakefield, 2012; Naughton, Prevost, & Sutton, 2008). It is possible that the finding of no relationship in the current review stems from poor reporting practices and lack of standardized methods for describing the frequency, duration, and dosage of intervention delivery. Thus, an important step for future research will be to develop better and more consistent guidelines for reporting on intervention intensity, in an effort to improve evidence synthesis on the relationship between intervention intensity and intervention outcomes. If it is concluded that higher intervention intensity is not associated with better smoking cessation outcomes among pregnant women, this would

indicate that significant resources could be saved by designing less intensive interventions without sacrificing effectiveness. However, analyses of intervention outcomes at different levels of intensity would be needed to determine the optimal dose, frequency, and duration. Considering patterns of smoking cessation and relapse, it is also possible that higher intensity intervention contact is needed for a discrete period of time as women initially quit smoking and deal with the acute effects of nicotine withdrawal, after which only low-intensity intervention is needed to sustain cessation and prevent relapse. Higher intensity intervention may be needed again during the postpartum period, when the risk of relapse increases. Regarding the ethics of research participation, asking women to take part in intensive interventions may be an avoidable and undue burden if greater intensity is not associated with improved outcomes.

Secondary Outcomes

Interventions in this review also yielded promising (significant) results for many secondary outcomes of interest, including additional measures of smoking behavior as well as perinatal outcomes. Specifically, treatment group participants were 1.44 times as likely as control group participants to significantly reduce (by at least 50%) their cigarette consumption, 1.54 times as likely to be smoke free in the early postpartum period, and 1.99 times as likely to be smoke free in the late postpartum period. Although complete smoking cessation during pregnancy is the optimal outcome, reduction in smoking is still associated with improved health outcomes for mother and fetus. Reducing cigarette consumption is particularly important during critical periods of fetal development, when nicotine and other toxic substances can restrict fetal oxygen supply, reduce nutrient absorption, and contribute to problems with organ development

(Crawford, 2008; Maritz, Morley, & Harding, 2005; Morales-Suarez-Varela, 2006). Given the significant harms associated with smoking during pregnancy, it has been suggested that significant reduction rates should be considered as a realistic harm-reduction outcome for heavy smokers (Windsor, Li, Boyd, & Hartman, 1999). Additionally, decreasing cigarette consumption can help reduce nicotine addiction and thus attenuate the symptoms of withdrawal during future quit attempts, which may increase the likelihood of successfully quitting. Pregnant women have been shown to experience more severe nicotine withdrawal due to accelerated nicotine metabolism, which is thought to make it harder to quit smoking during pregnancy (Dempsey, Jacob, & Benowitz, 2002). Thus, encouraging continuing smokers to reduce their nicotine consumption may help to address the unique physiological processes that make smoking cessation more challenging during pregnancy, while also mitigating some of the health risks associated with continued smoking.

The finding that women who participated in smoking cessation interventions had an increased likelihood of smoking abstinence in the postpartum period is especially notable in light of the fact that less than one-third of pregnant women who quit smoking remain abstinent one year after giving birth, with the majority relapsing within the first six months (Colman & Joyce, 2003; Fang, 2004). Maintaining smoking cessation during the postpartum period is particularly important for women who breastfeed, as nicotine and other chemicals from cigarettes can be passed along through breast milk (Abel, 1980). Additionally, women who smoke tend to stop breastfeeding earlier than women who don’t smoke (Amir, 2001; Scott, Binns, Oddy, & Graham, 2006). Heavy smokers are up to 2.5 times as likely as non-smokers to wean their infants off breastmilk before 10

weeks, thereby reducing infant exposure to the benefits of breastfeeding (Liu, Rosenberg, & Sandoval, 2006). Moreover, maintaining abstinence during the postpartum period (and beyond) reduces environmental smoke exposure, which is beneficial for new mothers and infants, as well as other family members (Yang, 2010). This effect was largest for contingent rewards-based interventions, indicating that the effects of such interventions may extend beyond the period during which rewards are offered, although the mechanisms of action in these maintenance effects remain somewhat unclear.

The results also revealed that smoking cessation interventions reduced the risk of poor two very common perinatal health outcomes: low birthweight and preterm birth. Specifically, treatment group participants had 73% less risk of delivering a low birthweight or very low birthweight infant and 67% less risk of preterm birth compared to control group participants. According to the U.S. Department of Health and Human Services, smoking may account for 20-30% of low birthweight deliveries and nearly 10% of preterm deliveries (Andres & Day, 2000; USDHHS, 2001). Unlike many other causes of preterm birth and low birthweight deliveries, smoking is a modifiable risk factor that can be addressed through behavior change interventions.

An unexpected positive finding from the review was the frequency with which published trials reported at least some results from process evaluations, which provided an indicator of implementation fidelity and completion rates. Additionally, several published reports included measures of intervention acceptability or usefulness, and a few incorporated the results of these evaluations into intervention design. Overall, women reported a high degree of acceptance across all types of intervention, although

one study found that low participation rates were due to concerns about stigma (Patten et al., 2012). In that trial, which focused on Alaska Native women, reports from women who did not participate indicated that there was a perception of stigma in the community associated with attending the smoking cessation program. Similar findings have been reported previously, particularly in marginalized and/or socially disadvantaged populations (Burgess et al., 2009; Greaves & Tungohan, 2007). Given that the intervention was conducted in a rural community of Alaska Natives in the Yukon Delta, perceptions of stigma may have been magnified by the small and interconnected social networks in which women were embedded. Women who participated in the program did so with the understanding that most members of their community would find out about their attendance, and thus would know that they were smoking during pregnancy. Future studies in similar settings should consider employing community-level strategies to address attitudes and stigma surrounding smoking cessation among pregnant women. Along similar lines, women in the study conducted by Hennrikus and colleagues (2010) reported that participation in the intervention resulted in an increase in negative support behaviors, including criticism and attempts to evoke guilt about smoking. As awareness of the harms of smoking has increased and the prevalence of smoking has decreased, stigma has increased for those who continue to smoke. An unfortunate consequence of this stigma is that some smokers may avoid seeking help. This may be especially true for pregnant women, who also face greater stigma, shame, and guilt related to smoking (Burgess et al., 2009; Greaves & Tungohan, 2007). Taken together, these findings point to a need for broader, community- and societal-level

messaging campaigns aimed at reshaping attitudes and encouraging positive responses to smoking cessation during pregnancy.

While acceptability of interventions was generally high, non-completion was a common problem. Across nearly all interventions in the review, dropouts increased as time went on. For interventions that included multiple sessions, participation was typically lower in later sessions. Compared to other risk factors, such as domestic violence and depression, women who smoked had the highest non-completion rates in the trial conducted by El-Mohandes and colleagues (2008). These findings suggest that intervention planners may need to develop better strategies to encourage sustained participation. Incentive-based interventions generally had higher rates of completion, suggesting that rewards may promote participation as well as smoking cessation. Additionally, it is likely that the greater completion rates in incentive-based interventions contributed to improved intervention outcomes. Given the challenge of non-completion despite high rates of acceptability, future studies should explore the barriers to sustained participation through qualitative research and attempts to follow up with women who drop out of interventions prematurely.

Limitations

Like every meta-analysis, study selection was based on search protocols and inclusion/exclusion criteria set by the author, and by the author's judgment of whether or not a particular study met those criteria. Therefore, the potential for bias exists in the selection procedures. However, we tried to minimize bias during study selection by using existing search protocols and inclusion/exclusion criteria to inform our own procedures, and by pre-specifying inclusion and exclusion criteria. Additionally,

decisions were based on the judgments of two coders who independently applied the inclusion/exclusion criteria and, when needed, discussed disagreements before reaching a determination. Bias and error may also be introduced during the process of data extraction and coding. To minimize this risk, two independent reviewers performed coding and data extraction using standardized forms and explicit instructions, as well as direct discussion when needed. Before applying the coding scheme to the studies included in the review, the coders tested the forms on a selection of related studies and revised them where confusion or difficulty were noted.

The generalizability of our findings is also limited by the study sample, which focused on trials conducted in the United States. Evidence from this review is not generalizable to developing nations due to significant differences on key variables across all levels of influence, including individual beliefs and attitudes (about pregnancy, motherhood, smoking, health, etc.), social factors (such as the woman's role in the family), cultural and societal issues (such as gender roles and gender inequality, collectivist versus individualistic worldviews, smoking prevalence, and religious influences), and policy-related factors (such as health care policies and smoking-related laws). Given that smoking and poor maternal and fetal health outcomes are more prevalent in developing nations, there is a pressing need for research on effective approaches to promoting smoking cessation among pregnant women in lower-income nations. Furthermore, even though the U.S. is comparable to other high-income nations on many key variables, there are differences in the quality and structure of prenatal care (and the health care system more broadly), cultural attitudes about smoking, beliefs about pregnancy, and other related factors that may make it difficult to translate

evidence from the U.S. to other high-income nations. Thus, these findings may not be not be generalizable to settings outside of the United States.

Given the goals of this review, we used strict search terms and inclusion criteria to limit the sample to high-quality, randomized controlled trials of behavioral and psychosocial interventions to promote smoking cessation among pregnant women in the United States. While this allowed us to examine the most rigorous evidence possible, it may also limit the generalizability of the findings, as most prenatal smoking cessation interventions are not randomized controlled trials. Furthermore, while our use of a standardized tool to assess methodological rigor reduced bias and limited the sample to high-quality studies, it also screened out potentially relevant but less rigorous studies. Some researchers have suggested that inclusion criteria for evidence syntheses involving theory-driven research questions and hypotheses should be based on the relevance to the research question, rather than the methodological quality (Perski, Blandford, West, & Michie, 2017).

Additionally, meta-analyses of published reports are susceptible to publication bias due to the tendency for journals to accept positive findings and reject negative findings. However, an assessment of the forest plot and two different markers of publication bias (classic fail-safe N and Orwin's fail-safe N) indicated that the results of this review were not likely influenced by publication bias. The results of the two fail-safe N formulas, when applied to this study sample, indicated that it would take anywhere from 131 to 432 missing (unpublished) studies with significant findings to reduce the effect size to a non-significant level.

Additionally, given that both of the cluster randomized trials in the review were incentive-based interventions, it is possible that leaving them out of the primary quantitative evidence synthesis could influence the interpretation of results regarding the effectiveness of incentive-based interventions. The context of the intervention, including the physical environment, organizational factors, and delivery-related characteristics may be more important in cluster randomized trials than in individual trials, so it is possible that leaving these two trials out may also influence the interpretation of moderator analyses (Donner & Klar, 2004).

Inconsistent and/or incomplete reporting and measurement were significant limitations that influenced the review in several key areas. For many important participant characteristics, such as age, income, and education level, variation in reporting practices made study-to-study comparisons impossible without modifications resulting in a loss of data through dichotomization or grouping based on scores on continuous variables. For example, some studies reported education level as a continuous variable reflecting the average years of education completed by participants, while other studies reported the percentage of participants with *12 years of education or less*, and others reported the percentage of participants with *less than 12 years of education*. Reporting on income was characterized by similar problems: most studies that included this variable reported the percentage of participants within certain income categories (e.g., less than $20,000/yr; $20,000 to $39,999; $40,000 to $59,999; $60,000 or more), but different studies used different income categories. To overcome these challenges, we chose to dichotomize data where appropriate and, in some circumstances, we created new variables (e.g., high psychosocial risk) to best describe

the sample based on the available data. However, this still resulted in a loss of data and less precise measurements, which may explain, at least in part, why many study-level variables failed to explain much or any between-study heterogeneity in effect sizes. Notably, we found that few studies included an assessment of whether the pregnancy was intended or unintended. This could have important implications for a number of reasons. First, women who plan their pregnancies are more likely to consider quitting before becoming pregnant; thus, women with intended pregnancies who are still smoking upon becoming pregnant may be more addicted or resistant to quitting. Secondly, unintended pregnancies are often characterized by high levels of stress and other psychosocial risk factors that may make smoking cessation more difficult. Furthermore, unplanned pregnancies may be accompanied by mixed emotions about becoming a mother, which may further complicate smoking cessation efforts. While efforts to standardize reporting have begun to increase consistency in published reports of intervention content, similar efforts are needed to improve reporting on participant and study delivery characteristics.

Similarly, it was sometimes difficult to categorize intervention content because of overlapping characteristics. For example, some social support interventions included a counseling component, and many counseling interventions included some form of social support. We coded interventions based on the *main* strategy, but in some cases, there was very little distinction between counseling and social support interventions. This was also true for intervention delivery and setting, as many face-to-face interventions also included some other form of contact, such as telephone or electronic content, and most interventions delivered in a clinic or hospital setting also included an at-home (e.g.,

telephone, mailer, electronic) component. Additionally, many interventions included multiple deliverers, such as a trained mental health professional for the counseling component and trained study staff for follow-up phone calls. Thus, while we coded for the main deliverer, delivery format, and setting, most interventions actually fell into multiple categories. This resulted in a significant degree of within-category variation, thus making it more difficult to determine whether (and which) factors related to delivery or setting influenced intervention effectiveness. While breaking down the study sample into more defined subgroups would have reduced within-group variability, the sample size and distribution of moderator variables did not allow for this.

Upon completion of the study, another search was performed to identify additional randomized controlled trials that may have been published during the time that this review was underway. Several additional trials that would meet inclusion criteria were identified, including three randomized controlled trials of text messaging programs for pregnant smokers (Abroms, Johnson, Leavitt, Cleary, Bushar, Brandon, & Chiang, 2017; Abroms, Chiang, Macherelli, Leavitt, & Montgomery, 2017; Forinash, Yancey, Chamness, Koerner, Intenso, et al., 2018), a telephone counseling intervention (Cummins, Tedeschi, Anderson, & Zhu, 2016), a biomarker feedback-based intervention (Patten et al., 2019), and a trial of behavioral counseling supplemented by bupropion (Nanovskaya, Oncken, Fokina, Feinn, Clark, et al., 2017). After the nine-session intervention, biochemically validated smoking abstinence was significantly greater in the treatment group compared to the control group in the telephone counseling intervention (Cummins et al., 2016), but not in any of the other trials. At the end of pregnancy, 38.8% of participants in the treatment group had achieved

cotinine-confirmed abstinence, compared to 22.5% of control group participants. While it is possible that these trials may alter the results of the meta-analysis, it is unlikely that they would significantly change the major conclusions, particularly with regards to incentive-based interventions. Additionally, given the lack of strong theoretical foundations in these studies, it seems unlikely that they would significantly impact the results of the theory-based components of this review.

Implications & Future Directions

The results of the meta-analysis provide evidence that interventions targeting psychosocial factors can promote smoking cessation during pregnancy among a diverse group of women, and that these effects can be sustained through the postpartum period. Importantly, smoking cessation interventions also reduced the risk of several adverse perinatal outcomes, though the mechanism of action for this effect is unknown. The findings also raise several key questions and provide useful insight to guide future research.

While contingent rewards were found to be the most effective type of intervention, there are still many unanswered questions about the mechanisms through which contingent rewards influence behavior and the conditions under which contingent rewards are most effective. Future research should explore how changes in the value and schedule of rewards influence outcomes, and whether contingent rewards are more effective for certain subgroups of women. Sustaining the cost of incentive-based interventions is a common concern, especially in certain settings such as local clinics. To address this issue, Donatelle and colleagues (2000a) purchased vouchers with funds donated by health care organizations, local businesses, and foundations.

Organizations were motivated to donate based on the social, health, and economic benefits of being smoke-free during pregnancy. For example, health organizations and insurers recognize that reducing smoking during pregnancy would reduce the burden of poor maternal and fetal outcomes in local clinics and emergency rooms. Donatelle and colleagues (2000a) suggest that this rationale could be used in future trials to elicit support from community service providers and health insurers, which in turn would increase the sustainability of incentive-based interventions. The feasibility of this approach should be explored in future studies.

The review also identified gaps between research and practice, particularly in the area of social support. While social support appears to be an important factor influencing pregnant women's smoking behavior, effectively increasing positive forms social support without also increasing negative forms of support may be challenging. This is an area that should be explored further. Future research should also explore factors that may influence the effectiveness of social support-based interventions, such as the type of supporter (e.g., family members, partner, peer, etc.), the intensity of the support, and the characteristics of the participants (e.g., SES, baseline levels of social support, parity, etc). Additionally, there is a need for meta-analyses examining how different types of social support (e.g., instrumental vs. emotional) are differentially associated with intervention effectiveness. It may be that there are interactions among these various factors, such that certain types of support or supporters may be more effective for certain women. Additionally, some women may benefit from more intense social support interventions, while the intensity of support may be less important for women with existing social support networks.

There is also a need for further research exploring higher-level (i.e., beyond the individual) factors such as provider training and system-level changes (e.g., implementing new types of record-keeping practices). In this review, provider- and/or organizational-level strategies were negatively associated with the effectiveness of the intervention. However, as noted previously, this may be explained by other factors such as the heterogeneity of intervention-types within this subgroup of interventions. Due to sample size constraints, we were unable to explore how specific types of provider training or organizational change were associated with effectiveness. This should be investigated further in future studies, as it seems likely that different types of strategies may be differentially associated with effectiveness.

Given the heterogeneity of intervention content both between and within subgroups of intervention type, there is a clear need for better systems of reporting, classification, and measurement. The overlap between intervention categories (e.g., social support and counseling interventions) makes it difficult to reliably classify interventions for the purpose of evidence synthesis, which in turns limits our ability to identify which approaches are most effective for promoting smoking cessation during pregnancy. Similarly, inconsistent and limited measurement and reporting of psychosocial outcomes makes it difficult to determine *why* interventions were effective. This points to a need to further specify intervention content, as outlined in the third aim of this project.

CHAPTER 4

Specific Aim 2

A large body of evidence demonstrates that theory-based interventions, or those that target theoretical mechanisms of behavior change, are more likely to be effective than non-theory-based interventions (Michie et al., 2008). Theory-based interventions specify an explicit causal pathway(s) involved in behavior change, which is what distinguishes theory-*based* interventions from 'theory-influenced' or 'theory-inspired' interventions (Michie & Abraham, 2004; Michie et al., 2008). As such, theory-based interventions provide the basis for theory-based evaluations testing the overall effectiveness of an intervention as well as the hypothesized underlying causal mechanisms, thereby greatly increasing the knowledge gleaned from such an analysis, and providing a much more thorough understanding of what works, including how the effects vary by population, context, and behavior (Michie & Abraham, 2004; Michie et al., 2008; Michie et al., 2009b).

Health Behavior Theory

The most successful public health initiatives are based on a thorough understanding of health behaviors and the context in which they occur. Theory provides a unifying framework for describing and understanding these factors and the relationships among them. In the field of social and behavioral sciences, the term 'theory' is generally understood as "a set of interrelated concepts, definitions, and propositions that present a systematic view of events or situations by specifying relations among variables, in order to explain and predict the events or situations" (Glanz, Rimer, & Viswanath, 2008, p.26). More specifically, "behavioral theories are composed of interrelated propositions,

based on stated assumptions that tie selected constructs together and create a parsimonious system for explaining and predicting human behavior" (DiClemente, Crosby, & Kegler, 2002, p.3). Thus, theories fulfill three primary functions:

1) Description: Theories provide a standardized approach to describing (and therefore understanding) the phenomenon of interest, "so that others can repeat [the] description with a high degree of agreement" (Denzin, 1970, p. 31).
2) Explanation: The explanatory nature of theories refers to "the construction of a system of interrelated propositions that permits the scientist to 'make sense' out of the events observed" (Denzin, 1970, p. 31).
3) Prediction: In addition to describing and explaining "*why* a given set of variables occurs together", theories also enable scientists to predict the future relationship(s) among these variables (Denzin, 1970, p. 31).

Theories of behavior change draw from a broad range of academic disciplines including psychology, sociology, communications, anthropology, marketing, economics, and more. A wide variety of approaches are included under this umbrella, from broad ecological models encompassing multiple levels of influence, to individual-level theories focusing on specific psychosocial processes such as risk perception, motivation, or readiness for change (Glanz & Bishop, 2010; Glanz, Rimer, & Viswanath, 2008). Among the most frequently used theories are the Health Belief Model (HBM, Rosenstock, 1974), the Theory of Planned Behavior (TPB, Ajzen, 1991), the Transtheoretical Model (TTM, Prochaska, Johnson, & Lee, 2009), Social Cognitive Theory (SCT, Bandura, 1997), Self-Determination Theory (SDT; Deci & Ryan, 2000), Protection Motivation

Theory (PMT; Rogers, 1983), and the Social Ecological Model (McLeroy, Bibeau, Steckler, & Glanz, 1988).

Optimally, the selection of a specific theory or theories to guide intervention design and evaluation should be guided by evidence. However, research indicates that the popularity of a theory is not necessarily associated with its foundation of empirical support. For example, Prochaska and DiClemente's (1983) transtheoretical model (TTM), which focuses on stages of readiness for change, is one of the most widely-used theories in smoking cessation research and practice (Sutton, 2000). However, the stages of the TTM have been described as "arbitrary" (Sutton, 2000, p. 209-211) and several reviews of stage-based interventions have concluded that the psychological processes underpinning the TTM are not supported by the available evidence (Bandura, 1998; Sutton, 2000; Weinstein et al., 1998). Thus, despite the popularity of the TTM, empirical support for the theory is quite limited.

Uses of Theory in Intervention Design, Implementation, & Evaluation

The use of theory in intervention design, implementation, and evaluation is advantageous for several reasons (Michie, Johnston, Francis, Hardeman, & Eccles, 2008). First, theories contribute to the effectiveness of interventions by specifying the causal determinants of behavior and behavior change. According to the tenets of behavioral theories, changing causal determinants of behavior (i.e., constructs) will promote behavior change (Hardeman et al., 2005; Michie et al., 2008). Thus, theory can be used to identify appropriate constructs to target in behavior change interventions. Second, the use of theory in intervention design and evaluation provides a framework for data collection and facilitates the accumulation and synthesis of evidence across a

variety of contexts, populations, and behaviors. Third, theory provides a mechanism for understanding *why* interventions are effective (or ineffective) and *how* behavior change techniques influence behavior, which in turn provides valuable insight for future intervention design and for the development and refinement of behavior change theories.

Evidence suggests that theory-based interventions are more effective in achieving health behavior change than interventions that do not utilize a theoretical foundation (Noar & Zimmerman, 2005). However, although more health behavior interventions reference theory now than in previous decades (Glanz, Rimer, & Lewis, 2002), a significant proportion of published interventions still make no reference to a theoretical basis (Albarracin et al., 2005; Davies, Walker, & Grimshaw, 2010; Hardeman, Johnston, Johnston, Bonetti, Wareham, & Kinmonth, 2002). In one review of the literature, Painter and colleagues (2008) found that from 2000 to 2005, theory was applied in only about one-third of published health behavior research. In another review, Grimshaw and colleagues synthesized the evidence from over 235 randomized controlled trials designed to improve the dissemination and implementation of evidence-based practice guidelines for health professionals (Grimshaw, Thomas, MacLennan, Fraser, Ramsay, et al., 2007). While the review found that interventions to improve implementation and dissemination were moderately successful, the authors noted that very few studies utilized a theoretical framework for intervention design and/or evaluation. As such, the investigators were unable to identify the processes underlying effective interventions and could not provide evidence-based guidelines for the design of new interventions to be delivered in different contexts, populations, and/or

medical practice areas. In a separate review of the same 235 trials, the investigators applied a coding scheme to classify the use of theory according to both type of use (explicitly theory based, some conceptual basis, and theoretical construct used) and stage of use (choice/design of intervention, process/mediators/moderators, and post hoc/explanation) (Davies, Walker, & Grimshaw, 2010). The review found that just 22.5% (n=53) of the trials utilized theory, and an additional 4.3% (n=10) used individual constructs from theories. The remaining 172 trials did not use theory or theoretical constructs. Of the 53 studies that used theory, the majority (n=42) used only one theory. When theory was used, it was almost always employed during the intervention choice/design stage (n=49). Very few studies utilized theory for process/mediator/moderator analyses (n=7) or for post hoc explanations (n=10). In the studies that utilized individual constructs from theories, all of them did so in process/mediator/moderator analyses, although the authors noted that very few of these studies actually performed statistical tests to analyze the mediating or moderating effects of the theoretical constructs. Furthermore, the rationale for why specific theories and/or constructs were used was not apparent in the majority of studies, and the quality of reporting on the use of theory was judged to be poor. Similarly, Painter and colleagues (2008) concluded that even when theory is applied in health behavior research, it is rarely used to its full potential. Specifically, among the roughly 30% of studies in their review that did use theory, a very small proportion employed rigorous methods such as theory testing (3.6%) or theory building (9.4%). Evidence also indicates that significant discrepancies exist between reported theory-use and actual application of theory. For example, Dombrowski, Sniehotta, Avenel, and Coyne (2007)

found that although 44% (n=34) of the trials in their review reported a theoretical basis for intervention development, none of these studies explained how theory was actually used to develop the intervention. Additionally, according to Noar and Zimmerman's (2005) review of 19 theory-testing studies (i.e. studies that compared two or more health behavior theories), the majority of research in this area has methodological weaknesses that greatly limit the potential for advancing the literature and state of knowledge on health behavior theory. For example, even when theory is applied to intervention development, it is often used only as a loose framework, and rarely used in its entirety. Other limitations in the literature on applied health behavior theory include insufficient explanations of the processes and criteria researchers use to select theories/theoretical constructs (i.e., rationale for choosing one theory/set of constructs over others), failure to explicate the links between behavior change techniques and the behavioral determinant(s) they target, inconsistent and/or poorly operationalized definitions of theoretical constructs, and inconsistent methods of measurement (e.g., wide variation in the methods, instruments, & design used to measure theoretical constructs) (Baranowski, Lin, Wetter, Resnicow, & Hearn, 1997).

An additional limitation stems from the methods employed to evaluate the use of theory. Many systematic reviews of health behavior interventions consider an intervention to be theory-based if the published report mentions a theory or theories in the context of intervention design (Michie & Prestwich, 2010). Often, reviews report on the use of theory using a simple categorical (Y/N) outcome (Ammerman et al., 2002), without evaluating how theory was used, at what stage(s), or to what extent. For example, Albarracin and colleagues (2005) conducted a meta-analytic review to

examine the impact of theory-use on the effectiveness of HIV-prevention interventions, and found that the use of theory was associated with a greater degree of behavior change (Albarracin et al., 2005). However, this finding was based only on reported use of theory (Yes/No), rather than the actual application of theory and the extent that theory was used to develop the interventions. As a result, many evaluations fail to distinguish between different uses of theory, and may conflate theory-based and theory-inspired interventions. This significantly limits the potential to perform theory-testing research and to accumulate detailed evidence on the use of theory, which in turn limits contributions to theory-building and refinement. As such, the specific associations between the use of theory and the effectiveness of interventions is not well understood, as there is insufficient evidence to determine how and when the use of theory contributes most to intervention effectiveness (Michie & Prestwich, 2010).

To address some of these limitations and advance the state of research on behavior change theory, Michie and Prestwich (2010) developed the first comprehensive guide for systematically coding reported use of theory in intervention design. The 19-item coding scheme specifies whether theory or theoretical constructs were mentioned, whether theory was used to directly inform intervention design via targeting of theoretical constructs, how theory was used to indirectly influence intervention design via participant selection or delivery to different groups of participants (tailoring), whether relevant theoretical constructs were measured, whether theory was tested to examine the association between theoretical constructs and outcomes (i.e., did changes in theoretical constructs explain and/or mediate intervention effects), and whether theory was refined based on study outcomes. As such, the coding scheme

specifies three main pathways through which the use of theory can influence intervention effectiveness: 1) by selecting specific behavior change techniques or combinations of techniques to target specific theoretical constructs; 2) by informing the selection of participants who are likely to benefit from the intervention; and 3) by tailoring the intervention to individuals based on theory-relevant characteristics. Some theory-based interventions may utilize theory for all of three purposes, while others may only apply theory to one or two of the potential pathways. While evidence is limited, Prestwich and colleagues (2014) hypothesize that interventions that apply theory more extensively in these domains may be more effective than those which apply theory less extensively.

Theory-based research may offer a promising approach to improving our understanding of the mechanisms by which prenatal smoking cessation interventions lead to changes in smoking behaviors, and, eventually, to developing more effective interventions informed by the evidence linking specific behavior change techniques with theoretical mechanisms of change. ***The promise of theory-based research informed the second aim of this project, which is to evaluate the use of theory in smoking cessation programs, as specified below***:

Aim 2: To evaluate the use of behavior change theory in prenatal smoking cessation interventions.

1) **Sub-aim 2a:** To assess the use of theory as a guiding framework in prenatal smoking cessation interventions, using Michie & Prestwich's (2010) coding scheme for evaluating the extent to which an intervention is theory-based.

2) **Sub-aim 2b:** To determine whether theory-based interventions are more effective at promoting smoking cessation among pregnant women than non-theory-based interventions by conducting meta-analyses on both types and comparing the pooled effect sizes.

Aim 2: Methods

Inclusion/Exclusion Criteria

Studies for this review were derived from the meta-analysis conducted in the first step of this project. In addition to the inclusion/exclusion criteria specified under Aim 1, studies for Aim 2 must have also contained an adequate description and measure of at least one theoretical construct or theory, where adequate is defined as any description of a theoretical construct (or theory) that provides enough detail and clarity for the reviewers to identify it as a distinct, not overlapping construct (or theory). If the study included a measure of a theoretical construct, the following minimum reporting requirements must be also be met:

1. Continuous outcomes:
    a. Means & SD's (Mean, SD, & N of Intervention and Control Groups)
    b. Means & SE's (Mean, SE, & N of Intervention and Control Groups)
    c. Means & Full Sample Size SD (SD of Full Sample; Mean & N of Intervention & Control Groups)
    d. t-test (t-value; N of Intervention & Control Groups).
    e. F-test (F-test statistic; N of Intervention & Control Groups)

f. Standardized & Unstandardized Regression Coefficient (B; SD of DV; N of Intervention & Control Groups

2. Dichotomous outcomes:

a. 2X2 Frequency Table (#Even & #Non-Event for Intervention & Control Groups)

b. Binary Proportions (Proportion w/Event; N for Intervention & Control Groups)

c. Chi Square and Marginal Distributions (X^2 statistic; Proportion of full sample w/Event; N of Intervention & Control Groups)

d. Standardized Mean Difference (d)

Measures

The 19-item Theory Coding Scheme (TCS; Michie & Prestwich, 2010) was used to code for reported theory use in the development and evaluation of interventions (see Appendix B for full coding scheme). As mentioned above, the TCS classifies theory-use into three main categories, according to function: 1) Selecting specific behavior change techniques or combinations of techniques to target specific theoretical constructs; 2) Informing the selection of participants who are likely to benefit from the intervention; and 3) Tailoring the intervention to individuals based on theory-relevant characteristics. The TCS also includes items that assess whether or not the published study mentions a theory or theoretical construct; whether the intervention was based on a single theory; whether theory-relevant constructs were measured and, if so, how reliable the measures were; whether the intervention led to significant change in at least one relevant theoretical construct (compared to control group); whether mediational

analyses were conducted and, if so, whether the mediator (or a change in the mediator) predicted the dependent variable (or a change in the DV); whether results were discussed in relation to theory; whether the study provided support for the theory or, alternatively, refuted the theory (by changing behavior without changes in theory-relevant constructs); and whether the results were used to refine theory by adding or removing constructs, or specifying that the theoretical pathways of change should be changed.

Items on the TCS are coded categorically (Yes/No/Don't Know) and demonstrated substantial agreement during initial development and validation (kappa $\geq$ 0.70 for 18/19 items; kappa = 0.64 for item 19d) (Michie & Prestwich, 2010). In addition to categorical codes, the TCS also calls for recording the name of the theory or theories mentioned in the relevant reference document.

Coding

Two trained coders independently applied the TCS to a subset of 10 studies to establish intercoder reliability, using Cohen's kappa (*k*) coefficient to assess agreement between coders (Cohen, 1960). Cohen's kappa is considered to be a stronger measure of agreement than simple percent agreement (i.e., the number of agreement scores divided by the total number of scores), as it accounts for the probability of agreement occurring by chance (McHugh, 2012). Kappa values can range from -1.0 to 1.0, with values between 0.61 and 0.80 reflecting substantial agreement, and values of 0.80 to 1.0 reflecting nearly perfect agreement between coders (Landis & Koch, 1977). Disagreements between coders were resolved through discussion and further examination of the studies and item content.

Scoring

Items on the TCS can be treated individually as well as grouped together to form composite measures reflecting the extent and function of their use. In this analysis, items were analyzed individually and, in some cases, composite measures were created to reflect specific uses of theory. Six composite measures were created based on the scoring criteria developed by Prestwich et al. (2014). The measures reflect the following:

1) **Was theory mentioned?**

   Three items on the TCS reflect whether theory and/or theoretical predictors of behavior were explicitly mentioned. Item 1 assessed whether the study mentioned a theory, even if theory was not used to inform the intervention. Item 2 assessed whether theoretical predictors of smoking behavior were explicitly mentioned (and also targeted). Item 3 assessed whether the intervention was based on single theory (rather than multiple theories or a combination of theoretical predictors). A total score was calculated by summing the scores of the three individual items, where 'yes' = 1 and 'no' = 0. Thus, total scores for this category ranged from 0 (no mention of theory or theoretical predictors) to 3 (optimal use of theory).

2) **Were relevant theoretical constructs targeted?**

   Six items on the TCS reflect whether relevant theoretical constructs were targeted in the intervention. Item 5 assessed whether intervention techniques were based on a theory, theoretical predictor, or combination of theories and/or predictors. Items 7-11 examined the extent to which the intervention targeted specific theory-relevant constructs. Items 7 and 10 reflect optimal use of theory, indicating that all

intervention techniques are linked to a theory-relevant predictor (item 7) and all theory-relevant predictors are associated with a specific intervention technique (item 10). Items 8, 9, and 11 reflect less optimal use of theory, indicating an indirect link between intervention techniques and theoretical constructs/predictors (and vice-versa). A total score was calculated by summing the scores on item 5 (“yes” = 1; “no” = 0), items 7-9, and items 10-11. Studies coded “yes” on item 7 were given a score of 3; studies coded as “yes” on item 8 were given a score of 2; studies coded as “yes” on item 9 (“Group of techniques are linked to a group of constructs”) were given a score of 1; and studies that were coded “no” on items 7-9 were given a score of 0. For item 10, studies coded as “yes” were given a score of 2. Studies coded as “yes” on item 11 were given a score of 1. Studies coded as “no” on items 10 and 11 were given a score of 0. Thus, total scores for this category ranged from 0 (no theory use) to 5 (optimal use of theory).

3) **Was theory used to select participants or tailor interventions?**

Two items assessed the use of theory to select participants and/or tailor intervention techniques for individual participants. Item 4 assessed whether theory was used to select participants based on their scores or levels on a particular theoretical construct or predictor. Item 6 assessed whether theory was used to tailor the intervention to the needs of individual participants. A total score was calculated by summing the scores on items 4 and 6, where “yes” = 1 and “no” = 0. Thus, total scores ranged from 0 (no use of theory) to 2 (optimal use of theory).

4) **Were relevant theoretical constructs measured?**

One item (12) assessed whether the targeted theoretical constructs were measured. If at least one of the targeted constructs/predictors was measured pre/post intervention *or* post-intervention, the item was coded as 'yes'. If the construct/predictor was not measured or if it was only measured pre-intervention, the item was coded as 'no'. Thus, total scores for this measure ranged from 0 (no theoretical constructs were measured) to 1 (at least one theoretical construct was measured pre-post or post-intervention).

5) **Is theory tested or refined?**

Four items on the TCS reflect the extent and nature of theory-testing. Item 15 assessed whether the intervention led to significant changes in at least one targeted theoretical construct, and items 16-18 assessed whether these changes explained the intervention effect. Item 16 assessed whether the study provided evidence of that changes in the theoretical construct led to changes in behavior through mediational analysis. Item 17 assessed whether the results were discussed in relation to theory, and item 18 assessed whether the results provide appropriate evidence to support or refute the theory. A total score for was calculated by summing the scores of items 15-18, where "yes" = 1 and "no" = 0. Thus, total scores ranged from 0 (no theory-testing or refinement) to 4 (optimal theory-testing and refinement).

6) **Overall use of theory.**

A total theory score was calculated by summing the totals of composite measures 1-5, where a score of zero reflected minimum (inadequate) use of theory, and a score of 15 reflected maximum (optimal) use of theory.

Analysis

Descriptive statistics were calculated to determine the extent to which theory was used and how theory was used in interventions in the sample. The extent of theory-use was assessed by calculating the percentage of studies that were coded as "Yes" for each item on the TCS.

To assess whether the use of theory, extent of theory-use, and/or specific uses of theory predicted the effectiveness of interventions, a series of subgroup analyses and univariate random effects meta-regressions were performed, using the same approaches explicated in detail in the methods section for Aim 1. Moderator analyses were conducted on the two categorical variables assessing theory use: explicit mention of theory (Y/N); based on a single theory (Y/N). Univariate meta-regression models were used to examine how much of the between-study variance could be explained by each continuous study-level variable (total theory-use score, and scores on each composite measure).

Results

Descriptive Statistics

Table 2.1.0 presents the results of the TCS. Cohen's kappa values ranged from 0.67-1.0 (Mean = 0.75) for the individual items on the TCS, indicating substantial to perfect agreement (Landis & Koch, 1977). Scores on individual items and composite measures are discussed below.

Category 1: Was Theory Mentioned?

This three-item composite measure assessed whether theory and/or theoretical predictors/constructs were explicitly mentioned and used to inform the development of

the intervention. Over two-thirds of the studies included in the review (68%; n=26) mentioned a theory, even if theory was not used to inform the intervention (item 1). The most common theories referenced were the Transtheoretical/Stages of Change Model (n=13), Social Cognitive/Social Learning Theory (n=8), and Operant Conditioning (n=7). Other theories that were mentioned included Empowerment Theory (n=1), the Health Belief Model (n=1), Community Mobilization Theory (n=1), and Marital Theory (n=1). A total of 26 studies (68%) explicitly mentioned and targeted theoretical predictors of smoking behavior, *and* provided appropriate evidence from the literature of the link between theory and behavior (item 2). Only 24% (n=9) of the interventions included in the review were based on a single theory (rather than multiple theories or a combination of theoretical predictors), indicating that most trials did not use theory in an optimal manner (item 3). For theory-testing purposes, interventions based on a single theory are considered optimal, as the use of multiple theories (or a combination of theoretical predictors) tends to obscure the theorized pathways of change. The mean score on this composite measure was 1.55, on a scale of 0 (no theory use) to 3 (optimal theory use).

Category 2: Are Relevant Theoretical Constructs Targeted?

This 5-item composite measure reflected the degree to which theory was used to inform the selection of intervention techniques, and the degree to which intervention techniques were explicitly linked to theory-relevant constructs/predictors. Two-thirds of trials included in the review (n=25) reported using theory or theoretical predictors to inform the selection of intervention techniques (item 5). Only one trial (Stotts, 2004) reported an explicit link between all intervention techniques and at least one theory-relevant construct or predictor (item 7), while 19 trials (50%) reported an explicit

link between at least one, but not all, of the intervention techniques and at least one theory-relevant construct or predictor (item 8). Only two trials (Donatelle, 2000a; Hennrikus, 2000) reported targeting all of the theoretical constructs within a specified theory (or all theoretical constructs mentioned in the study) with specific behavior change techniques (item 10), while 21 trials (55%) reported targeting at least one, but not all, of the theoretical constructs with at least one behavior change technique (item 11). Five trials (13%) used theory to link a group of techniques to a group of theory-relevant construct or predictors (item 9). The mean score for this composite measure was 2.5, on a scale of 0 (no theory use) to 5 (optimal theory use).

Category 3: Is Theory Used to Select Participants or Tailor Interventions

This two-item composite measure reflected the degree to which theory was used to select participants for the intervention and/or to tailor intervention techniques for individual participants. Only one intervention (Cinciripini, 2010) reported using theory to select participants based on their scores or levels on a particular theoretical construct or predictor (item 4). In this trial, participants were selected based on meeting a threshold for depressive symptomology. Eight studies (21%) utilized theory to tailor intervention techniques for individual participants (item 6). Most frequently, intervention techniques were tailored according to participants' stage of change/readiness to quit smoking. The mean score for this composite measure was 0.26, on a scale of 0 (no use of theory) to 2 (optimal use of theory).

Category 4: Are the relevant theoretical constructs measured?

While many studies included measures of theory-relevant constructs/predictors at baseline, very few included follow-up assessments during the post-intervention

period. As such, only 13% of trials (n=5) were coded ‘yes’ on item 12. Of the five trials that included post-intervention measures of theoretical constructs, two trials used measures that were all previously validated and included evidence of their reliability. The other three trials used at least one measure that was previously validated and had some evidence for its reliability, but also used measures that were not validated and did not have evidence of reliability. The mean score for this measure was 0.26, on a scale of 0 (no use of theory) to 1 (optimal use of theory).

Category 5: Is theory tested?

Item 15 assessed whether the intervention led to significant changes in at least one targeted theoretical construct, and items 16-18 assessed whether these changes explained the intervention effect. Only three trials presented evidence that the intervention produced significant changes in one or more theoretical constructs or predictors in favor of the treatment group. In the intervention conducted by Stotts and colleagues (2004), participants in the treatment group reported significant increases in self-efficacy, while also reporting significant decreases in depression and temptation to smoke. In the trial conducted by Hennrikus and colleagues (2010), treatment-group participants reported significant increases in both positive- and negative-support behaviors by their designated supporter (a female friend or family member selected by the subject to help her quit smoking). Finally, in the trial conducted by Ondersma and colleagues (2012), treatment-group participants reported significant increases in likelihood to quit smoking, confidence to complete a successful quit attempt, and readiness to quit smoking. However, none of these studies provided evidence, through mediation analyses, that smoking outcomes were explained by these changes. Thus,

while the interventions produced significant changes in theoretical constructs/predictors, it is not possible to determine whether these variables accounted for observed changes in behavior. The next two items assessed whether the results were discussed in relation to theory (item 17) and whether the results support or refute the theory (item 18). Just over half of the trials included in the review (n=20) were coded 'yes' on item 17, while none of the studies were coded 'yes' on item 18. The final item (19) on the TCS assessed whether the authors attempted to refine the theory upon which the intervention was based by either adding or removing constructs, or specifying that relationships between the theoretical constructs should be changed. None of the trials included in the review met the criteria for coding 'yes' on item 19. The mean score for this composite measure was 0.61, on a scale of 0 (no use of theory) to 4 (optimal use of theory).

Overall Theory Score

A total theory score was calculated by summing the totals of categories 1-5, where a score of zero reflects no use of theory, and a score of 15 reflects optimal use of theory. Observed scores ranged from zero to 11, with a mean score of 5.05.

Moderator Analyses and Meta-Regression

Moderator Analyses

Subgroup analyses on the two categorical theory variables did not reveal a significant moderating effect of either variable. The effect size for interventions that did explicitly mention theory (n = 20) did not differ significantly from the effect size for interventions that did not explicitly mention theory (n = 14) ($Q_b[1] = 0.882$; $p= 0.348$). Similarly, the effect size for interventions based on a single theory (n = 8) did not differ

significantly from the effect size for interventions that were not based on a single theory (n = 8) (Q_b[1] = 1.21; p= 0.271). The non-significant results of the subgroup analyses were further confirmed by overlapping confidence intervals, indicating that intervention effectiveness did not differ significantly between levels of these two theory-related variables. See **Tables 2.1.1** and **2.1.2** for full results.

Meta-Regression: Theory Coding Scheme Scores

Univariate meta-regression models were used to determine how much heterogeneity in effect sizes for the primary outcome of late pregnancy smoking cessation could be accounted for by theory-related variables. The results of the meta-regression analyses are described below.

Model 1: TCS Category 1: The first meta-regression model revealed that scores on the TCS Category 1 ("was theory mentioned?") accounted for 8% (R^2 analog = 0.08) of the total between-study variance in late-pregnancy smoking abstinence. The regression coefficient (b=-0.151; 95%CI: -0.360-0.059; p= 0.159) indicated a non-significant, negative association between scores on TCS category 1 and late-pregnancy smoking abstinence. See table 2.2.1 for full results.

Model 2: TCS Category 2: The second meta-regression model revealed that scores on the TCS Category 2 ("Are theory-relevant constructs mentioned?") did not account for any of the between-study variance in late-pregnancy smoking cessation (R^2 analog = 0.0). The regression coefficient (-0.045; 95%CI: -1.33-0.043; p= 0.319) indicated a non-significant, negative association between scores on TCS category 2 and late-pregnancy smoking abstinence. See table 2.2.2 for full results.

Model 3: TCS Category 3*: The third meta-regression model revealed that scores on the TCS Category 3 ("Was theory used to tailor or select participants?") accounted for 25% of the between-study variance in late-pregnancy smoking cessation (R^2 analog = 0.36). The regression coefficient (-0.360; 95%CI: -0.685- -0.035; p= 0.023) indicated a significant, negative association between scores on TCS category 3 and late-pregnancy smoking abstinence. However, it is possible that these results were influenced by the characteristics of study participants in interventions that used tailoring, as many of these interventions focused on high-risk populations such as pregnant women with depression or substance use disorders. See table 2.2.3 for full results.

Model 4: TCS Category 4: The fourth meta-regression model revealed that scores on the TCS Category 4 ("Were relevant theoretical constructs measured?") did not account for any of the between-study variance in late-pregnancy smoking cessation (R^2 analog =0.00). The regression coefficient (b= -0.372; 95%CI: -0.853-0.109; p=0.130) indicated a non-significant, negative association between scores on TCS category 4 and late-pregnancy smoking abstinence. See table 2.2.4 for full results.

Model 5: TCS Category 5*: The fifth meta-regression model revealed that scores on the TCS Category 5 ("Is theory tested?") accounted for 45% of the between-study variance in late-pregnancy smoking abstinence (R^2 analog = 0.45). The regression coefficient (b= -0.379; 95%CI: -0.636- -0.123; p = 0.004) indicated a significant, negative association between scores on TCS category 5 and late-pregnancy smoking abstinence. See table 2.2.5 for full results.

Model 6: TCS Total Score: The sixth meta-regression model revealed that total scores on the TCS accounted for accounted for 19% of the total between-study

variance in late-pregnancy smoking cessation (R^2 analog = 0.19). The regression coefficient (b=-0.055; 95%CI: -0.112-0.001; p= 0.057) revealed a non-significant, negative association between total theory score and late-pregnancy smoking cessation. See **Table 2.2.6** for full results.

Discussion

The purpose of this chapter was to use a standardized coding scheme to evaluate the use of theory in intervention design, implementation, and evaluation. To our knowledge, this is the first study to assess the use of health behavior theories in the published literature on prenatal smoking cessation interventions.

Based on the scores of the individual items and composite measures, it is apparent that theory is not being utilized to its full capacity in the development, implementation, and evaluation of smoking cessation interventions for pregnant women. While many studies mentioned theory and/or theoretical predictors of smoking-related behavior, most interventions were only loosely based on theory and did not allow for theory to be tested or refined. Of the 26 published trials that explicitly mentioned theory in the introduction or methods, only nine were based on a single theoretical framework. Five of these studies utilized the learning-based theory of operant conditioning (Cinciripini et al., 2010; Heil et al., 2008; Higgins, et al., 2004; Higgins et al., unpublished; Higgins et al., 2014), two studies utilized the transtheoretical/stages of change model (Stotts et al., 2004; 2009), one study used social cognitive theory (Patten et al., 2012), and one study used social learning theory (Secker-Walker et al., 1997). Even among these nine trials, theory was used primarily in a descriptive manner, as

opposed to an explanatory or predictive manner. While a theoretical basis for the intervention was stated in many studies, there was a general failure to explain how theory was used to inform intervention design, how the evaluation tested theory, and how the results may support or refute the stated theory. Even when explicit pathways of change were described, none of the studies included in this review provided evidence that changes in smoking behavior could be explained through the theorized pathways. Thus, the results of these studies have limited utility in terms of theory-building and refinement.

Many studies measured theory-relevant constructs at the baseline assessment, but only five studies included a follow-up assessment during the post-intervention period. Baseline assessments can be used to analyze whether participants' scores on a particular theory-relevant construct are associated with their likelihood of quitting smoking. However, intermediate and post-intervention assessments are necessary to determine whether the intervention led to significant changes on measures of theory-relevant constructs, and whether changes on theory-relevant constructs can explain the observed changes in behavior. Only five studies included post-intervention assessments, and only three of these provided evidence of significant changes on theory-relevant constructs in favor of the intervention group. Furthermore, conceptual and methodological differences in the measures used to assess theory-relevant constructs at baseline prohibited their inclusion in the meta-analysis. While many studies measured constructs such as self-efficacy and motivation, there was significant variation in psychometric properties (most notably, many studies constructed their own measures instead of using previously-validated measures), units of measurement,

terminology, and conceptual definitions (for example, some participants were asked how motivated they were to stop smoking generally, while other participants were asked how motivated they were to stop smoking within a specific time period).

The results of the subgroup analyses and meta-regression models were counter to the hypothesis that use of theory would be positively associated with intervention effectiveness. Scores on two categories of the theory coding scheme ("Was theory tested?" and "Was theory used to tailor or select participants?") were significantly associated with the primary outcome of late-pregnancy smoking abstinence, but both of the associations were negative, indicating that greater use of theory was associated with a lower likelihood of smoking abstinence during the late-pregnancy period.

The overall Theory Coding Scheme score was not significantly associated with the likelihood of achieving late-pregnancy smoking abstinence, but the regression coefficient was negative, which is in line with the finding that use of theory was negatively associated with the effectiveness of prenatal smoking cessation interventions in this review. However, this finding may be a reflection of the overall poor use of theory, rather than the true relationship between theory-use and intervention effectiveness. As described above, theory was rarely used optimally to inform intervention design. More often than not, theory was mentioned but not used explicitly to select 1) targets of change (constructs) and 2) the techniques to target these constructs (BCTs). Furthermore, even when theory was used to identify targets of change and/or BCTs, few studies included appropriate measures of the targeted constructs. As a result, it was not possible to determine whether the selected BCTs were effectively changing the targeted constructs. Additionally, none of the studies included in this review discussed

parameters of effectiveness, or the conditions that must be satisfied for the intervention to be effective (Kok et al., 2016; Peters, Ruiter, & Kok, 2013). For example, when measured, most women in the studies included in this review indicated at baseline that they already perceived the risks of smoking during pregnancy as high, that they wanted to stop smoking, and that they were motivated to do so. As such, BCTs targeting perceived risk or motivation to quit would likely yield only limited effects on the targeted constructs and subsequently, on behavior.

On a similar note, the results of the TCS only reflect the utility of the theories used in the studies included in the review. Thus, in addition to inadequate and suboptimal applications of theory, the observed lack of association between use of theory and intervention effectiveness may also reflect a poor choice of behavior change theories. For example, the most common theory mentioned by the studies included in the review was the Transtheoretical/Stages of Change Model, (TTM) which is widely used in the context of smoking cessation. However, previous studies describe the stages of the TTM as “arbitrary” (Sutton, 2000, p. 209-211) and several reviews of stage-based interventions have concluded that the psychological processes underpinning the TTM are not supported by the available evidence (Bandura, 1998; Sutton, 2000; Weinstein et al., 1998). Furthermore, evidence suggests that interventions tailored based on the stages of change are no more effective than interventions than do not include stage-based tailoring (Riemsma 2003). Thus, even if the theory is used optimally, the TTM may not enhance the effectiveness of the intervention. Given that the TTM was so widely used among the studies in this review, it is possible that this may explain (at least in part) the lack of association between TCS scores and intervention

effectiveness. This is in line with previous research, which suggests that the selection of theory is often based on the popularity of the theory, rather than evidence of its effectiveness in a specific context or behavioral domain (Sutton, 2000).

While theory-based interventions are considered to be more effective in achieving health behavior change than interventions that do not utilize a theoretical foundation (Noar & Zimmerman, 2005), other recent reviews examining the use of theory in behavior change interventions have also found mixed and even negative results. In a systematic review of behavior change interventions based on the Theory of Planned Behavior, Hardeman and colleagues (2002) found that intervention effectiveness was unrelated to use of theory to develop the intervention; specifically, that use of theory was not associated with significant changes in behavioral intentions or behaviors. The study found that, although the TPB was often used descriptively, it was rarely used to select intervention targets and it was often used incompletely. Furthermore, many of the interventions also used other theories and models to inform intervention design, which complicates theory testing and may obscure the relationship between theory-use and intervention effectiveness (Prestwich et al., 2014). As the authors noted, “it is difficult to assess the true effectiveness of using the TPB, as interventions were rarely designed on the basis of the theory, and often also other theories and models were used to develop the intervention” (Hardeman et al., 2002, p. 149). Colquhoun and colleagues (2013) reported similar findings in a systematic review of the use of theory in randomized controlled trials of audit and feedback interventions. Overall, the explicit use of theory in audit and feedback trials was found to be rare. A range of theories were used as the conceptual basis for such trials but there was a lack

of consistency in the application of theory, which made it difficult to determine whether certain theoretical approaches were superior, and to explore the mechanisms through which audit and feedback interventions work (i.e., the causal pathways) (Colquhoun et al., 2013). In another review of audit and feedback interventions, Gardner and colleagues (2010) examined the use of two specific theories (Feedback Intervention Theory and Control Theory) to see if they could link intervention components to specific theoretical constructs in an effort to determine which factors contribute to effectiveness. The authors found that in most studies, theory was either not used sufficiently or not described in enough detail to allow for a clear analysis of whether theory-use contributed to intervention effectiveness (and if so, how) (Gardner et al., 2010).

To date, the strongest evidence supporting the use of behavior change theories to inform intervention design comes from observational studies (e.g., cross-sectional or longitudinal). Using meta-analysis, researchers have confirmed the predictive power of theoretical constructs such as behavioral intentions (from the Theory of Planned Behavior; e.g., McEachan, Conner, Taylor, & Lawton, 2011), the stages of the Transtheoretical Model (Bui, Mullan, & McCaffery, 2013; Marshall & Biddle, 2001), and self-efficacy (from Social Cognitive Theory; e.g., Spence, Burgess, & Cutumisu, 2006). However, because these findings come from observational studies, they fail to meet Prochaska et al.'s (2008) efficacy criterion, which states that a theory-based intervention "is demonstrated to have significant efficacy" if it produces "greater behavior change than a placebo or control" (p. 565).

Overall, these results suggest that there are important parameters that must be considered when developing interventions based on behavior change theory. They also

underscore the importance of adequately using theory to inform decision-making during the intervention design process, rather than simply discussing theory in published reports of the intervention. This includes steps such as including strategies designed to target relevant theoretical constructs, ensuring adequate measurement of behavior and theoretical constructs (i.e., using validated measures of theoretical constructs at baseline and including at least one follow-up assessment to measure changes in targeted constructs), as well as explaining how and why a specific theory was chosen to inform intervention design, and thoroughly describing the hypothesized mechanisms of change in published reports of intervention trials. Finally, it is possible that certain applications of theory (e.g., to target certain participants, to identify targets of change, to select behavior change techniques, etc.) may be effective when used simultaneously but not in isolation.

CHAPTER 5

Specific Aim 3

Because of the complexity of the determinants and correlates of maternal smoking behavior, there are conflicting perspectives about the best approaches to address the problem. Interventions that employ a greater number of behavior change techniques tend to be more effective than those with fewer components (Webb, Joseph, Yardley, & Michie, 2010). However, interventions comprised of numerous, interacting components also present a unique challenge, as researchers must isolate the effects of these components to establish *why* a certain intervention worked and if its effects were contingent on certain conditions, and to replicate its effectiveness in the future. This problem was described by Bryant and colleagues, who explained that "smoking cessation interventions typically incorporate substantial behavioral components that are difficult to both describe and reproduce" (Bryant, Passey, Hall, & Sanson-Fisher, 2014, p. 2).

Components of interventions include both the techniques used to promote behavior change ("active ingredients") and the procedures for delivering the intervention techniques. These procedures include information about who delivers the intervention and to whom, as well as the recommended frequency, dose, format, and duration of delivery, and the contexts in which the intervention is designed to be delivered (Davidson, Goldstein, Kaplan, Kaufmann, Knatterud, Orleans, et al., 2003). To determine how an intervention worked or why it did not work, all of these components must be clearly described and defined. Currently, however, there is no clear consensus on guidelines for specifying the content of interventions (Michie et al., 2011a; 2011b).

Guidelines such as the CONSORT Statement for the reporting of evaluation trials, the TREND Statement for the reporting of evaluations with non-randomized designs, and the STROBE statement for the reporting of observational studies, all call for intervention content to be described in published manuscripts but lack explicit guidance on what to report and how to report it (Michie et al., 2009b; Michie et al., 2011a; 2011b). As a result, there is great variation in the details provided in published intervention studies, making it difficult to synthesize evidence and identify the mechanisms of action underpinning effective behavior change interventions. This is in stark contrast to biomedical and pharmaceutical interventions, which mandate explicit and precise directions for delivery, dosing, and mechanisms of actions, as well as complete information on the drug's active ingredients.

Furthermore, inconsistent use of terminology limits the potential to draw conclusions even among studies that include detailed descriptions of intervention components. For example, the terms 'psychosocial counseling' and 'behavioral counseling' are used interchangeably throughout the prenatal smoking cessation literature to describe a wide variety of techniques ranging in content, delivery, intensity, and duration (Chamberlain et al., 2013; Lumley et al., 2009). With such variation encompassed under one term, coupled with the confusion of overlapping terminology describing the same phenomenon, valid comparisons and replication are often not possible (Michie et al., 2009b).

A similar problem arises when the content of interventions is described in a way that conflates intervention techniques with the characteristics of delivery (Davidson et al., 2003). For example, descriptions of behavior change techniques used in published

reports of smoking cessation interventions for pregnant women include “peer support,” (Donatelle et al., 2000; Hajek et al., 2001; Hennrikus et al., 2010; Solomon et al., 2000), “clinic-based counseling,” (El-Mohandes et al., 2011), “home-based visiting” (Graham 1992), “computer-based counseling” (Ondersma et al., 2012), “telephone counseling” (Bullock et al., 1995; Solomon et al., 2000; Stotts et al., 2002; Rigotti et al., 2006), and “nurse-delivered telephone support” (Bullock et al., 2009), all of which make it impossible to distinguish the effects of specific behavior change techniques (e.g., counseling, support, etc.) from the characteristics of delivery (e.g., telephone-based, computer-based, nurse-delivered, etc.) and the context of delivery (e.g., home-based, clinic-based, digital, etc.).

Specifying Intervention Components

Establishing reliable methodology for specifying intervention components is a key starting point for evidence synthesis, allowing investigators to identify and evaluate the effectiveness of distinct behavior change techniques, as well as the factors that may influence their effectiveness. The recent development of taxonomies of theory-linked behavior change techniques provides a novel framework to reliably code the content of interventions (Abraham & Michie, 2008; Michie et al., 2011, 2013) as described in intervention reports and guidelines. In addition to providing a foundation for the synthesis of evidence across interventions targeting a specific health behavior in a particular population and/or context, behavior change taxonomies have also been utilized to investigate the moderating effects of empirically or theoretically derived features and/or clusters of intervention techniques (Dombrowski et al., 2012; Greaves et al., 2011; Michie et al., 2009; Taylor, Conner, & Lawton, 2012; Webb, Joseph, Yardley,

& Michie, 2010). This has produced a growing body of evidence on the most (and least) effective intervention components across different theoretical domains, thus helping to inform the development of more effective interventions in the future (Michie & Johnston, 2012).

In a review of smoking cessation treatment manuals utilized by the English Stop Smoking Services, Michie and colleagues (Michie, Hyder, Walia, & West, 2011) identified 43 unique behavior change techniques used to provide individual behavioral support for smoking cessation. These 43 techniques were grouped by their primary function, resulting in the following four categories of behavior change strategies: 1) techniques that directly address motivation, such as contingency management and positive reinforcement; 2) techniques that focus on maximizing skills or self-regulatory capacity, such as problem-solving and goal-setting; 3) techniques that promote adjuvant activities, such as providing advice on pharmacological cessation aids and facilitating the development of social support networks; and 4) techniques that focus on supporting and enhancing other intervention components, such as building rapport and tailoring materials.

This taxonomy, which was developed based on written protocols of intervention trials, was later used to specify the content of smoking cessation behavioral support interventions as actually delivered in practice (Lorencatto, West, Seymour, & Michie, 2013). Using transcripts of audio-recorded consultations delivered by the English Stop Smoking Services, Lorencatto and colleagues (2013) established the feasibility and reliability of applying such a taxonomy to identify behavior change techniques and evaluate variability in the provision of behavioral support interventions for smoking

cessation interventions in a general patient population. However, as the authors acknowledged, “[t]his study is only a starting point in the labeling and classification of BCTs for smoking cessation. The list was identified and analyzed using guidance documents and treatment manuals from just one country and represented the current practice in that country. It is possible that different techniques may be used in other contexts or added in the future” (Michie, Hyder, Walia, & West, 2011, p. 318). Additionally, Michie and colleagues’ (2011) review focused only on behavioral support interventions, the goal of which is to “change the balance of impulses and inhibitions by reducing impulses to smoke and increasing motivation and capacity to resist those impulses on all relevant occasions (p. 316).” The authors used the PRIME theory to inform the development of their coding manual, which provided a coherent structure but may not have captured the entire range of theories and behavior change techniques used in smoking cessation interventions. Furthermore, there is also a need to consider the role of environmental influences in the process, as health behavior and behavior change interventions are embedded within social, cultural, and/or physical systems (Golden & Earp, 2012). Finally, given that the vast majority of smoking cessation research is based on published reports of intervention trials (as opposed to intervention manuals or observations of intervention implementation), there is a need to develop and refine taxonomies for the specific purpose of applying them to published studies.

Behavior Change Techniques

The purpose of developing taxonomies of behavior change techniques is to establish a systematic method for classifying and defining intervention components, with the goal of advancing a cumulative science of behavior change. The effectiveness of

any given technique is not part of its definition. Rather, evidence of effectiveness is the product of research *using* reliable definitions to identify common and distinctive behavior change techniques across published intervention studies. When applied in conjunction with meta-analytic methods, taxonomies of behavior change techniques can be used to test the effectiveness of specific techniques, as well as to test for potential moderating factors. Albarracin and colleagues (2005) were among the first to demonstrate the feasibility of such an approach. In a meta-analysis of interventions designed to promote condom use, the investigators first identified 10 common behavior change techniques that were included in published reports of intervention trials. Next, they demonstrated which techniques were associated with effectiveness, and then conducted moderator analyses to determine how technique effectiveness was influenced by characteristics of the recipients. The results showed that certain techniques, such as provision of normative arguments, were effective only when used with younger participants (under age 21), which allowed the investigators to make recommendations for future intervention design. This approach also allowed the investigators to test the assumptions of relevant behavior change theories and make recommendations based on the results. For example, the analysis revealed that the provision of attitudinal arguments and normative arguments were associated with increased effectiveness, lending support for the use of theories such as Fishbein & Ajzen's (1975) Theory of Reasoned Action. On the other hand, the provision of threat-inducing arguments was not associated with effectiveness, suggesting that theories based on fear appeals may not provide useful guidance in the context of promoting condom use. In a meta-analysis conducted the same year, Hillsdon and colleagues (2005) applied a similar

methodology to identify effective behavior change techniques used in community-based physical activity interventions (Foster, Cavill, Crombie, and Naidoo, 2005). The results revealed that interventions that included telephone support, encouragement of self-monitoring, and/or provision of written instructional materials were most likely to be effective. Importantly, these three techniques were *not* identified in Albarracin and colleagues' (2005) meta-analysis, indicating the need for a more comprehensive set of behavior change techniques.

In 2008, Abraham and Michie addressed this need with the development of a reliable, comprehensive and theory-linked taxonomy of behavior change techniques, which provides the foundation for categorizing intervention content and synthesizing evidence across published intervention studies. The promise of this approach was demonstrated in a recent review of physical activity and dietary change interventions (Michie, Abraham, Whittington, McAteer, & Gupta, 2009). Researchers first used a standardized taxonomy to describe and classify the behavior change techniques employed by studies in the review. Next, meta-regression analysis was used to isolate and quantify the effects of these techniques, leading to the conclusion that interventions using the technique of self-monitoring explained the greatest amount of heterogeneity among studies. Finally, using Control Theory to identify theoretically-linked behavior change techniques, the investigators found that interventions using self-monitoring in combination with at least one other technique derived from Control Theory were more effective than other interventions, including those which used self-monitoring in combination with other, non-theory-derived techniques (Michie et al., 2009). Using similar methods, Dombrowski and colleagues (2012) identified and analyzed the

effectiveness of behavior change techniques used in interventions targeting dietary and/or physical activity change for obese adults. The analysis revealed that four techniques (provision of instruction, self-monitoring, relapse prevention, and prompting practice) were linked to more successful intervention outcomes. Most interventions included in the meta-analysis employed multiple behavior change techniques, but simply increasing the number of techniques was not necessarily associated with better outcomes. However, the use of multiple techniques derived from Control Theory was associated with greater weight loss when compared to other combinations of techniques. Gardner and colleagues (2010) applied a similar methodology in a meta-analysis of audit and feedback interventions. The investigators found that audit and feedback interventions were effective overall in changing behavior, but there was significant among-study variation in effectiveness. Using meta-regression, the authors assessed whether the variation in effectiveness could be explained by the use of techniques linked to Control Theory. The results revealed an overall dearth of theory in the design and evaluation of audit and feedback interventions, which limited the potential to explore whether techniques linked to Control Theory led to better intervention outcomes. However, the authors were able to determine that the addition of goal-setting and action-planning (at the same time) increased the effectiveness of feedback. In a meta-analysis of HIV risk reduction interventions, Smoak and colleagues (2006) used a multivariate meta-regression model to evaluate the predictive utility of Fisher and Fisher's (1992) information-motivation-behavioral skills (IMB) model (Smoak, Scott-Sheldon, Johnson, & Carey, 2006). Consistent with the theory, interventions that included informational, motivational, and behavioral skills components led to greater risk

reduction than interventions that did not include all three IMB model components. Additionally, interventions were found to be more effective when they included higher doses (hours of exposure) of the components.

To date, nearly all evaluations of prenatal smoking cessation interventions have focused on the entirety of the intervention, rather than its individual components. The results of such evaluations can be used to determine whether or not an intervention worked, but not *why* or *under what conditions* it worked. To develop more effective interventions, it is necessary to understand what makes effective interventions work in the first place. In addition to informing the development of better interventions, identifying the effective components of interventions also has important implications for resource allocation, as it may be possible to design more parsimonious interventions without sacrificing results.

As Homish and colleagues noted in a 2012 review of social and environmental factors related to smoking during pregnancy, there is a need for additional research not only on which intervention technique or techniques are effective, but also on when these techniques are most appropriate and for which population(s) of pregnant women (Homish, Eiden, Leonard, & Koszlowski, 2012).

With these goals in mind, the third major aim of this meta-analysis is to identify, describe, and quantify the effects of individual techniques described in published reports of prenatal smoking cessation interventions, and to explore factors that may serve as parameters of effectiveness for each technique. These aims are described in further detail below:

Aim 3: To identify, describe, and quantify the effects of behavior change techniques used in prenatal smoking cessation interventions.

1) **Sub-aim 3a:** To identify and describe standardized, theory-linked behavior change techniques used in published randomized controlled trials of prenatal smoking cessation interventions, using a coding process described by Michie and colleagues' (Abraham & Michie, 2008; Michie et al., 2008; Michie et al., 2009a).
2) **Sub-aim 3b:** To evaluate the effectiveness of each technique using subgroup analyses to calculate the effect size of interventions that used the technique compared to those that didn't use the technique, and to determine whether the total number of active BCTs used in an intervention is associated with effectiveness using a univariate meta-regression model.
3) **Sub-aim 3c:** To explore whether the effect size estimates of BCTs identified as effective in sub-aim 3b differ according to characteristics of the study design, intervention, or participants.

Methods

Sample

Studies for this section were derived from the meta-analysis conducted in the first step of this project.

Inclusion/Exclusion Criteria

In addition to meeting the criteria specified for inclusion in the meta-analysis, studies in this section must include at least one distinct **behavior change technique**, defined as "a replicable component of an intervention designed to alter or redirect

causal processes that regulate behavior" (Michie, Abraham, Eccles, Francis, Hardeman, & Johnston, 2011, p. 2). According to the operational definition proposed by Michie and colleagues (2011), behavior change techniques share three primary defining characteristics: observability; replicability; and irreducibility (Michie, Abraham, Eccles, Francis, Hardeman, & Johnston, 2011). Behavior change techniques specify the minimum content that must be delivered to allow for identification of the technique, but they are not attached to a specific mode of delivery (Abraham & Johnston, 2013). That is, they specify *what* content must be delivered, but now *how* it is delivered. Examples include goal-setting, contingent rewards/incentives, graded tasks, and prompts/cues (Abraham & Michie, 2008).

Measures

Behavior change techniques were identified using Abraham & Michie's (2008) 26-item taxonomy, which describes and defines 26 unique, theory-derived techniques. The 26 techniques reflect a variety of theoretical foundations and have been applied across different behaviors and behavior change interventions. In a series of 78 reliability tests (applying the 26 items across three reviews), the average kappa per technique was found to be 0.79 (Abraham & MIchie, 2008). Comparing 13 intervention manuals to 13 published articles describing the same intervention, average agreement was higher for techniques identified in manuals (85%) than in published studies (74%). Mismatches between treatment manuals and accompanying published reports were common; three-quarters of these mismatches arose from identification of a technique in the intervention manual that was not identified in the published study, indicating the need

for further exploration of applications of the taxonomy to different types of reference documents.

The taxonomy is also accompanied by a manual with definitions of each technique and instructions for coding and applying the taxonomy to reference documents.

Coding

Coding for behavior change techniques was performed according to the procedures specified by the authors of the taxonomy. Reference materials from a website created for training purposes were used to practice identifying and coding behavior change techniques before applying the taxonomy to the studies in this review (http://www.bct-taxonomy.com). Coding forms and instructions are presented in Appendix B.

The goal of the coding process is to capture as many techniques as possible by analyzing text from published documents, and assessing specific words and phrases that identify or describe intervention content. Once the presence of a behavior change technique was identified, the relevant text was highlighted and categorized based on the standardized definitions and terminology specified in the taxonomy. Using these procedures, behavior change techniques can be identified by name (e.g., “Nurses delivered 1-hour motivational interviewing sessions” was coded as “Motivational Interviewing”) or by analyzing the description and/or function of a technique, and then matching this to the appropriate named technique (e.g., “Nurses described the effects of smoking on fetal development” was coded as “Providing Information on Behavior-Health Link”). Some passages described more than one technique, and were coded

accordingly. For example, “counseling women about the dangers of continued smoking for the health of the fetus” was coded as “Providing Information on Behavior-Health Link” and “Providing information on Consequences (negative)”.

To ensure standardization, the wording, labels, and definitions of behavior change techniques were kept constant from the original taxonomy, with the exception of one modification: While the original taxonomy was comprised of 26 behavior change techniques, the final analysis was performed on a modified 27-item version. The additional item was created by parsing one technique (“provide information about consequences”) into two separate techniques to capture the difference between information about positive and negative consequences. While there was significant overlap (as many trials provided information about both positive and negative consequences), more trials included information about negative consequences (K = 16) than about positive outcomes (K = 12). Research suggests that presenting pregnant women with negative information about smoking, especially in the absence of accompanying positive information, may sometimes backfire and cause women to reject the information altogether (Flemming, Graham, Heirs, Fox, & Snowden, 2013), so we considered this an important distinction to make.

Intervention and control arms were coded separately. For the eight trials with multiple intervention arms, the presence of behavior change techniques was coded separately for each arm (as discussed previously, only one intervention arm was included in the meta-analysis to avoid unit-of-analysis problems associated with multiple comparisons). For the purposes of *categorizing* intervention content, we coded for the presence of each technique even if the same technique was included in the intervention

and control arms. For example, many trials included a basic educational component comprised of providing verbal and/or written information about the link between smoking and harmful health effects ('providing information about health-behavior link'), the consequences of continued smoking ('provide information on consequences -negative'), and/or the benefits of quitting smoking ('provide information on consequences – positive'). However, in order to isolate the effects of the ***active techniques***, the technique was only identified as an active component if it was not included in the control arm *or* if it was delivered in a more intensive dose than in the control arm. For example, women in the control arms of most trials were offered basic guidance on quitting smoking ("Providing Instruction"), but women in the treatment arms were often provided with more detailed, tailored, and/or specific instruction on smoking cessation; in this case, providing instruction was still identified as an active technique even though it was present in both the intervention and control arms.

Analysis

Effectiveness of Behavior Change Techniques

Treatment vs Control*.* To quantify the effectiveness of each BCT, we used random effects meta-analyses to calculate risk ratios for late pregnancy smoking cessation for the subsets of interventions that included each BCT (comparing the treatment arm to the control arm). A BCT had to be present and identified as an *active technique* in at least three studies to be included in the analysis.

BCT [Y] vs BCT [N]. Based on the results of the random effects meta-analysis models when grouped by intervention technique, a secondary analysis was conducted on techniques that were found to have a significant effect size in the first step. In the

second step, we utilized univariate random effects meta-regression models to compare the effect size for interventions that included each technique to a reference group of interventions that did not include the technique, and to explore how much between-study heterogeneity each technique explained. The specific use of meta-regression in this review is based on several recent studies that have demonstrated that univariate meta-regression can be applied successfully in this manner to quantify the unique contribution of various intervention components to intervention effects (e.g., Abell, Glasziou, & Hoffman, 2017; Dombrowski et al., 2012; Uddin et al., 2016). In this context, meta-regression extends traditional subgroup analyses to facilitate more detailed exploration of associations between study characteristics (in this case, BCTs) and intervention outcomes (RR for late pregnancy smoking cessation). Using this method, a significant p-value indicates a significant association between the study outcome and the explanatory variable (in this case, BCT [yes] compared to BCT [no]), with the direction of the regression coefficient providing an indicator of whether inclusion of a specific BCT was associated with a larger or smaller effect size (i.e., a greater or lower likelihood of achieving late pregnancy smoking abstinence).

In summary, we first examined the effect size of subsets of interventions that included each BCT, comparing the treatment group to the control group. Based on the results from the first step, BCTs with significant effect sizes were identified for inclusion in secondary analyses. In the second step, we used meta-regression models to compare the effect size for subgroups of interventions that included each BCT to those that did not include the BCT. Thus, the second step of the analysis allowed for the

determination of whether interventions that included a specific BCT were associated with a greater likelihood of achieving late pregnancy smoking abstinence than interventions that did not include the BCT.

Total number of BCTs: We also used a univariate meta-regression model to explore whether the total number of active BCTs was associated with intervention effectiveness.

Assessment of Heterogeneity

Based on the results of the secondary analyses comparing interventions that included each technique to interventions that did not include the technique, we undertook further analyses to evaluate whether the effectiveness of BCT 15 was influenced by characteristics of the intervention, participants, or the provision of additional BCTs. Specifically, we performed moderator analyses on categorical variables, using the subgroup method described previously, and then utilized random effects meta-regression models to explore continuous covariates.

Effectiveness Ratios

To further examine the effectiveness of BCTs used in smoking cessation interventions for pregnant women, we calculated 'effectiveness ratios' for every technique. For each BCT, we divided the total number of active uses of the BCT by the total number of effective uses of the BCT (as determined by a significant risk ratio when compared to its respective control group). For example, BCT 1 was utilized as an active ingredient in 12 interventions, but the risk ratio for BCT 1 was only significantly different between the control group and treatment group in three of these interventions. Thus, the ratio of effective BCT use to active BCT use was 1:4. In comparison, BCT 15 was used

as an active ingredient in 9 interventions, and had a significant risk ratio in 6 interventions, resulting in an effectiveness ratio of 1:3. The purpose of calculating effectiveness ratios was to provide a more detailed indicator of effectiveness that may be useful for intervention planners choosing among a variety of techniques. While the risk ratio provides an indicator of statistical significance, the effectiveness ratio provides an indicator of how often the technique is used successfully, relative to the frequency of its use.

Results

Descriptive Statistics

As seen in **Table 3.1.0**, almost all of the behavior change techniques were utilized in at least one trial. The only techniques that were not identified in any trials were “time management” and “prompt identification as a role model”. Average interrater reliability across all techniques (K = 0.74) was moderate to high, indicating an acceptable level of agreement and providing evidence for the feasibility of applying the behavior change taxonomy to published reports of prenatal smoking cessation interventions. (However, as described in further detail in the discussion, reliable and thorough coding of BCTs was limited by poor specification in published reports.) Interrater reliability for the specific intervention techniques ranged from k = 0.62 (for “provide information on health-behavior link”) to k = 0.91 (for “agree to a behavioral contract”). The most common behavior change techniques represented in the study sample were ‘providing instruction’ (K=29), ‘prompting specific goal setting’ (K=25), and ‘providing information on the health-behavior link’ (K=19). Two techniques (‘prompt practice’ and ‘provide information about others’ approval’) were only identified in one

study, and two techniques were not identified in any studies ('prompt identification as a role model' and 'time management'); as such, these were not included in analyses of effect size.

Table 3.1.0. Behavior Change Techniques: Intercoder Reliability and Frequencies

Behavior Change Technique	Associated theory(ies)	Intercoder Reliability (*k*)	Number of studies: Total *K* (out of 38)	Number of studies: Active *K*
1: Provide info on health-behavior link	IMB	0.62	19	12
2: Provide info on consequences (negative)	TRA, TPB, SCogT, & IMB	0.65	16	10
3: Provide info on consequences (positive)	TRA, TPB, SCogT, & IMB	0.71	12	7
4: Provide information about others' approval	TRA, TPB, IMB, SCogT	0.82	1	1
5: Prompt intention formation	TRA, TPB, SCogT, & IMB	0.66	13	7
6: Prompt barrier identification	SCogT	0.71	10	7
7: Provide general encouragement	SCogT	0.69	17	12
8: Set graded tasks	SCogT	0.83	3	2
9: Provide instructions	SCogT	0.63	29	8
10: Model/demonstrate the behavior	SCogT	0.85	6	5
11: Prompt specific goal setting	CT	0.71	25	10
12: Prompt review of behavioral goals	CT	0.7	12	5
13: Prompt self-monitoring of behavior	CT	0.68	7	1
14: Provide feedback on performance	CT	0.73	11	8
15: Provide contingent rewards	OC	0.94	9	9
16: Teach to use prompts/cues	OC	0.73	4	3
17: Agree to behavioral contract	OC	0.91	4	3
18: Prompt practice	OC	0.76	1	1
19: Use follow-up prompts	OC	0.71	10	8
20: Provide opportunity for social comparison	SCogT	0.75	4	3
21: Plan social support/social change	Social support theories	0.71	10	7
22: Prompt identification as role model	Stress & coping theories	0.90	0	0
23: Prompt self-talk	IMB	0.67	7	3
24: Relapse prevention	Relapse prev. therapy	0.73	10	5
25: Stress management	Stress & coping theories	0.71	6	2
26: Motivational interviewing	SCogT, IMB	0.84	11	8
27: Time management	IMB	1.0	0	0

The number of unique behavior change techniques (“active ingredients”) varied substantially between trials, from a minimum of one active ingredient to a maximum of 12, with a mean of 4.7 per treatment arm, as seen in **Table 3.1.1.**

Effectiveness of BCTs

Treatment vs. Control

When grouped by the inclusion of each BCT (i.e., using the “select if” command to limit the analysis to subsets of the overall sample), random-effects meta-analyses revealed significant differences in late pregnancy smoking cessation in favor of the intervention group for subsets of interventions that included any of the following techniques: BCT 1 (“Provide general information about health-behavior link”), BCT 2 (“Provide information about consequences [negative]”), BCT 5 (“Prompt intention formation”), BCT 9 (“Provide general instruction”), BCT 11 (“Prompt specific goal setting”), BCT 15 (“Provide contingent rewards”), BCT 16 (“Teach to use prompts/cues”), or BCT 17 (“Agree to behavioral contract”). See **Table 3.1.2** for full results. Subgroup analyses on eligible BCTs were repeated after removing four studies identified as potential outliers (El-Mohandes et al., 2013; Pollak et al., 2007; Secker-Walker et al., 1997; Tuten et al., 2012), but the significance of the results was unchanged.

BCT [Y] vs BCT [N]

Based on the results of the random effects models when grouped by intervention technique, further analyses were limited to the eight techniques that demonstrated effectiveness in comparisons of the treatment vs control conditions. In univariate

random effects meta-regression analyses comparing the effect size for subsets of interventions that included each technique to a reference group of interventions that did not include the technique, only one of the techniques (BCT 15: “provide contingent rewards”) was associated with a significantly increased likelihood of achieving late-pregnancy smoking abstinence. The results of the meta-regression model revealed that BCT 15 explained 72% of the between-study variance in late-pregnancy smoking cessation. The regression coefficient (b=0.785; 95%CI: 0.419-1.152; p<0.0001) was significant, indicating that the provision of contingent rewards was associated with a significantly greater likelihood of achieving late pregnancy smoking abstinence, compared to a reference group of interventions that did not provide contingent rewards. These findings were confirmed in subgroup analyses, which revealed that interventions that provided contingent rewards had a larger effect size (n=9; RR=2.82; 95%CI: 2.05-3.88) than interventions that did not provide contingent rewards (n=25; RR=1.30; 95%CI: 1.12-1.49). A significant between-group heterogeneity statistic ($Q_b[1]=19.07$; p<0.001) and non-overlapping confidence intervals indicated that the difference in effect sizes was significant, such that interventions that provided contingent rewards were more effective than those that did not. (See **Table 3.2.6** for full results).

Univariate random effects meta-regression models revealed that BCT 1 (provide information about health-behavior link; n=12), BCT 2 (provide information about consequences [negative]; n=10), BCT 5 (prompt intention formation; n=7), BCT 9 (provide general instruction; n=9), BCT 11 (prompt specific goal setting; n=10), and BCT 16 (teach to use prompts/cues; n=3) did not explain any between-study variance in effect sizes (R^2= 0.00). Interventions that included any one of these techniques were not

significantly more effective than interventions that did not include the respective technique. BCT 17 (agree to behavioral contract) explained 6% of the between-study variance in effect sizes (R^2= 0.06), but the regression coefficient was not significant (b= 0.410; 95%CI: -0.240-1.06; p=0.216), indicating that the effect size for interventions that included BCT 17 was not significantly different when compared to the reference group of interventions that did not include BCT 17. (See **Tables 3.2.1 to 3.2.8** for full results).

Total Number of BCTs

A univariate random-effects meta-regression model indicated that the total number of behavior change techniques used ("total BCT's) did not explain any of the between-study variance (R^2 analog = 0.0) in late-pregnancy smoking abstinence. The regression coefficient for total BCTs (b = -0.049; 95%CI: -1.09-0.011; p= 0.105) was not statistically significant, indicating that the number of techniques used within an intervention was not associated with the likelihood of achieving late-pregnancy smoking abstinence.

See **Table 3.3.1** for full results.

Contingent Rewards

Having established that BCT 15 ("contingent rewards/incentives") was the only behavior change technique that demonstrated evidence of a moderating effect (i.e., that the effect size for interventions providing contingent rewards was significantly greater than for interventions not providing contingent rewards), additional analyses were carried out to explore study-level variables that may influence the effectiveness of contingent rewards. The results of the analyses examining BCT 15 are described below, beginning with a description of the subset of nine interventions that included BCT 15.

Subgroup and meta-regression analyses should be interpreted as exploratory, given that the subset of studies providing contingent rewards was a homogenous subset to begin with.

Descriptive Statistics

Among the subset of nine interventions that provided contingent rewards, eight were categorized as ‘low-SES’, while only one was categorized as ‘not low-SES’. The participants in this subset were generally healthy, with eight of nine studies categorized as ‘healthy’ and only one study focused specifically on participants with mental health and/or substance use disorders. However, seven studies in this subset were categorized as ‘high psychosocial risk’, as indicated by low social support, high stress, or depression among at least 50% of participants. One study in this subset was categorized as ‘majority minority’, while the other eight were not. When grouped by deliverer, eight of the nine interventions in this subset were delivered by trained study staff, and one was delivered by trained volunteers. Eight of the contingent-rewards-based interventions were delivered within the context of routine prenatal care, while one was delivered outside of this context. When grouped by contact intensity, four interventions in this subgroup were categorized as level 2 (moderate intensity), while five were categorized as level 3 (high intensity).

Four interventions in this subgroup were based on a single theory, while five were not. Similarly, four interventions explicitly mentioned the name of a theory, while five did not.

Compared to other BCTs, contingent rewards had relatively little overlap with other BCTs. The most common behavior change techniques provided alongside

contingent rewards were BCT 1 (“provide information about behavior-health link”), which was included in four of the nine studies in the contingent rewards subset, and BCT 14 (“provide feedback on performance”), which was included in three of the nine studies. Two studies included BCT 9 (“provide instruction”), while BCTs 2, 3, 5, 7, 8, 10, and 11 were each included in one of the nine studies in this subset.

Assessment of heterogeneity

Due to the homogeneity of this subset of interventions, we were unable to perform subgroup analyses and meta-regression models on many of the moderators/covariates of interest. Additionally, due to the lack of overlap with other BCTs, we were unable to test multiple, theoretically-derived clusters of BCTs to compare their effectiveness. Our ability to test pathways of change was similarly limited by the small number of studies that measured changes in theory-relevant constructs. However, we *were* able to analyze whether the effect size for the subset of nine studies that included BCT was influenced by the provision of general information about the link between smoking and health (BCT 1), or by the provision of feedback on performance (BCT 14).

Behavior Change Techniques as Moderators: To determine whether the effectiveness of contingent rewards was influenced by the presence of other BCTs, a series of subgroup analyses were conducted on the sample of nine studies that provided contingent rewards. As stated previously, sample size constraints limited our analyses of BCT clusters, such that we were only able to assess BCTs 1 and 14 as potential moderators. The results of subgroup analyses revealed that neither BCT 1 (“provide information about behavior-health link”) nor BCT 14 (“provide feedback on

performance") moderated the effectiveness of contingent reward-based interventions, such that interventions that provided contingent rewards *plus* BCT 1 or BCT 14 were no more effective than interventions that provided contingent rewards alone. (Please see **Tables 3.4.1 and 3.4.2** for full results).

There are several potential explanations for these results. Previous research suggests that information about the harms of smoking may sometimes have a backfire effect, especially when it is provided in the absence of strategies to enhance self-efficacy to mitigate the negative effects. This backfire effect may explain why the provision of information about the link between smoking and health did not enhance the effectiveness of contingent rewards. It is possible that the increased intensity and/or frequency of participation required by interventions that provided feedback may have acted as a barrier for participants, which could explain why feedback did not enhance the effectiveness of contingent rewards.

Intervention/participant characteristics as moderators: Subgroup analyses were also performed to explore whether characteristics of the intervention and/or participants influenced the effectiveness of interventions that provided contingent rewards. Because of sample size constraints and homogeneity within this subset of studies (see descriptive statistics), we were only able to assess three categorical variables as potential moderators: 1) Assessed smoking in social network; 2) Referred participants to community resources; and 3) Contact intensity.

A random effects model revealed no significant difference in effect sizes between interventions that provided contingent rewards *and* assessed smoking habits in the participants' social network (n=3) (RR=2.43; 95%CI: 1.26-4.68) versus those that did

not assess smoking in the social network (n=6) (RR=3.02; 95%CI: 2.03-4.47), as evidenced by overlapping confidence intervals and a non-significant between-groups heterogeneity statistic (Q_b[1]=0.30, p=0.579). There were also no significant differences between interventions that provided contingent rewards *and* referred participants to community resources (n=3) (RR=2.81; 95% CI: 1.83-4.31) versus those that did not provide such referrals (n=6) (RR=2.83; 95%CI: 1.75-4.58) (Q_b[1]=0.001; p=0.982). Similarly, there was no evidence of a significant difference in the likelihood of achieving late pregnancy smoking abstinence when comparing moderate-intensity contact (n=4) (RR = 2.50; 1.62-3.86) to high intensity contact (n=5) (RR=3.36; 95%CI: 2.0-5.65) (Q_b[1] =0.726; p=0.394) among interventions that provided contingent rewards. Please see **Tables 3.3.3, 3.3.4, and 3.3.5** for full results.

Intervention and participant characteristics as covariates*:* Two continuous variables were assessed as covariates in univariate meta-regression models within the subset of 9 studies that provided contingent rewards: 1) gestational age at baseline, and 2) cigarettes per day at baseline.

The first random effects meta-regression model revealed that gestational age at study entry did not account for any of the between-study variance in late-pregnancy smoking cessation in the subset of studies that provided contingent rewards (R^2 analog = 0%), and the regression coefficient was not significant, indicating that gestational age at study entry did not significantly influence the likelihood of achieving late-pregnancy smoking abstinence among the subset of studies that provided contingent rewards. Please see **Table 3.4.6** for full results. The next model revealed that cigarettes per day at baseline accounted for 100% of the between-study variance in late-pregnancy

smoking cessation in the subset of studies that provided contingent rewards, but the regression coefficient was not significant. Additionally, only five studies were included in this model because of limited data on baseline cigarette consumption within this subset of studies. The results of this model suggest that baseline smoking habits did not significantly influence the likelihood of achieving late-pregnancy smoking abstinence among the subset of studies that provided contingent rewards. Please see **Table 3.4.7** for full results.

Effectiveness Ratios

Effectiveness ratios comparing effective uses of each BCT to total active uses of each BCT (effective uses: active uses) varied greatly, but only four techniques were found to be effective 25% or more of the time: BCT 1 (provide information on health-behavior link); BCT 8 (set graded tasks); BCT 15 (provide contingent rewards); and BCT 17 (agree to behavioral contract). Based on the effectiveness ratio, the most promising technique was BCT 15 (provide contingent rewards), which had a ratio of 2:3. Thus, for every three active uses of BCT 15, two of those were effective uses. This supports the results of the subgroup and meta-regression analyses, in which BCT 15 was found to be the only technique associated with increased effectiveness when compared to studies that did include BCT 15.

Notably, many of the most commonly used BCTs had the lowest effectiveness ratios. For example, BCT 7 (provide general encouragement) was used as an active technique in 12 interventions, but only one of those uses was effective, resulting in an effectiveness ratio of 1:12. Similarly, BCT 2 (provide information on consequences

[negative]) and BCT 11 (prompt specific goal setting) were each as active ingredients in 10 interventions, but only one out of ten uses was effective.

Also of note was the finding that BCT 8 (set graded tasks) was only used as an active technique in two interventions, but one of the two uses was effective. Because of its limited use, we were unable to calculate an effect size for BCT 8 in the main analysis. However, the effectiveness ratio suggests that it could be a promising but underutilized technique.

Finally, given that the application of theory was inadequate in most studies in this review, it is possible that the effectiveness of BCTs utilized by these trials was limited by poor implementation. Optimally, the selection of BCTs should be based on a solid theoretical foundation and should be linked to specific theoretical constructs (or targets of change). However, this practice was not employed by most interventions in the current review. It is possible, therefore, that the effectiveness of BCTs was limited by poor implementation and may reflect poor translation from theory to practical application, rather than a failure of the BCT itself.

Table 3.5.1 presents the effectiveness ratios for all behavior change techniques.

Discussion

The purpose of this chapter was to identify, isolate, and quantify the effects of distinct, theory-derived behavior change techniques in the published literature on prenatal smoking cessation interventions, with the goal of determining whether the use of certain BCTs was associated with better intervention outcomes. A secondary aim was to determine whether specific theories were supported based on the resulting evidence.

Main Results

Overall, the evidence was mixed with regards to use of behavior change techniques. The total number of techniques used was not associated with late pregnancy smoking abstinence, indicating that more is not necessarily better. This may have important implications for intervention design, as interventions utilizing more techniques are generally more costly and time-consuming, but may not provide any additional benefit. By identifying the most effective intervention components and leaving out the others, intervention planners may be able to save time and resources without sacrificing results.

Effect sizes were significantly larger for the treatment group than the control group for subsets of interventions that 1) provided information about the link between smoking and health [BCT 1]; 2) provided information about the negative consequences of smoking [BCT 2]; 3) prompted the formation of intentions to quit smoking [BCT 5]; 4) provided instructions [BCT9]; 5) prompted specific goal setting [BCT 11]; 6) provided contingent rewards [BCT 15]; 7) taught participants to use prompts and/or cues [BCT 16]; and/or 8) had participants agree to a behavioral contract [BCT 17]. Notably, many of the most commonly-used techniques, including the provision of encouragement, the use of follow-up prompts, and motivational interviewing, did not demonstrate evidence of effectiveness in comparisons with their respective control groups. However, in some studies, terms such as “motivational interviewing” were used loosely, thus making it difficult to actually evaluate the technique. Additionally, only one technique (provide contingent rewards) demonstrated evidence of effectiveness above and beyond other techniques, such that the effect size for interventions that provided contingent rewards

was significantly greater than the effect size for interventions that did not provide contingent rewards. It is possible, however, that examining BCTs in isolation may produce different results than examining them in clusters. For example, the use of follow-up prompts may not be effective on its own, but could be effective when applied in combination with techniques such as goal-setting and teaching participants to use prompts and cues. However, because of sample size limitations, we were unable to examine clusters of techniques in the current review.

The effectiveness ratios associated with each technique provide a descriptor indicator of how often a technique is used effectively relative to the frequency with which it is used as an active ingredient in interventions. Notably, several of the most commonly used active techniques had the lowest effectiveness ratios. Three techniques (BCT 2, BCT 7, & BCT 11) were used as active ingredients in 10 or more interventions, of which only one of these uses was effective, resulting in effectiveness ratios ranging from 1:10 to 1:12. On the other hand, setting graded tasks was only used as an active technique in two interventions, but one of these uses was effective, resulting in an effectiveness ratio of 1:2. For intervention planners choosing among a wide variety of techniques, effectiveness ratios provide a useful indicator that may help save resources through the identification of techniques with the highest relative likelihood of success.

Limitations

The current study employed univariate, single-predictor meta-regression to examine the association between behavior change techniques and the primary outcome of late-pregnancy smoking cessation. Like any statistical or methodological procedure, there are both strengths and limitations to this approach. These considerations are the

subject of ongoing debate in the meta-analysis community. As summarized in a 2011 event organized by the Royal Statistical Society, the benefits of meta-regression with multiple covariates or outcomes come at the price of making more assumptions that do not necessarily result in better inference (Jackson, Riley, & White, 2011). One common problem encountered in meta-analyses is that not all studies provide data on the same covariates and outcomes (Thompson & Higgins, 2001). As noted by Borenstein and colleagues, meta-regression—like simple regression—requires an adequately large ratio of studies to covariates in order to produce meaningful results (Borenstein, Hedges, Higgins, & Rothstein, 2009). If multiple covariates are used in the same model, meta-regression is typically not recommended when the sample size (number of studies) is small. While there is not an explicit rule delineating how many studies must be present for each covariate added to the model, Borenstein and colleagues recommend that each additional covariate should correspond to at least 10 studies. This standard was employed by Hysong (2009), who used single-predictor meta-regression models in a study of audit and feedback interventions aimed at improving health care service quality. Due to a lack of overlap between behavior change techniques, our sample size limited our ability to run meta-regression models with multiple covariates. Put differently, among the studies that included a behavior change technique *x* as an active ingredient, only a limited number of the same studies also included behavior change technique *y* as an active ingredient. Thus, the danger of overfitting the meta-regression model was a significant constraint in the current study. While it is possible to impute missing data, there are also limitations to doing so. For example, if data are missing due to non-random causes, estimating the missing

data using an assumption that data are missing at random can exacerbate publication bias and other biases (Jackson, Riley, & White, 2011).

Conceptually, this study sought to establish a starting point for further investigation. Thus, the findings should be interpreted as hypothesis-generating rather than as conclusive evidence. Similar statistical approaches have been used in other studies seeking to establish which components of interventions were associated with intervention effectiveness. For example, O'Brien and colleagues (2015) analyzed how individual behavior change techniques were associated with the effectiveness of physical activity interventions among older adults. The results of the meta-analysis revealed that feedback was the only behavior change technique that moderated intervention effectiveness, such that interventions that used feedback were more effective than interventions that did not use feedback. In another meta-analysis, West and colleagues (2010) identified behavior change techniques used by the English Stop Smoking Services and examined their association with intervention effectiveness (West, Walia, Hyder, Shahab, & Michie, 2010). In that study, the authors analyzed behavior change techniques individually to determine how each technique contributed to the success of the intervention. Similarly, Michie and colleagues (2009) utilized univariate, single-predictor meta-regression to examine the association between behavior change techniques and intervention outcomes in trials of healthy eating and physical activity interventions (Michie, Abraham, Whittington, McAteer, & Gupta, 2009). The results of the meta-regression models were then used to inform further analyses. Specifically, the five techniques found to be associated with intervention effectiveness in single meta-regression models were later analyzed in a multiple meta-regression model.

However, as the authors noted, the number of studies required to undertake such an analysis is much greater than the number of studies required for the single meta-regression analysis. In Michie and colleagues' meta-analysis, the broader subject area (healthy eating and physical activity interventions) yielded a much larger sample size compared to the current study. Given our much smaller sample size, we were limited in our ability to run such analyses. This limitation was also noted by Achterberg and colleagues' (2010) meta-analysis of behavior change techniques to promote healthy eating. Even working with a significantly larger number of studies than were included in the current review, the authors were unable to analyze combinations of behavior change techniques due to limited sample size.

Other researchers have noted that, in order to best understand mediators of intervention effectiveness, starting with single-component analyses and working towards more complex, multi-component analyses may be the most appropriate strategy (Bauman, Sallis, Dzewaltowski, & Owen, 2002). Given the goals of the current study, breaking down behavior change techniques into the smallest unit of analysis was deemed to be appropriate. However, future studies with larger sample sizes should explore these techniques in pre-specified clusters. Clusters of techniques may be conceptualized differently according to different theories, and specific search criteria could be used to maximize sample size. It is possible that, when used in combination, the effect size of some behavior change techniques would be different than the effect size when analyzed in isolation. However, it is also important to note that the number of active behavior change techniques employed was not associated with intervention effectiveness in the current study, suggesting that analyzing behavior change

techniques in clusters rather than individually may not have significantly changed the findings.

Theoretical Implications

The results of the meta-analyses and meta-regression models exploring the unique contribution of behavior change techniques to intervention effects have important implications for behavior change theory. Within the literature on prenatal smoking cessation interventions, little work has been done to refine and build theories that explain and/or predict smoking behaviors and cessation among pregnant women. The lack of primary research in this area, combined with inconsistent and sometimes poor reporting practices, limited our ability to test theorized mechanisms of behavior change. However, this review represents an important step towards improving the science through identification of limitations and challenges, and exploration of promising avenues for future research.

Due to the lack of overlap between active behavior change techniques and limited measurement of theoretical constructs, we were unable to analyze theoretically-linked clusters of behavior change techniques and theory-derived mediators as a test of key tenets of behavior change theories. However, we were able to provide preliminary evidence in support of certain behavior change theories based on our analyses of individual behavior change techniques, using Abraham & Michie's (2008) guidelines for linking specific techniques to their theoretical underpinnings. A similar approach has been used in previous meta-analytic reviews, including Albarracin and colleagues' (2005) review of interventions promoting condom use, though the most common approach involves starting with the goal of testing a specific theory or parts of

a theory (i.e., an a priori approach), rather than inductively exploring which theory or theories are supported by the results of the review (e.g., Dombrowski et al., 2012; Hardeman et al., 2002; Johnson et al., 2011; Michie et al., 2009; Smoak et al., 2006). To our knowledge, neither of these approaches have been applied in a meta-analysis of smoking cessation interventions for pregnant women.

Theoretical explanations for prenatal smoking cessation

Based on effect sizes and effectiveness ratios for each technique, BCT 15 (provide contingent rewards) was clearly the most effective technique for promoting late pregnancy cessation. This suggests that smoking cessation during pregnancy may be driven primarily by factors such as motivation and expectancies, although tests of theorized mediators of change will be necessary to further evaluate the mechanisms underlying the behavior change process. The following section reviews several theoretical perspectives that may explain the results of the current review.

The use of contingent rewards as a behavior change technique is often grounded in the principles of **operant conditioning theory**, which posits that behavior is a direct function of its consequences, including punishment and reinforcement (Skinner, 1953). According to Skinner, people are most likely to engage in a behavior if it is immediately followed by positive reinforcement, such as material rewards or encouragement. On the other hand, people are less likely to engage in a behavior (or, put differently, more likely to stop engaging in a behavior) when it is not rewarded or is punished. Thus, the promising results for interventions that provided contingent rewards may be explained by the principles of operant conditioning and related learning theories. Most of the studies in this subgroup provided rewards in the form of vouchers for groceries,

transportation, formula, or other necessities, or in some cases, cash or tokens that could be exchanged for cash. The effectiveness of BCT 17 (agree to behavioral contract) lends further support for the applicability of operant conditioning theory, given that written and/or oral contingency contracts strengthen and make explicit the link between behavioral performance and reinforcement, thus increasing the likelihood that the desired behavior will continue.

The effectiveness of contingent rewards as a behavior change technique could also be understood by examining the principles of **expectancy-value theories**, which posit that behavior change is a function of beliefs about the expected consequences of performing a behavior and the value assigned to those consequences (i.e., costs or benefits) (Eccles & Wigfield, 2002). Expectancy-value theories assume that individuals will engage in or change a behavior if they expect that the consequences of doing so will yield more personal benefits than costs. Thus, changing beliefs about the likelihood of behavioral consequences and the value associated with them can change the likelihood of behavioral performance. In the case of contingent rewards as a technique to promote smoking cessation, clearly defined rewards for performing the desired behavior may increase the perceived likelihood of reaping positive consequences for quitting smoking. Although the health benefits of quitting smoking are also positive, they are delayed consequences and thus may not have the same impact as immediate rewards.

The results may also be understood within the context of the **Information-Motivation-Behavioral Skills model**, which posits that behavior and behavior change are driven by three primary determinants: 1) Information about the

behavior and its outcome(s), as well as cognitive heuristics and related mental "shortcuts" that influence decision-making; 2) Motivation, which is comprised of personal beliefs and attitudes about a particular behavior and/or intervention outcome(s), as well as social motivation in the form of positive social norms and social support; and 3) Behavioral skills, or the specific skills and strategies needed to successfully perform and/or maintain a behavior (Fisher & Fisher, 1992; Fisher, Fisher, & Shuper, 2009). Of the eight techniques that demonstrated effectiveness in comparisons of treatment versus control arms, two were related to the provision of information (BCT 1 and BCT 2), two involved instruction or teaching (BCT 9 and BCT 16), and one involved enhancing motivation through the use of positive reinforcement (BCT 15).

It should be noted, however, that contingent rewards-based programs may reflect a variety of theoretical approaches, depending on the characteristics of the intervention. There is also evidence to suggest that factors involving the delivery process may play an important role in the effectiveness of incentives as a behavior change technique. For example, the provision of incentives may trigger processes related to social desirability stemming from the anticipation of contact with an intervention deliverer. This is supported by evidence from previous studies that have found that only a small proportion of financial incentives offered even in successful interventions are actually redeemed, indicating that intrinsic reward (rather than tangible or extrinsic reward) may be a driving force behind behavior change in rewards-based interventions (Kane, Johnson, Town, & Butler, 2004). The findings from the meta-analysis revealed that contingent rewards had a significant effect on late pregnancy smoking cessation as well as sustained abstinence in the postpartum period, indicating that the effects of

contingent rewards continued beyond the period during which tangible rewards were offered. Additionally, in comparisons of contingent rewards versus non-contingent rewards, treatment effects were greater for participants in the contingent-rewards group, suggesting that the contingency component contributes to intervention effectiveness above and beyond any effects of the reward itself (Heil, 2008; Higgins et al., 2004; Higgins et al., unpublished; Higgins, 2014; Tuten, 2012).

While we did not have sufficient sample size to perform quantitative analyses examining how reward schedule and type influenced the effectiveness of contingent reward-based interventions, a review of effect sizes and intervention characteristics reveals several important patterns. First, the largest effect sizes were observed in trials with more frequent distribution of rewards and more rigorous monitoring schedules (Heil et al.,2008; Higgins, 2004; Tuten et al., 2012). The trial conducted by Heil and colleagues (2008) employed a strict schedule of check-ins and rewards that began with daily monitoring (CO levels) for the first five days, then transitioned to urine cotinine monitoring twice a week for the next seven weeks, then weekly for four weeks, and every other week for the rest of pregnancy. Women set a quit date, and then reported to the clinic for the scheduled check-ins and immediate voucher distribution. Vouchers were dependent on biochemically validated abstinence, starting at a value of $6.25 and increasing by $1.25 per check-in. The maximum voucher size was $45, and a positive cotinine test reset the voucher back to its original value. Similarly, Tuten and colleagues (2012) employed a schedule that called for collection of urine and breath tests three times a week (Monday, Wednesday, and Friday) for the duration of the study period. In the treatment group, rewards were contingent on meeting smoking reduction targets

that escalated every 2 weeks for the 12-week period. By the end of the study period, the target goal was smoking abstinence. Women were eligible to receive vouchers starting at a value of $7.50 for the first reduction target, and increasing in value by $1/day for each consecutive target met, reaching a maximum of $41.50. If a woman missed one of the reduction targets, no reward was distributed and the voucher level was reset to the original amount. However, if the participant met the target reduction on five consecutive occasions, the voucher level went back up to the previously attained level. Finally, the trial conducted by Higgins and colleagues (2004) employed a very rigorous schedule that began with daily abstinence monitoring and reward distribution (for the first week), then moved to twice weekly (for the next 7 weeks), weekly (for 4 weeks), and then every other week until delivery. The initial voucher value ($6.25) escalated by $1.25 per consecutive negative specimen, up to a maximum value of $45.00.

Of the three contingent rewards-based interventions with non-significant effect sizes, two trials employed less frequent monitoring and distribution of rewards (Donatelle, 2000c; Ondersma, 2012). Ondersma and colleagues (2012) only required cotinine testing at prenatal care visits, and the total number of rewards was limited to five, distributed at least a week apart, up to $50 in total value. Donatelle and colleagues (2000c) only required monthly testing, and reward size was limited to $25/month. Both of those trials also used a fixed reward size, rather than increasing the size of the reward if smoking abstinence was maintained over time.

In the three trials conducted by Donatelle and colleagues (2000a; 2000b; 2000c), larger reward size (dollar amount) appeared to be associated with a greater likelihood of achieving late pregnancy smoking cessation. In the two trials for which the effect size

for late pregnancy smoking cessation was statistically significant, the reward size was $50/month (Donatelle et al., 2000a; 2000b). In the one trial for which the effect size was not statistically significant, the reward size was only $25/month (Donatelle et al., 2000c). Almost all other intervention characteristics were the same, so it is reasonable to hypothesize that the decrease in reward size may be associated with the smaller effect size.

Taken together, these results suggest that more frequent monitoring and distribution of rewards may enhance the effectiveness of contingent reward-based interventions. Additionally, increasing the value of rewards, contingent on smoking reduction or abstinence, also appears to be associated with increased effectiveness, These observations should be investigated further, however, as they are based on a qualitative review, not a quantitative analysis.

Limitations & Suggestions for Future Research

The results from the analyses in this chapter are subject to a number of limitations, including inconsistent and poor reporting practices in published trials, inability to test clusters of techniques due to lack of overlapping BCTs, the possibility of unspecified (and therefore, unmeasured) intervention content accounting for or influencing the observed effects, variation within BCT categories, and the subjective nature of coding (even when using structured, standardized forms).

Reporting practices

One problem we encountered was a lack of specificity in describing intervention components, including behavior change techniques and their implementation. Although we were able to establish interrater reliability when identifying behavior change

techniques, we had to make many decisions that included some degree of subjective judgment, and thus other researchers may come to different conclusions, even using the same coding scheme on the same sample. Additionally, the degree of subjectivity may vary by technique, given that reporting quality was better for certain behavior change techniques than for others (i.e., for techniques described more thoroughly, it was easier to apply the coding scheme in a more straightforward manner, with limited subjective judgment calls). For example, published reports describing contingent rewards and motivational interviewing generally included detailed descriptions of the characteristics of delivery, including the duration, frequency, and scheduling, as well as the fidelity of implementation. These descriptions provided enough clarity to identify the techniques and distinguish them from similar techniques. On the other hand, provision of information, instruction, and social support were often reported with limited detail and without distinct features distinguishing one technique from another. Often, terms such as “counseling” were used to describe a process that included the provision of information as well as some form of instruction and/or social support. In these cases, more subjectivity was required in our decisions, and we often coded for multiple behavior change techniques. It is possible, therefore, that in some cases (especially those involving the provision of information, instruction, and support) we coded for two behavior change techniques when only one was used. In other cases, we may have coded for one behavior change technique when in fact, two distinct techniques were used.

Variation within BCT categories

Another problem we encountered was a great degree of variability within categories of behavior change techniques. Each behavior change technique on the taxonomy is meant to represent a distinct technique, but in some cases, the category encompassed a wide range of intervention content. This was particularly true for descriptions of intervention content related to the provision of information and guidance. Almost every study provided some type of information about smoking cessation, in the format of brochures, booklets, mailers, videos, and/or tailored materials. However, the quality and quantity of such information varied greatly, as did the mode of delivery and deliverer. For example, the most common type of information-delivery was in the form of written material provided to all women, including those in the control arm. In some trials, the written material was mailed to participants; in other trials, it was given to women at prenatal care visits. Additionally, in some cases, written information was given at the beginning of the trial only; in other cases, it was distributed throughout the trial, sometimes using multiple modes of delivery (e.g., in person at the start of the trial and by mail later on). While we coded for intensity, primary deliverer, and main mode of delivery for the intervention, we did not code for these variables for each behavior change technique. It is possible, therefore, that these factors may have influenced the effectiveness of the techniques. Future studies should explore promising behavior change techniques in more detail to identify potential parameters or moderators of effectiveness. This could be achieved through more focused meta-analyses exploring only one technique or a pre-determined group of techniques, rather than the range of all possible techniques.

Unspecified intervention content

The presence of unspecified intervention content is another limitation that could have influenced the results of the analysis. This is particularly true given that we were unable to assess specific pathways and mechanisms of change (e.g., the pathways through which behavior change techniques influenced behavior). While we isolated each behavior change technique in the analysis, we cannot rule out the possibility that unspecified content accounted for some of the observed effects. It is possible, for example, that the effectiveness of BCT 15 (contingent rewards) was due to the increased intensity and frequency of contact associated with contingent rewards-based interventions. Compared to other techniques, contingent rewards were delivered more frequently and on a much more structured schedule. Furthermore, the provision of contingent rewards may often include subtle or indirect elements of other behavior change techniques such as social support, feedback, and review of behavioral goals. Although these techniques were not discussed in published reports and thus were not coded as active techniques in the review, it seems likely that the effectiveness of contingent rewards may be due at least in part to processes of behavior change stemming from separate but related (unmeasured) techniques. It is also possible that intervention deliverers provided informal instruction and/or information that was not reported in the published trial, but may have influenced the effectiveness of contingent rewards. This may be true for other intervention techniques, as well. In future studies, detailed process evaluations may allow researchers to identify and account for unspecified and unmeasured intervention content. Again, meta-analyses focusing on one technique or group of techniques could explore these questions in further detail.

Identifying Active BCT's

In order to identify “active” behavior change techniques, we only included a technique as an active ingredient in the treatment arm if it was not present in the control arm or if it was present only in a lower dose in the control arm. While this allowed us to account for the control group condition, it may have limited our ability to test for clusters of techniques used in the treatment arm, as it reduced the total number of BCTs identified as active ingredients in the analysis. Some studies included many behavior change techniques in both the control and treatment arms; thus, when we excluded those techniques included in both arms, the number of active techniques was much lower than the total number of techniques. It is possible that the effects of active techniques could interact with the effects of “inactive” techniques (i.e., those included in both the treatment and control arms). We acknowledge this as a limitation and suggest that in future studies, analyses should include an exploration of all techniques included in the treatment arm as well as an exploration of techniques identified as active ingredients.

CHAPTER 6

Overall Discussion

Smoking remains one of the leading preventable causes of adverse maternal and fetal health outcomes, and thus represents an important target for behavior change interventions. Effective behavior change interventions have the potential to significantly reduce poor pregnancy outcomes, as well as to improve the health of women and children by promoting long-term smoking cessation. However, using the current best practice standard of brief counseling, only about one out of every 20 pregnant women quits smoking, and relapse remains a significant challenge (Lumley et al., 2009; U.S. Department of Health and Human Services, 2020). Furthermore, even when interventions are found to be effective, it is often difficult to identify which intervention techniques are responsible for promoting observed changes in behavior, and whether these effects are dependent on characteristics of the intervention, participants, or environmental context (Chamberlain et al., 2013; Lumley et al., 2009).

The purpose of this project was to evaluate the effectiveness of prenatal smoking cessation interventions, with a specific focus on advancing our understanding of *what* worked, *when* it worked, and *why* it worked. The first aim was to conduct a meta-analysis to produce quantitative estimates of intervention effect sizes and to identify factors that may explain the observed heterogeneity in intervention effectiveness. The second aim was to evaluate the use of the health behavior theory in intervention design, implementation, and evaluation, and to assess whether the use of theory was associated with intervention effectiveness. The third aim was to isolate the "active" ingredients in prenatal smoking cessation programs by first applying a

standardized taxonomy of behavior change techniques to identify the techniques, then quantifying the effectiveness of each individual technique. The overarching goal of this project was to build upon and expand the existing literature on prenatal smoking cessation by combining recent developments in intervention categorization and specification with meta-analytic methods to facilitate a more thorough exploration of the mechanisms of change underlying prenatal smoking cessation interventions, with the aim of informing better intervention design and ultimately, helping more pregnant women quit smoking through the use of evidence-based behavior change techniques. The following section presents a summary of the main results and key findings, followed by a discussion of the implications, limitations and considerations, and recommendations for future research.

Summary of Main Results

This project involved three primary steps. First, a meta-analysis was conducted to produce quantitative estimates of the effectiveness of published reports of prenatal smoking cessation interventions. The meta-analytic review served as the starting point for the next two steps, which involved the use of standardized frameworks and coding schemes to extract data for the purpose of answering new research questions that were not addressed in the primary studies. In the second step, we used a coding scheme to evaluate the use of behavior change theory in the design, implementation, and evaluation of prenatal smoking cessation interventions, and then examined the relationship between the use of theory and intervention outcomes. In the third step, we used a validated coding scheme to identify behavior change techniques used in prenatal smoking cessation interventions, and then used meta-analytic methods to

explore the effectiveness of the individual techniques. Each step in this project yielded several key findings, with important implications for research and practice. These findings are described below.

Meta-Analysis Results

The sample for the meta-analysis included a total of 38 trials representing over 12,000 pregnant smokers. The primary outcome of interest was late pregnancy smoking abstinence (defined as point prevalence abstinence measured anywhere from 28 weeks of pregnancy through birth), but additional outcomes including smoking reduction and perinatal health were also assessed when possible. Looking at the primary outcome, a random effects meta-analysis model revealed a significant effect in favor of the treatment groups, such that women in the treatment arms were 1.53 times as likely to achieve smoking abstinence before giving birth relative to women in the respective control groups.

Effect sizes for late pregnancy smoking abstinence varied according to several study-level characteristics, including intervention type, setting, and participant socioeconomic status. With regards to intervention type, incentives/rewards-based interventions were found to be the most effective category of intervention. Women in the treatment arm of incentives/rewards-based interventions were 2.82 times as likely to achieve late pregnancy smoking abstinence than women in the respective control groups. This finding is in line with other recent studies, which have identified incentives as the most promising approach to promoting smoking cessation in pregnancy (Bauld & Coleman, 2009; Lumley et al., 2014). With regards to setting, interventions delivered within the context of routine prenatal care were found to have a larger effect size than

those delivered outside of routine prenatal care. With regards to participant characteristics, interventions delivered to primarily low-SES women were found to have a larger effect size than interventions delivered to non-low-SES women. This is a particularly notable finding, given that low-socioeconomic status is often identified as a predictor of continued smoking during pregnancy and lower quit rates in prenatal smoking cessation programs (McLeod, 2004; Tong et al., 2009; Tong et al., 2013).

In addition to the primary outcome of late pregnancy smoking cessation, significant results in favor of the treatment group were also found for other measures of smoking behavior including significant reduction in smoking (by at least 50%) and point prevalence abstinence in the early and late postpartum periods. Lastly, the results of the meta-analysis also indicated that prenatal smoking cessation interventions were associated with a significantly lower risk of both low-birthweight and preterm birth deliveries.

Use of Theory

In the second step of the project, we used Michie & Prestwich's (2010) Theory Coding Scheme (TCS) to evaluate the use of behavior change theory in the design, implementation, and evaluation of prenatal smoking cessation interventions in the sample of 38 published reports derived from the meta-analysis search strategy. The TCS classifies theory-use into three main categories, according to function: 1) Selecting specific behavior change techniques or combinations of techniques to target specific theoretical constructs; 2) Informing the selection of participants who are likely to benefit from the intervention; and 3) Tailoring the intervention to individuals based on theory-relevant characteristics. It also assesses whether the published report mentions

a theory or theory-relevant constructs; whether the intervention was based on a single theory; whether and how theory-relevant constructs were measured; whether the intervention led to significant change in at least one relevant theoretical construct; whether meditation analyses were conducted, and if so, whether a change in the mediator predicted a change in the outcome variable; whether the results were discussed in relation to theory; whether the study provided support for or refuted a theory or theories; and whether the results were used to refine theory. Five composite scores and a total score were calculated to reflect the degree to which theory was used for various purposes, as well as the degree of overall theory-use. The names of theories mentioned in published reports were also recorded.

On a scale of zero (no use of theory) to 15 (optimal use of theory), total theory scores ranged from zero to 11, with a mean score of 5.05. Composite scores were highest for the measures reflecting whether theory or relevant theoretical constructs were mentioned (Mean = 1.55 on a scale of 0-3), and whether relevant theoretical constructs were targeted in the intervention (Mean = 2.5 on a scale of 0-5). Overall, 68% of studies (n=26) mentioned a specific behavior change theory, even if it was not actually used to inform the intervention. The most common theories mentioned by studies in this review were the Transtheoretical/Stages of Change Model (n=13), Social Cognitive/Social Learning Theory (n=8), and Operant Conditioning (n=7).

While many studies mentioned theory, far fewer studies actually utilized theory to inform intervention design, and most did not utilize theory in an optimal manner. For example, only 24% of studies (n=9) were based on a single theory rather than multiple theories or a combination of theoretical predictors. Interventions based on a single

theory are typically considered optimal, as the use of multiple theories and/or combinations of theoretical predictors can make it difficult to test and refine theory by obscuring theorized pathways of behavior change (Michie & Prestwich, 2010). Similarly, while 33% of studies (n=25) reported using theory or theoretical predictors to inform the selection of intervention techniques, only one study reported an explicit link between all intervention techniques and at least on theory-relevant construct or predictor, and only two studies reported targeting all of the theoretical constructs within a specified theory (or all theoretical constructs mentioned in the study) with at least one behavior change technique.

Few studies followed optimal guidelines for measuring relevant theoretical constructs. Optimally, theorized mediators of behavior change would be measured pre- and post-intervention; at a minimum, theorized mediators must at least be measured post-intervention to facilitate theory testing. Only 13% of studies (n=5) included post-intervention measures of theoretical constructs, and only two of these studies used measures that were previously validated and included evidence of their reliability. As such, the ability to test theorized mediators and mechanisms of behavior change research was limited. Only three trials presented evidence that the intervention produced significant changes in one or more theoretical constructs or predictors in favor of the treatment group, but none of the studies in the review provided evidence, through mediation analyses, that smoking outcomes were explained by these changes. Thus, while these studies found significant changes in theoretical constructs/predictors associated with the intervention, they did not provide evidence that these variables accounted for observed changes in smoking behavior. As such, none of the studies in

the review provided evidence that directly refuted or supported a theory, and none attempted to refine the theory upon which the intervention was based.

Contrary to the results of previous studies, our review did not find that greater use of theory was associated with greater intervention effectiveness. In a univariate meta-regression model, overall TCS score was not significantly associated with the effect size for late pregnancy smoking cessation. However, given that theory was rarely used in an optimal fashion, these results should be interpreted cautiously and should not be taken as evidence that the use of theory is unrelated to intervention effectiveness. It is more likely that these findings are a function of the underutilization of theory, such that simply mentioning theory and/or using it minimally or in piecemeal fashion is unlikely to enhance intervention effectiveness.

Behavior Change Techniques

In the third major phase of this study, we applied Abraham and Michie's (2008) 26-item taxonomy of theory-derived behavior change techniques to the sample of 38 published trials of prenatal smoking cessation interventions. First, we identified the techniques in published descriptions of intervention content. Next, we identified the *active* techniques, defined as those techniques which were present in the treatment arm and not in the control arm, or delivered in a more intensive dose in the treatment arm than in the control arm. Finally, we used subgroup and moderator analyses to quantify the effectiveness of each technique, and calculated effectiveness ratios to reflect the number of active uses of a technique in relation to the number of effective uses.

In subgroup analyses comparing the treatment group to the control group, eight BCT's were associated with a significantly larger effect size for late pregnancy smoking cessation, in favor the treatment group:

- BCT 1: Provide information about the link between smoking and health
- BCT 2: Provide information about the negative consequences of smoking
- BCT 5: Prompt the formation of intentions to quit smoking
- BCT 9: Provide instructions
- BCT 11: Prompt specific goal setting
- BCT 15: Provide contingent rewards
- BCT 16: Teach participants to use prompts and/or cues
- BCT 17: Have participants agree to a behavioral contract

In moderator analyses comparing interventions that used each of the eight BCTs mentioned above to interventions that did not use the BCT, only one technique (BCT 15: provide contingent rewards) demonstrated evidence of a significant moderating effect. Specifically, the effect size for late pregnancy smoking cessation for interventions that provided contingent rewards was significantly larger than the effect size for interventions that did not provide contingent rewards. This supports the results from the meta-analysis conducted in step 1, which found that incentives/rewards-based interventions were the most effective category of intervention.

Contingent rewards also had the most promising effectiveness ratio (ER), with 6 effective uses out of a total of nine uses (ER = 2:3). Three other behavior change techniques were effective in 25% or more of their active uses:

- BCT 1: Provide information on the health-behavior link (ER = 1:4)

- BCT 8: Set graded tasks (ER = 1:2)
- BCT 17: Agree to behavioral contract (ER = 1:3)

Implications

More is not always better

While it is often assumed that higher intensity interventions and the use of more behavior change techniques are positively associated with intervention effectiveness, the results of this review did not find support for that proposition. Intervention effectiveness did not vary by level of intensity, and the total number of active behavior change techniques was not associated with intervention effectiveness, nor did it explain any of the between-study variability in effect sizes. These findings indicate that the effectiveness of prenatal smoking cessation interventions is not a function of the quantity of intervention content, but rather the quality of intervention content. Currently, intervention techniques tend to be chosen without a clear rationale for the selection of specific techniques or combinations of techniques, and in some cases, more techniques are used in the hopes that *something* will work. However, as other researchers have noted, there may be a limit to what women will accept in terms of intervention intensity, and pushing this limit risks lowering participation, adherence, and/or compliance rates (Chapman, 2012).

The most common BCT's may not be the most effective

The purpose of calculating effectiveness ratios was to assess the frequency of active BCT uses relative to the frequency of effective BCT uses. Effectiveness ratios provide important context that isn't calculated by risk ratios alone, by accounting for how often a BCT was used overall versus how often a BCT was used successfully. Notably,

the findings revealed that some of the most common techniques had the lowest effectiveness ratios. For example, the provision of information on the negative consequences of smoking/not quitting (BCT 2) was identified as an active technique in 10 interventions, but only one of those uses was effective (when comparing the treatment group to the control group), resulting in an effectiveness ratio of 1:10. Similarly, goal-setting (BCT 11) was used as an active technique in 10 interventions, but only one of those uses was effective, and the provision of general encouragement (BCT 7) was used as an active technique in 12 interventions, with just one effective use. Thus, just as the most common behavior change theories do not necessarily have the strongest empirical support (Sutton, 2000), it should not be assumed that the most common behavior change techniques are the most effective.

Advancing the state of behavior change theory

Behavior change theories postulate that changing the causal determinants of behavior (i.e., theoretical constructs) will promote behavior change (Hardeman et al., 2005; Michie et al., 2008). Theories are a useful tool for intervention planners, as they can be used to identify the key theoretical constructs and mediators to target in an intervention as well as the mechanisms of action that make specific behavior change techniques work. Theory can also be used to design evaluations that facilitate the exploration of how, why, and when interventions succeed or fail at effectuating behavior change. The use of theory-based research also allows for the application and integration of evidence across different populations, contexts, and even behaviors by specifying the common mechanisms underlying behavior change. As such, the use of theory is widely recommended for researchers and intervention planners alike.

Despite the calls for greater use of theory in behavior change research and practice, many published intervention trials still make no reference to a theoretical basis (Albarracin et al., 2005; Davies, Walker, & Grimshaw, 2010; Hardeman, Johnston, Johnston, Bonetti, Wareham, & Kinmonth, 2002) and almost none explain how theory was actually used to inform the design, implementation, and/or evaluation of the intervention (Davies, Walker, & Grimshaw, 2010). As a result, most studies do not yield evidence that can be used to refine existing theories or build new ones. Additionally, the potential to accumulate and evaluate evidence across contexts, populations, and behaviors is limited when the use of theory is absent or poorly specified, which is a barrier to scientific and clinical progress.

Understanding how theory is being used (or not used) in a given field is an important step towards improving its use and, ultimately, advancing the state of theory and intervention science. Thus, we sought to explore the use of theory in published trials of prenatal smoking cessation interventions in an effort to identify how and where theory could be used more optimally to advance research and practice. To our knowledge, this review is the first attempt to systematically evaluate the use of behavior change theory in the field of prenatal smoking cessation interventions. Using a coding scheme developed by Michie & Prestwich (2010), we evaluated how theory was used to select behavior change techniques targeting specific theoretical constructs, to inform the selection of participants most likely to benefit from the intervention, to tailor the intervention based on theory-relevant characteristics, and/or to guide the selection of measures of theoretical constructs. We also examined whether theory was tested or refined, and calculated a total score reflecting the degree to which theory was

mentioned and used in published reports of smoking cessation interventions. The results revealed that behavior change theories are not being used in their full capacity, leaving a great deal of room for improvement. While many studies mentioned specific behavior change theories, few studies actually described how theory was used to guide the development of the intervention, and none of the studies in the review attempted to test theorized mechanisms of behavior change.

In line with other reviews evaluating the use of theory in published research on health-related behavior change interventions (e.g., Painter et al., 2008; Grimshaw et al., 2007; Davies et al., 2010), we found that theory is often mentioned but rarely accompanied by a detailed explanation of why that specific theory was selected, or how it was used to inform decisions such as the selection of behavior change techniques and the theoretical constructs they are supposed to target. While over two-thirds of the studies in this review (n =26) explicitly mentioned and targeted predictors of smoking behavior (such as motivation, self-efficacy, and intentions), only half (n=19) reported an explicit link between at least one behavior change technique and at least one of the targeted predictors. Only one trial reported an explicit link between all intervention techniques and at least one theory-relevant construct or predictor. Similarly, 55% of studies (n = 21) reported targeting at least one of the theoretical constructs mentioned in the published report with at least one behavior change technique, but only two trials reported targeting all of the constructs within a specified theory (or all of the theoretical constructs mentioned in the study). When describing the selection of behavior change techniques, the authors rarely provided a thorough or theory-based explanation for why a specific technique or set of techniques was chosen. Additionally, only nine studies

used a single theory or set of constructs from a single theory, while 17 studies used multiple theories or a combination of constructs from multiple theories. Although it is not inherently bad to use multiple theories (and in some cases, it may be entirely appropriate to use different theories at different stages of the intervention), there was often no rationale given for the selection of more than one theory. Mixing and matching theories often results in overlapping and sometimes conflicting assumptions about behavior change, and the mechanisms of action become obscured. The result is a complex web of techniques and theorized predictors that do not add up to a whole theory, and which cannot be tested as a theory. This unsystematic approach may explain why composite scores reflecting the use of theory in the selection of intervention techniques and the targeting of theory-relevant constructs/predictors were not associated with intervention effectiveness, nor was the overall use of theory score.

Use of theory was most limited in the areas of measurement and theory-testing. Testing theorized pathways of behavior change depends on adequate measurement of theoretical constructs targeted in the intervention. For example, an intervention that uses incentives to promote smoking cessation by modifying outcome expectancies and motivation must measure outcome expectancies and motivation to examine 1) if they changed as a result of the intervention and 2) if changes in smoking behavior can be explained by changes in these variables. In this review, only five trials included post-intervention measures of targeted theoretical constructs, and only two used measures that were previously validated and demonstrated evidence of their reliability. Three trials provided evidence that the intervention produced significant changes in one or more theoretical constructs or predictors in favor of the treatment group. However,

none of the trials provided evidence, through mediation analyses, that smoking outcomes were explained by changes in these constructs/predictors. As such, the findings could not conclude whether or not the targeted theoretical constructs actually accounted for observed changes in behavior. Similarly, none of the trials in this review attempted to refine the theory upon which the intervention was based, as the results did not yield sufficient evidence for doing so. Finally, because of the limited number of studies that included post-intervention measures of targeted constructs, we were unable to examine theorized pathways of behavior change in the meta-analysis and thus could not determine *why* interventions were successful or unsuccessful.

Our evaluation of theory use in the prenatal smoking cessation literature identified several key areas for improvement. First, the use of theory must move from general discussions of behavior change theories and related constructs to detailed explanations of why a given theory was chosen to guide intervention design, and how it was used to inform decision-making. At a minimum, published reports should include the following information: 1) A detailed rationale for why the specific theory was selected instead of others (and if multiple theories are used, the authors should provide a rationale for this decision); 2) Evidence that the theory's key constructs are associated with smoking behavior; 3) A description of the behavior change techniques used in the intervention; 4) A description of how the interventions targets the theoretical constructs (i.e., the causal processes targeted by behavior change techniques); 5) A description of the theoretical assumptions that underlie the intervention (i.e., the process[es] through which behavior change is theorized to take place), optimally in the form of a detailed logic model; 6) A description of the parameters of effectiveness, or the conditions that

must be satisfied for the intervention to be effective (e.g., fear appeals are only effective when delivered to populations with high self-efficacy, and may even be counterproductive when delivered to populations with low self-efficacy) (Kok et al., 2016; Peters, Ruiter, & Kok, 2013); 7) A description of if and how the theory was used to select participants and/or tailor intervention content for specific groups of participants; and 8) A description of how and when the theoretical constructs targeted in the intervention will be measured; and 9) A description of how the evaluation will test theorized mechanisms of change. While limitations on page length in scientific journals make it difficult to include such information in the body of published articles, there are several potential solutions to this problem. First, researchers may choose to publish this information in its own standalone article, which could then be referred back to in future publications, as we encountered several times while conducting this review. Alternatively, this information could be included as supplementary material and published online alongside the primary article. More broadly, academic journals could encourage better reporting practices by requiring the publication of intervention protocols and related information as online supplementary material before allowing the publication of additional studies, including outcome and impact evaluations.

In addition to more rigorous use of theory and more detailed reporting on how theory was used to inform intervention design and evaluation, there is a need for more research focused on theory testing and theory comparison. Theory-testing research provides the basis for understanding the mechanisms underlying behavior change, and is necessary for refining and building theory, as well as rejecting existing theories and developing new ones. Theory-testing research can answer important questions such as

whether the addition of a new construct to an existing theory adds to the utility of theory, and whether removing a construct has any effect on the theory's explanatory or predictive power.

Theory-comparison studies can help integrate separate lines of research and lead to a greater understanding of the process of behavior change than research on any single theory alone, and can thus provide critical insight about when a particular theory may be most appropriate, whether a specific theory is a better fit than others, and for whom a particular theory may be more effective than others. This type of research could answer questions such as whether different theories are needed to inform interventions at different stages of the smoking cessation process. For example, theories that explain and predict smoking cessation among pregnant women may not be appropriate when applied to relapse-prevention or cessation maintenance during the postpartum period. Recent research suggests that motivational factors may be more relevant during the process of trying to quit, while self-regulatory processes may be more relevant to maintaining smoking abstinence (Herd, Borland, & Hyland, 2009). Similarly, there is also evidence that the determinants of trying to quit smoking may be different from the determinants of successfully quitting (Borland, Yong, Balmford, Cooper, Cummings, O'Connor, et al., 2010). It is also possible, for example, that certain theories may be appropriate for explaining behavior change in light smokers, while other theories may be more appropriate for heavy smokers, who may require different and more intensive intervention to promote and maintain behavior change.

Based on our analysis of the most promising behavior change techniques (and the theoretical determinants they target), this review provides initial support for learning

theories such as operant conditioning, as well as expectancy value theories and the Information-Motivation-Behavioral Skills model, as potentially promising theories for explaining processes of change involved in prenatal smoking cessation. Although the Transtheoretical/Stages of Change model was commonly used, there was a lack of evidence supporting its use in this context and previous studies indicate that interventions tailored based on the stages of change are no more effective than interventions than do not include stage-based tailoring (Riemsma, 2003). This does not necessarily mean that the theory is inappropriate for use in the design of prenatal smoking cessation interventions, but it does indicate the need for further research examining how it is used, and whether its use is associated with intervention effectiveness.

Our evaluation of theory also highlighted a pattern of focusing primarily on psychological determinants of smoking cessation, with a notable lack of attention given to variables at higher levels of influence. For example, although several studies referred participants to community resources, most of them did so as an ancillary service that was not listed as a key intervention component and was thus not considered as a contributor to intervention effects. Additionally, even when the training provided to deliverers was described in detail and documented in process evaluations, it was not measured or categorized as an intervention component that might influence effectiveness. While psychological variables are certainly important contributors to smoking cessation, using theories that include higher-level constructs may help to better explain and understand smoking behavior during pregnancy. For example, social cognitive theory considers how factors in a person's physical and social environment

may influence their behaviors (and vice versa), while social ecological models explain individual behaviors within their interpersonal, organizational, community, and political/policy-related contexts. This may include assessing relationships such as how certain government policies make it easier or harder to purchase tobacco products or deliver smoking cessation interventions, whether workplace tobacco policies influence smoking behaviors, how shifting social norms shape smoking behaviors, or how poverty-related stress serves as a barrier to sustained smoking cessation.

Advancing the science of behavior change

As with most health behavior change interventions, prenatal smoking cessation interventions typically consist of multiple intervention strategies and techniques (Lumley et al., 2014). While this may contribute to the likelihood of promoting behavior change, it also makes it difficult to identify which intervention components are contributing to effectiveness. To date, evaluations of prenatal smoking cessation interventions have focused largely on whether or not interventions were effective, but have not looked at *what made them effective*. Identifying the active ingredients in interventions, or which techniques contributed to intervention effectiveness, has many important implications, including facilitating better intervention design, saving resources, reducing undue burden on participants, and refining behavior change theories. Advancing the science of prenatal smoking cessation (and other domains of behavior change) requires systems of synthesizing evidence. While standard meta-analytic methods contribute greatly to the accumulation of evidence, they are limited in terms of their ability to determine which behavior change processes are responsible for observed changes in behavior.

Given the significant heterogeneity in the outcomes of maternal smoking cessation programs, developing a more thorough and systematic understanding of the effectiveness of various behavior change techniques and the mechanisms through which they influence behavior could yield key insight for improving intervention design, evaluation, and synthesis. A critical first step in determining "what works" is to establish consistent terminology for describing intervention components and their relevant theoretical influences (Michie et al., 2011a; 2011b). This study sought to expand upon existing meta-analyses of prenatal smoking cessation interventions by incorporating recent developments in intervention categorization and specification to facilitate the identification of discrete behavior change techniques that contribute to the effectiveness of interventions. Using standardized definitions of behavior change techniques and other intervention components facilitates the accumulation of evidence and allows for the assessment of when, how, and why interventions worked. This line of research has the potential to address several major problems that have been noted in previous reviews and meta-analyses.

One such problem is that intervention content is not adequately described in the published literature (Dombrowski, Sniehotta, Avenell, & Coyne, 2007). Reviews of reporting practices of trials across numerous domains of behavior indicate that only 5% to 30% of published studies actually provide detailed descriptions of intervention content (Michie, Fixsen, Grimshaw, & Eccles, 2009). Furthermore, even when intervention content is adequately described, very few studies systematically measure the implementation of intervention content, and even fewer studies include an analysis of how intervention content is associated with intervention outcomes. Thus, there is a lack

of clarity regarding the specific components of interventions and how they are related to intervention outcomes. A related problem stems from lack of consistent and standardized terminology to describe intervention content. Even when published reports provide detailed descriptions of interventions, inconsistent terminology limits the accumulation of evidence, as the same terms may be used to describe very different concepts (and alternatively, different terms may be used to describe conceptually similar content). This not only limits the accumulation of evidence, but also the replication of effective behavioral interventions (Michie, Fixsen, Grimshaw, & Eccles, 2009).

The findings of the current review indicate that these problems are present in the literature on prenatal smoking cessation interventions. While there was significant variation in the quality of reporting, most studies did not describe intervention content in enough detail to be replicated by other researchers. Furthermore, while many studies included some type of process evaluation assessing implementation fidelity, most of these were informal, qualitative evaluations that could not be used for the purpose of quantitative evidence synthesis. Despite these limitations, we were able to reliably identify distinct behavior change techniques used in prenatal smoking cessation interventions. To our knowledge, this is the first attempt to isolate, categorize, and quantify individual behavior change techniques used to promote smoking cessation among pregnant women in the U.S., and thus represents an important first step that will help inform future research and practice.

New Literature

An updated literature search for relevant studies published after 2015 returned six new randomized controlled trials of smoking cessation programs for pregnant women that would have met our criteria for inclusion. These studies are briefly described below.

Forinash and colleagues (2018) used the transtheoretical/stages of change model to identify pregnant women in the preparation stage, then randomized them to receive standard care (pharmacist-driven education with or without nicotine patch or bupropion) or standard of care plus motivational text messages encouraging smoking cessation. Although quit rates were higher among women in the intervention group, the difference was not significant. However, as the authors noted, the study was underpowered and there was a high dropout rate, which may have made it more difficult to detect intervention effects.

Patten and colleagues (2019) developed and tested a phone-based biomarker feedback intervention for pregnant Alaska Native women. Intervention messages were based on social cognitive theory and designed to give women feedback on their baby's likely exposure to carcinogens. Participants were randomly assigned to receive three study calls (10-20 min each), either as part of the feedback intervention or as part of usual care. No significant differences in cessation were found between the two groups.

Abroms and colleagues (2017) tested a text message-based smoking cessation program, Quit4baby, in a sample of pregnant women already enrolled in an existing mobile health program. Text messages were grounded in social cognitive theory and designed to improve self-efficacy for quitting, describe the outcome expectations from quitting, increase social support via an ex-smoker "quitpal", and increase behavioral

capability for quitting. Texts were scheduled around enrollment into Quit4baby, the quit date, and the baby's due date. Significantly more women in the intervention group reported not smoking at the 1-month follow-up and 3-month follow-up periods, but biochemical verification of smoking status at the 3-month mark revealed no significant differences.

In another study of a text-based intervention, Abroms and colleagues (2017b) (Abroms, Chiang, Macherelli, Leavitt, & Montgomery, 2017) tested an automated program called SmokefreeMOM, which is specifically designed for pregnant smokers. Although it was highly rated by participants, the program did not produce any significant differences in smoking outcomes when compared to a control text message quitline.

Cummins and colleagues (2016) tested a telephone-based counseling intervention designed specifically for pregnant smokers. Women in the study were randomly assigned to the intervention group (telephone counseling plus self-help materials) or the control group (self-help materials only). The nine-session counseling program was designed to address pregnancy-specific topics such as "misunderstanding of health risks, perceived loss of control over timing of quitting, emerging self-image as a non-smoking parent, management of mood, and remaining smoke-free following the birth." The results of the program were promising, with intervention participants showing significantly higher abstinence rates than control group participants at the end of pregnancy and into the postpartum period.

Finally, in a trial of behavioral counseling supplemented by twice-daily doses of a medication called bupropion, Nanovskaya and colleagues (2017) found evidence that the combination of two techniques significantly reduced pregnant women's use of

tobacco products (Nanovskaya, Oncken, Fokina, Feinn, Clark, et al., 2017). The counseling component included 35-minute counseling sessions at each of the first 2 visits and 10 minutes of smoking cessation counseling at subsequent visits, and was designed to address cravings and withdrawal. Although the program helped women reduce their use of tobacco, there was no significant difference in abstinence rates between groups at the end of the intervention or at end of pregnancy.

Since this dissertation was first completed, the Theoretical Domains Framework and/or BCT taxonomy have been used in several studies to advance our understanding of smoking cessation among pregnant or postpartum women. None of these studies focused on the same research questions as this dissertation, nor do the results of those studies overlap with the results presented in this meta-analysis. However, the results from this new line of research do provide support for many of the findings in this dissertation, and the limitations encountered are remarkably similar.

Campbell and colleagues (2018) used the Theoretical Domains Framework to identify potentially effective BCTs related to known barriers and facilitators to smoking cessation during pregnancy. In consultation with 12 smoking cessation experts, the researchers came to a consensus on the barriers and facilitators most modifiable through behavioral support, then mapped existing BCT taxonomies against TDF domains to assess the extent to which BCTs used in existing interventions target key barriers and facilitators. The expert panel ranked ‘smoking [is] a social norm’ and ‘quitting [is] not a priority’ as the most important barriers and ‘desire to protect baby’ as a key facilitator to quitting. From a sample of 14 trials, the study identified 23 potentially effective BCTs targeting the key barriers and facilitators, most of which fell into one of

four TDF domains: ‘Social Influences’, ‘Knowledge’, ‘Emotions,’ and ‘Intentions’. Few potentially effective BCTs mapped onto every TDF domain, leading the researchers to conclude that key barriers and facilitators are “not sufficiently targeted” by BCTs used in existing smoking cessation interventions for pregnant women.

In an extension of Campbell and colleagues’ 2018 study, the same group of researchers conducted a modified Delphi survey to form an expert consensus on the potential influence (on behavior) of 34 pre-identified barriers and facilitators to smoking cessation during pregnancy, as well as the difficulty of addressing these barriers and facilitators (Fergie, Campbell, Coleman-Haynes, Ussher, Cooper, and Coleman, 2019). Forty-four practitioners with experience providing smoking cessation support to pregnant women were recruited for the study, which employed a three-round modified Delphi survey aimed at first forming an expert consensus on the influence of and difficulty of addressing 23 pre-identified barriers and 11 facilitators to smoking cessation during pregnancy, then identifying techniques to address the barriers and facilitators and forming a consensus on the appropriateness for their use in practice. The expert panel identified barriers and facilitators related to women’s motivation and self-efficacy, as well as the influence of significant others and social norms, as the most important in terms of their influence on smoking cessation during pregnancy. The panel considered having a supportive partner to be the most influential facilitator of smoking cessation, while a lack of partner support was the only barrier that reached consensus as being difficult to manage or address. In total, 14 of the 34 pre-identified barriers and facilitators were identified as being extremely or very important in influencing pregnant women’s smoking behavior, of which six were also identified as being very easy or easy to

address with existing BCTs. Despite reaching a consensus that barriers related to social norms were highly influential with regards to women's smoking behaviors, the study found that these barriers are very difficult to target and poorly covered by existing BCTs. As such, future research aimed at identifying and/or developing BCTs to effectively address social norm-related barriers to smoking cessation during pregnancy could prove to be of great importance.

In a meta-analysis looking at studies designed to improve health care providers' provision of smoking cessation care during pregnancy, Bar-Zeev and colleagues (2019) found that using audit and feedback and behavior change theories "may improve effectiveness," but concluded that it is still not clear which intervention components are most effective in improving smoking cessation care during pregnancy. The results of the meta-analysis did suggest that having 3 or more intervention components may be associated with increased intervention effects on specific care components. Similar to the findings of this dissertation, Bar-Zeev and colleagues also noted that the studies included in their analysis varied substantially in "design, intervention components, and outcome measurement," which impacted their ability to interpret the synthesized results, as did poor reporting of intervention content.

In a study that was described as the first review of BCTs to prevent postpartum relapse, Brown and colleagues (2019) conducted a study to identify BCTs and delivery modes used to prevent returning to smoking during the postpartum period. The researchers used the Behavior Change Technique Taxonomy, Version 1, to extract BCTs, then identified which were potentially effective by looking at which BCTs were both frequently occurring and used in interventions that had evidence of long-term

effectiveness. Out of 32 total trials, six demonstrated long-term effectiveness. The six effective trials all used self-help, sometimes in conjunction with counseling, and often delivered remotely. From those six trials, the researchers identified six potentially promising BCTs: ‘problem solving’, ‘information about health consequences’, ‘information about social and environmental consequences’, ‘social support’, ‘reduce negative emotions’ and ‘instruction on how to perform a behavior’. Additionally, the study found that tailored self-help approaches, with or without counseling, may be effective modes of delivery of BCTs aimed at preventing relapse during the postpartum period.

Other recent studies have focused on identifying the most effective behavior change techniques for modifying other risk-related behaviors during pregnancy. In a study that employed methods similar to those used in this dissertation, Fergie and colleagues (2019) examined RCT’s aimed at reducing alcohol consumption and illicit substance use during pregnancy, with the goal of identifying effective BCT’s and assessing the extent of theory use in intervention design and measurement. The researchers calculated effectiveness percentages to reflect potential effectiveness of each technique. These were calculated by dividing the total number of times a BCT had been a component of an effective component by the total number of times the BCT was used as an intervention component. Ultimately, 13 BCTs showed potential effectiveness for reducing alcohol consumption, and six of the nine alcohol trials reported using theory, but not extensively. None of the trials for reducing illicit drugs showed positive results. The BCTs that showed potential effectiveness included: Action planning, behavioral contract, prompts/cues, self-talk, offer/direct toward written material, problem

solving, feedback on behavior, social support (unspecified), information about health consequences, behavior substitution, assess current readiness and ability to reduce excess alcohol consumption, goal setting (behavior), and tailor interactions appropriately.

Limitations and Considerations

Applications of behavior change taxonomies

Behavior change taxonomies may be used to code for behavior change techniques specified in intervention and treatment manuals, published reports of interventions, or to actual implementation of techniques in an intervention setting (i.e., through direct observation). To our knowledge, this is the first use of the 26-item behavior change taxonomy in the published literature on prenatal smoking cessation interventions. We chose to use published reports of interventions because the vast majority of research projects and practical endeavors are based on evidence from the published literature (as opposed to treatment manuals or direct observation). Using published articles makes these findings more generalizable and applicable for researchers and intervention planners, but it also meant that coding and data extraction were based on less-than-optimal descriptions of intervention content. It is also important to note that, because we used published reports of intervention evaluations as the basis for evidence synthesis, it is possible that the lack of effectiveness associated with certain behavior change techniques was due to poor implementation fidelity, rather than the technique itself. While many studies included some type of process evaluation, the published reports did not go into detail about the implementation of individual behavior change techniques; rather, they focused on the delivery of the intervention as

a whole. As Abraham and Michie (2008) noted in their original reliability tests and reporting on the 26-item behavior change technique taxonomy, “Although larger samples are required to confirm this finding, the data indicate that pressure on journal space may curtail intervention descriptions in published articles. This may threaten replication fidelity because detailed manuals are not always accessible and are not presented in standardized formats. It also means that reviewers synthesizing findings on the basis of published evaluations may not be able to accurately and comprehensively identify intervention content” (p. 385). Thus, to deal with these practical realities, there may be a need to develop different taxonomies or at least different instructions for specifying intervention content based on published reports versus treatment manuals, given that published reports rarely adequately specify intervention content.

Choosing among taxonomies

The 26-item taxonomy used in this review is one of many different taxonomies available for specifying intervention content. The same group who developed this taxonomy has also developed a 43-item, smoking-specific taxonomy (Michie, Hyder, Walia, & West, 2011), as well as a 93-item taxonomy of behavior change techniques common to multiple domains of behavior (Michie, Johnston, Abraham, Francis, Hardeman, Eccles, Cane, & Wood, 2013).

Michie and colleagues conducted a review of treatment manuals from the English Stop Smoking Services and, using the PRIME theory as a guide, identified 43 techniques used to provide individual behavioral support for smoking cessation (Michie, Hyder, Walia, & West, 2011). This 43-item taxonomy was later used to specify the

content of smoking cessation behavioral support interventions as actually delivered in practice, based on transcripts of audio-recorded consultations delivered by the English Stop Smoking Services (Lorencatto, West, Seymour, & Michie, 2013). Although this taxonomy was developed specifically for smoking cessation, we chose to use the original 26-item taxonomy for two key reasons. First, the 26-item taxonomy reflects a broad range of theoretical approaches, while the 43-item smoking-specific taxonomy was developed based on one underlying theory. Given that many behavior change theories are relevant to the process of smoking cessation, we wanted to let the data lead to our conclusions about theory, rather than the other way around.

More recently, Michie and colleagues (2013) developed a 93-item, hierarchically clustered taxonomy of distinct behavior change techniques used in behavior change interventions. While this taxonomy provides a more comprehensive list of techniques, its practical application may be limited due to the volume of information and training required to reliably identify 93 different but often conceptually similar techniques. Additionally, many of these techniques are not used frequently in behavior change interventions, while the 26 core behavior change techniques were all identified at least five times across multiple domains of behavior (Michie et al., 2013). Thus, the 93-item taxonomy may be more appropriate for long-term projects attempting to classify all identifiable existing behavior change techniques, while a more parsimonious taxonomy may be a more practical and replicable tool for researchers seeking to code intervention content in meta-analyses and systematic reviews. In a recent attempt to apply the 93-item taxonomy to interventions aimed at preventing pediatric obesity, researchers were unable to establish adequate intercoder reliability, even after intensive training

(Jakicic et al., 2017). In order to reliably identify the techniques, the research team had to make significant changes to the taxonomy and coding protocol, including reducing the total number of techniques coded from the original 93 to 22 techniques that were identified at least once. While they specified their methods and rationale for modifying the taxonomy, there was no way to keep the original structure in place with so many major modifications. However, the feasibility of using the 93-item taxonomy has been demonstrated in more recent studies (Tate, Lytle, Polzie, Diamond, Leonard, Jakicic, et al., 2019).

The taxonomies discussed above are works in progress and are still being refined and improved through various applications and extensions, including additional guidance such as the behavior change wheel, as well as ongoing efforts to develop ontologies of behavior change techniques. Some have criticized the taxonomies because they do not explicitly link the behavior change techniques to features of theory such as the construct(s) targeted by the technique or the parameters of effectiveness (e.g., Peters et al., 2013; Peters & Kok, 2016). Additionally, the definition of a behavior change technique does not include evidence of its effectiveness, which Peters and colleagues (2013) cited as a weakness. Instead of taxonomies, they promote the use of an intervention mapping approach, which conceptualizes methods for behavior change as techniques or processes that have been demonstrated to change one or more determinants of behaviors (Bartholomew, Parcel, Kok, Gottlieb, & Fernández, 2011). Thus, unlike Michie and colleagues' behavior change taxonomies, the intervention mapping approach includes evidence of effectiveness in the definition of behavior change methods. The intervention mapping approach also specifies how theory-based

methods of behavior change can be translated for practical application for specific populations and contexts, and describes the parameters of effectiveness for each method (or the conditions that must be satisfied for successful application of a behavior change method) (Kok et al., 2016). As Peters and Kok (2016) note, the intervention mapping approach provides a series of practical steps, beginning with problem identification and moving towards problem-solving or mitigation, and including specific guidance for identifying theory-based determinants and matching them with appropriate behavior change methods. Kok and colleagues (2015) cite the lack of explicit specification of targeted determinants in behavior change taxonomies as a major limitation in both research and practice. However, the purpose of behavior change taxonomies is to establish a basic set of behavior change techniques, using standardized terminology and definitions, to serve as a basis for conducting research on the effectiveness of techniques and their hypothesized mediators. Thus, effectiveness is not included in the definition because there is still a need to identify and categorize techniques that may *not* be effective. Additionally, although the taxonomy does not explicitly link each technique with the determinant(s) it targets, it does include supplemental material with directions for coding determinants targeted by each technique. Finally, the taxonomy is meant to be used as a guide for categorizing intervention content for future research investigating the factors that may influence the effectiveness of each technique. Thus, it was designed as the starting point, not the endpoint, for identifying effective intervention techniques and parameters of effectiveness. Overall, intervention mapping may prove to be a more useful tool for

direct application to intervention design, while the behavior change taxonomy may be more appropriate for research and classification purposes.

Evaluating use of theory

Theory may be used in a number of different ways to inform the design, implementation, and evaluation of interventions. The Theory Coding Scheme (TCS) attempts to capture many different uses of theory, but it does not go include a detailed evaluation of these various uses. For example, while the TCS includes an assessment of whether some or all theoretical constructs are targeted by behavior change techniques, it does not evaluate the appropriateness or suitability of these links. Studies evaluating specific applications of theory may yield more descriptive findings about the best ways to use theory to improve intervention effectiveness. For example, Noar and colleagues (2007) evaluated the use of theory for tailoring print materials delivered in health behavior change interventions (Noar, Benac, & Harris, 2007). This study included a detailed evaluation of tailoring characteristics, such as the number and type of theoretical concepts tailored on, the type of print material, and whether demographics were measured and tailored on, as well as potential moderating characteristics. The study found that tailoring on three characteristics combined (theoretical constructs, behavior, and demographics) was more effective than tailoring on any of those characteristics alone or in pairwise combinations with each other. While the TCS assesses whether or not tailoring was not used, it does not include this type of detailed examination and thus yields less descriptive findings on specific applications of theory and how they are associated with intervention effectiveness.

Additionally, different approaches to using theory may be appropriate depending on the goals of the research or practical application. As Lippke and Ziegelmann (2008) note, when the goal is advancing theoretical knowledge, interventions based strictly on one theory may be most appropriate, but when the goal is maximizing the intervention effectiveness, using several theoretically derived behavior change techniques from multiple theories may be the most appropriate approach. Thus, while the TCS considers interventions based on one theory to be optimal, there may be situations when using multiple theories is more appropriate.

Similarly, the TCS does not evaluate what type of theory was used (e.g., explanatory or change theory). For selecting behavior change techniques to use in an intervention, change theories such as the Transtheoretical Model may be more directly applicable, but explanatory theories such as the Health Belief Model may be more useful for understanding the processes of change. Depending on the goals of the study, both types of theory may be appropriate to use at the same time. For example, an explanatory theory may be used to identify key determinants of change to target with behavior change techniques, while a change theory could be used to tailor messages and intervention content. Along the same lines, it may be appropriate to combine an individual-level theory with a broader ecological model to identify and target behavioral determinants at all levels of influence.

Our ability to test theoretical mechanisms of behavior change was limited by several factors, including variability in measurement and specification of components of behavior change theories. Identifying theoretical mediators of behavior change requires pre- and post-intervention measurements of the theorized mediators. These

measurements can then be used to determine whether observed changes in behavior can be explained by changes in the mediating variable. Unfortunately, most published trials of smoking cessation interventions did not include measures of theoretical constructs during the pre- and post-intervention period. Thus, we were unable to examine theorized mechanisms of behavior change. As a result, we were unable to draw conclusions or make specific recommendations with regards to refining theory.

The use of meta-analyses to evaluate the utility of health behavior theories is subject to a variety of constraints. For example, theories are rarely used in their entirety, so important theoretical constructs may be omitted from intervention evaluations; as such, it may not be possible to test the full theory. In one meta-analysis of interventions designed to reduce sexual risk for HIV among adolescents, the investigators attempted to test the utility of the three components of the Information-Motivation-Behavioral skills model but were unable to test the information component because of limited variability in information provision (Johnson, Scott-Sheldon, Huedo-Medina, & Carey, 2011). The authors found that provision of motivation and behavioral skills reduced HIV risk through increased condom use, but because information was provided in nearly every intervention, they could not determine whether its inclusion was associated with increased effectiveness. Additionally, many behavior change theories include conceptually overlapping constructs. Examples of overlapping categories of theoretical constructs include: 1) self-efficacy, perceived behavioral control, perceived control, beliefs about capabilities, and confidence; 2) benefits/barriers and pros/cons; 3) attitudes, outcome expectancies, and beliefs about consequences; and 4) intention, motivation, and readiness for change. Lack of specification and inconsistent terminology

in the published literature makes it difficult to assess whether studies are measuring the same construct, or a similar (but conceptually distinct) construct. Measurement problems also present significant limitations. Common measurement problems encountered in this study and noted in previous reviews include the use of unreliable or non-validated instruments to measure theoretical constructs, the use of non-comparable instruments to assess similar theoretical constructs, the use of insensitive measurements with poor discriminative properties, failure to establish baseline measurements, and incomplete and/or inconsistent reporting of measurement instruments and psychometric properties, (Bauman, Sallis, Dzewaltowski, et al., 2002; Cohen, Underwood, & Gottlieb, 2000; Redding, Maddock, & Rossi, 2006; Wilson, Allen, & Li, 2006a; Wilson, Allen, & Li, 2006b). In other instances, intervention design and measurement may interact to create new challenges for evaluating theory in meta-analyses. For example, the provision of social support may not produce the same effects for individuals with different levels of baseline social support. Specifically, individuals with low levels of social support are likely to benefit from techniques designed to increase social support, while those who already have high levels of social support may experience little to no detectable benefits from additional support (Cohen, Underwood, & Gottlieb, 2000). Similarly, studies of interventions employing audit and feedback indicate that the technique may only be effective in motivated populations who endorse positive attitudes about making the intended behavior change (Jamtvedt, Young, Kristoffersen, O'Brien, & Oxman, 2006). If these constructs are not adequately and reliably specified and measured, as they often weren't in the published literature, it

is not possible to determine whether they influenced or mediated the effectiveness of intervention technique(s).

Future Directions

The results of the current study point to several areas of improvement that should be addressed in future research, as well as several opportunities to advance intervention design and expand on our existing knowledge of smoking cessation behaviors during pregnancy.

Areas for improvement

Reporting Practices: There is a significant need to improve reporting practices in the literature on prenatal smoking cessation interventions. Regardless of how many high quality studies are conducted, the state of science and practice cannot move forward if the research is not presented in such a way that is accessible to both investigators and practitioners. Currently, a lack of specificity in reporting on intervention content and high levels of variability in reporting practices make it difficult, and in some cases impossible, to reliably identify intervention components for purposes such as evidence synthesis. Similar problems exist in practices of reporting on the use of theory in intervention design and evaluation, as discussed previously. While recent advances in reporting guidelines and recommendations have certainly led to improvements in this area, there is still a great need for greater standardization and detail in reporting of intervention content and delivery, as well as how theory was used in each stage of intervention design and evaluation. Furthermore, incomplete outcome reporting was a limitation noted in several studies included in this meta-analysis. This often occurred when non-significant results were found. Instead of reporting the full results, authors

would sometimes simply note the non-significant findings in the text of the article. In other instances, authors combined the results of different intervention arms and failed to report the results of each intervention arm independently. Such practices are problematic for many reasons, including lack of transparency and inability to include results in meta-analyses, and thus should be avoided.

Measurement: Similar to the problems we encountered with reporting practices, we also found that variability in measurement of theoretical constructs greatly limited theory-testing and evidence synthesis. While it may not be practical to suggest that all evaluations of prenatal smoking cessation interventions use the same measures to evaluate theoretical constructs, developing more standardized and validated measures of core constructs would greatly enhance the quality and scope of future meta-analyses. Unless there is a clear need to do so, studies should also avoid using their own measures for theoretical constructs when validated measures already exist. Perhaps most importantly, evaluations of interventions should optimally measure all relevant theoretical constructs at the beginning, middle, and end of the intervention to assess whether the intervention actually produced changes in theorized mediators of behavior change. This will be an integral step to facilitate theory-testing research and enhance our understanding of how behavior chance actually takes place.

Use of Theory: The recent calls by granting agencies to use a theoretical framework in intervention design appears to have resulted in more studies citing a specific theoretical basis, but not actually using theory to guide the development of interventions. Future trials of prenatal smoking cessation interventions should seek to use theory to its full capacity, including to identify key determinants of behavior change,

select techniques to target those determinants, and, when appropriate, tailor the provision of techniques to specific participant characteristics. Theory can also be used to develop and guide the selection of appropriate instruments to measure key theoretical constructs. Granting agencies could encourage improvements in the use of theory by specifying requirements for researchers to provide a rationale for their choice of theory, as well as to describe how the theory was used throughout intervention development.

Standardizing Terminology: Using standardized vocabulary and definitions to describe and measure intervention components and theoretical constructs is a key step in advancing our understanding of the processes involved in prenatal smoking cessation and the components of effective prenatal smoking cessation interventions. Many studies in the current review discussed and measured conceptually similar concepts but used different labels to describe them. The same problem was found in descriptions of behavior change techniques. In other instances, umbrella terms were used to describe intervention techniques that actually encompassed multiple behavior change techniques (e.g., “counseling” was commonly used to describe techniques involving social support, information provision, instruction, and elements of goal setting, intention formation, and encouragement). Using taxonomies of behavior change techniques and theoretical constructs could help solve this problem, thus facilitating advances in evidence synthesis, theory building and refinement, and intervention science.

Multiple behavior change and risk factor research

Given the overlap between smoking and other risk behaviors and risk factors (e.g., mental health disorders, intimate partner violence, late entry to prenatal care), an

important step for future research will be to develop and evaluate interventions that target multiple risks simultaneously. Before this can be done, there is a need for research exploring conceptually similar and interacting behaviors and risk factors, as well as different approaches to intervening on multiple behaviors and risk factors at once. For example, it will be important to understand if intervention techniques should target behaviors sequentially or simultaneously. If a sequential method is identified as the most promising approach to behavior change, researchers will also need to determine the order in which behavior(s) or risk factor(s) should be targeted to maximize effectiveness. Another important step will be identifying common mediators that explain or predict changes in behavior, risk status, and/or key health outcomes. For example, perceived social support has been identified as a key mediator of change across numerous maternal health behaviors, such that low social support reduces the likelihood of (positive) behavior changes such as smoking cessation (Elsenbruch, Benson, Rucke, Rose, Dudenhausen, et al., 2006; McCormick, Brooks-Gunn, Shorter, Holmes, Wallace, & Heagarty, 1990). Low social support is also associated with other risk factors such as depression, and is associated with a higher risk of poor perinatal outcomes (Collins, Dunkel-Schetter, Lobel, & Scrimshaw, 1993; Feldman, Dunkel-Schetter, Sandman, & Wadwha, 2000). Similarly, evidence suggests that perceived stress may act as a mediator of smoking cessation and other behavioral changes during pregnancy. High levels of perceived stress during pregnancy are also associated with risk behaviors such as poor eating habits and high caffeine consumption, and also independently contribute to the risk of poor birth outcomes (Glynn, Schetter, Hobel, & Sandman, 2008; Lobel, Cannella, Graham, DeVincent,

Schneider, & Meyer, 2008). Thus, interventions targeting factors such as social support and perceived stress have the potential to promote behavior change and improve pregnancy outcomes through multiple mechanisms.

El-Mohandes and colleagues (2008) are taking the first steps in this direction in the area of maternal smoking and related risks. In an intervention designed to reduce psychosocial and behavioral risks, they used an integrated approach to target smoking, environmental smoke exposure, depression, and intimate partner violence among pregnant women. They concluded that targeting multiple risk factors contributed to the maintenance of behavior change in the postpartum period, likely by enhancing coping and behavioral modification skills. They also identified additional risk factors (alcohol and drug use) that appeared to interfere with the effectiveness of the intervention, leading to the recommendation that future interventions should target these risk factors in addition to the four risk factors targeted in the initial intervention. However, the researchers also warned that risks must be selected carefully to avoid overwhelming patients and/or providers with too much at once.

The use of standardized terminology to describe and define behavior change techniques, theoretical determinants, and other intervention components is particularly important in the area of multiple behavior change/risk factor interventions. Interventions targeting multiple risk factors inherently involve conceptually similar content. It will be important for researchers to resolve discrepancies such as the use of different terms to describe the same concepts and techniques, as well as the use of the same terms to describe distinct concepts and techniques. Similarly, it will be important to minimize variation in measurement by establishing validated measures with high discriminative

properties to assess these related concepts. Taxonomies of behavior change techniques may be of particular importance in this emerging line of research.

Behavior Change Ontologies

Given the complexity of this field of study, systems of organizing and accumulating evidence are needed to facilitate the synthesis and comparison of findings across different studies. As mentioned previously, the behavior change taxonomy utilized in the current study is meant to establish a common vocabulary, but does not include the effectiveness of techniques or the interrelationships among them as part of the definition of a behavior change technique. An important line of future research will be to use this taxonomy as the basis for developing behavior change *ontologies* that seek to answer the question, “What works to change behaviors, for whom, in what situations, how, and why?” (West & Michie 2016, as cited by Larsen et al., 2017, p. 16). An ontology is a clearly defined, shared vocabulary of terms and the specific relationships between those terms (Srivastava & Shu, 2014). Behavior change ontologies link together behavior change techniques, mechanisms of actions, target behaviors, context, and usage, as well as the interrelationships between and among these classes (Larsen et al., 2017). Each one of the aforementioned classes could be depicted by its own ontology, and then combined in a unified ontology of behavior change, as proposed by Larsen and colleagues (2017). An ontology of behavior change techniques would include a controlled vocabulary of behaviors, problem types (e.g., starting a new behavior, stopping an existing behavior, modifying the level of engagement in a behavior, or maintaining behavioral performance), behavior change techniques, the mechanisms of action and targets of change associated with each

technique, and the mediators of effectiveness such as context, dose, delivery, and participant characteristics (Larsen, Michie, Hekler, Gibson, Spruijt-Metz, Ahern, et al., 2017; Srivastava & Shu, 2014). Applied examples of behavior change techniques/clusters and their targets could also be included to help intervention designers generate ideas for techniques and targets of change that are relevant for specific behaviors and problem types. This information could then be codified into a format that is computer accessible and readable in order to facilitate efforts to refine the ontology, to encourage collaborative knowledge generation and evidence accumulation, and to provide a searchable "toolbox" for intervention designers. Ultimately, the goal would be to develop an ontology that could be searched using specific syntax to define the parameters of a query, much like the functions employed by Google Scholar, PubMed, and other search engines.

A behavior change ontology would be useful for both research and practical applications, and could contribute greatly to our understanding of the processes involved in behavior change by offering a platform for systematically collecting new evidence and storing data from different types of studies that might otherwise not be comparable (Srivastava & Shu, 2014). Importantly, ontologies also facilitate the collection, storage, and comparison of evidence from different fields of study. Currently, evidence from different fields tends to be stored in isolated bodies of literature, with each field advancing on parallel tracks. The same pattern exists in the literature on behavior change theories, with a separate body of research for each theory and few attempts to unify the literature. By offering a standardized format for the collection and storage of data, ontologies could help researchers compare the properties and

mechanisms of behavior change techniques, thus facilitating insights such as the identification of interchangeable techniques or the discovery of a certain attribute or attributes of techniques that may be particularly applicable to a specific type of intervention, behavior, or mechanism of change. For example, it may be found that behavior change techniques targeting motivation are key for interventions promoting the adoption of a new behavior, while techniques targeting outcome expectancies are key for interventions aimed at stopping an existing behavior. It may also be found that behavior change techniques targeting outcome expectancies are effective for certain domains of behavior, but are not sufficient (on their own) when used in the context of other domains, such as addictive behaviors. Similarly, ontologies could help researchers identify mechanisms that explain how behavior change interventions work, which may be particularly useful for identifying common and unique mechanisms of change in multiple behavior change interventions. For intervention designers, ontologies provide a readily accessible and systematic method of quickly identifying the behavior change techniques and targets of change that are most relevant for a given type of behavior/behavior change. As such, the practical application of ontologies by intervention designers would contribute to more effective behavior change interventions, as well as to the advancement of the ontology itself as evidence from such interventions is added back into the ontology.

A behavior change ontology could also help aggregate findings across different health behavior theories, and thus facilitate theory refinement and integration, as well as hypothesis generation (Larsen et al., 2017). While a variety of behavior change theories have been developed to explain the complexities of human behavior, this has resulted

in a problematic lack of shared terms and definitions, with each theory offering its own vocabulary to describe various constructs and the relationships among them. In a 2015 review of behavior change theories in the social and behavioral science literature, Davis and colleagues identified 1,725 different theoretical constructs across 83 theories, with a mean of 21 constructs per theory (and a maximum of *91 constructs* in one theory) (Davis, Campbell, Hildon, Hobbs, & Michie, 2015). Behavior change theories often share overlapping constructs with other theories, use different terms to describe the same constructs, and use different items and scales to measure the same constructs. Theoretical constructs are often inadequately defined, as are the relationships between different constructs, and between constructs and behavior. This greatly limits our ability to synthesize evidence, refine theory, and apply the theory to intervention design, implementation, and evaluation. Just as an ontology could facilitate the accumulation of evidence across behaviors and fields of study, it could also advance the integration of evidence across theories by providing a systematic method of articulating theoretical constructs, mechanisms, and the relationships between them.

The codification of ontologies into a computer-readable format is an important step that would allow researchers to use information science techniques such as Natural Language Processing to improve definitions and better specify the relationships among constructs. This, in turn, can be used for advances such as improving the measurement of constructs. In one recent study, researchers applied Natural Language Processing to examine similarities and differences among words and phrases used in measurement scales to determine whether the same construct label was being used to describe two conceptually distinct constructs (Larsen & Bong, 2016). The same

approach could also be used to assess whether two different labels are being used to describe the same construct. As demonstrated by Staunton and colleagues (2014) and described by Larsen et al. (2017), Natural Language Processing techniques can also be used to enhance meta-analyses when applied to the extraction of data and the comparison of different ways of structuring and organizing theoretical constructs. Using this approach, Natural Language Processing would allow researchers to extract operational definitions of constructs based on all available information from primary studies, and then link them to various labels representing different organizational structures. These structures, which represent different theoretical approaches to defining conceptually similar constructs (e.g., self-efficacy vs. perceived control vs. perceived behavioral control), could then be compared to determine the best fit (i.e., the most appropriate theoretical model). Finally, this information would be used as input for an "automated meta-analysis," which uses a bottom-up approach to extract all relevant details related to a study question, matches the extracted data with appropriate labeling structures, and then computes effect sizes to quantify the relationships among the various constructs and with other variables (e.g., demographic variables) (Larsen et al., 2017). This allows researchers to reliably and efficiently categorize and quantify relationships among theoretical constructs using descriptions in the published literature that may otherwise be incomparable, overlapping, or underspecified. Such an approach facilitates evidence aggregation and theory refinement through the identification of new relationships between constructs, as well as shared constructs found in two (or more) theories (Larsen & Bong, 2016). This approach has also been used to assess the appropriateness of effect size benchmarks for relationships across and within domains

of theoretical constructs, to identify appropriate effect size cutoffs for tests of theoretical relationships, and to inform better power analyses for theory-testing purposes (Bosco, Aguinis, Singh, Field, & Pierce, 2015).

An important and related step for future research will be to determine the parameters of effectiveness of behavior change techniques used in prenatal smoking cessation programs. As defined by Kok and colleagues (2016), parameters of effectiveness are “the conditions that must be satisfied in practical applications for the method to be effective” (p. 301). We sought to begin the process of identifying these parameters by evaluating whether the effectiveness of techniques was influenced by study-level characteristics, such as participant socioeconomic status or intervention intensity, but there is much more research to be done in this area. For example, future studies should seek to determine whether the effectiveness of providing contingent rewards depends on the delivery schedule or level of reward. West and Michie (2016) recommend including these parameters in the behavior change ontology.

It will also be important to incorporate theoretical explanations of behavior change into this process, as behavior change theories typically specify parameters of effectiveness. For example, the Theory of Planned Behavior specifies that intention formation depends on attitudes, subjective norms, and perceived control. Thus, it is unlikely that behavioral intentions can be changed without first addressing related attitudinal, social normative, and control beliefs. Similarly, the Health Belief Model specifies that behavior change is most likely to occur when high perceived risk and susceptibility are also accompanied by high self-efficacy. Exploring these theory-based parameters of effectiveness is an important step towards developing a more thorough

understanding of the conditions under which certain behavior change techniques work, and thus how they can best be utilized in practice.

Digital & Mobile Healthcare: As more and more healthcare institutions integrate digital technologies and e-health interventions into regular practice, there is a growing need to determine whether these innovative approaches can be applied effectively to smoking cessation during pregnancy, and if so, which techniques and delivery modes are most effective and for which populations, along with other similar questions about intervention effectiveness. These questions may become even more important in light of the ongoing, global COVID-19 pandemic, which has disrupted the delivery of health care and greatly increased reliance on remotely-delivered health care services. If this trend continues, it will be critical to understand if and how existing smoking cessation interventions for pregnant women can be adapted to be delivered remotely, and how this can be done to maximize effectiveness.

There is already work underway in this area that could be built upon in future studies. Recent meta-analytic research suggests that digital interventions can be used effectively to promote smoking cessation during pregnancy, with computer-based and text-message interventions showing the most promise (Griffiths, Parsons, Naughton, Fulton, Tombor, and Brown, 2018). In a meta-analysis of digital interventions for smoking cessation during pregnancy, Griffiths and colleagues (2018) identified seven BCTs associated with effectiveness: information about antecedents; action planning; problem solving; goal setting (behavior); review behavior goals; social support (unspecified); and pros and cons. Using a meta-regression model, the researchers also found that the number of BCTs used in digital interventions was positively associated

with the effectiveness of the intervention. This finding stands in contrast to the results of this dissertation, which found that more BCTs were not necessarily better.

Pollak and colleagues (2020) tested a text messaging program that compared quit rates among pregnant women who received supportive messages plus scheduled gradual reduction (SGR) messages (intervention group) to women who received supportive messages only (control group). The SGR messages ("alert texts") were designed to help women gradually reduce or stop smoking over a period of 3-5 weeks, while the support messages were designed to address key determinants derived from social cognitive theory, including self-efficacy, outcome expectations, motivation, problem-solving, partner support, stress, and emotional factors such as guilt and shame. At the end of the study, there were no significant differences in smoking cessation or reduction between the two groups, with a quit rate of about 10% in both groups, and about half of women reporting that they reduced smoking by at least 50%. The study did not include a true control group, but as the authors noted, women in both groups had higher quit rates than would be expected with no intervention at all.

In another published report, Timbor and colleagues (2017) described the development of a smartphone app called "SmokeFree Baby", which was designed to identify and modify five key intervention targets to help pregnant women stop smoking. The app design was grounded in two integrative behavior change theories (COM-B and PRIME theories) in addition to widely-used frameworks for designing complex interventions, including the Medical Research Council, Multiphase Optimization Strategy, and Behavior Change Wheel. The five key intervention targets included identity change, stress management, health information, promoting face-to-face

support, and behavioral substitution. Before launching the first trial of the app, researchers conducted extensive qualitative research to get women's views on the design, content, and usability of the app (Wu et al., 2017). Next, the app was tested in a group of 565 pregnant smokers, but even after usability testing, engagement with SmokeFree Baby was found to be low, and the app did not increase smoking abstinence during pregnancy (Timbor, 2019).

Overall Conclusions

The purpose of the current study was to assess the state of the literature on prenatal smoking cessation interventions and to advance our understanding of the processes involved in prenatal smoking cessation by combining advances in intervention and theory categorization and specification with meta-analytic methods of evidence synthesis. While previous meta-analyses have assessed whether or not prenatal smoking cessation interventions were effective, this review expanded on existing findings by using a recently developed taxonomy to identify, isolate, and quantify the effectiveness of individual behavior change techniques used in interventions, as well as applying a coding scheme to evaluate how theory is being used in the literature and whether the use of theory is associated with the effectiveness of interventions. To our knowledge, this is the first study to demonstrate the feasibility of applying the 26-item coding scheme to published reports of prenatal smoking cessation interventions. This review also provides the first systematic overview of the use of theory in the published literature on prenatal smoking cessation.

While the results of the review were subject to limitations stemming from reporting and measurement practices, several key findings emerged. First, behavior change theory is not being utilized to its full capacity in the development and evaluation of prenatal smoking cessation interventions. This is a significant limitation, but it also presents an opportunity to improve intervention design and possibly increase the effectiveness of interventions. To maximize the utility of behavior change theory, published reports of interventions should include detailed descriptions of how theory was selected and used to inform intervention design. Secondly, many of the most common behavior change techniques used in prenatal smoking cessation interventions were not associated with better intervention outcomes. Thus, significant resources are being expended on behavior change techniques (and delivery formats) that have not been shown to increase effectiveness in many circumstances. In the future, it may be possible to design more parsimonious interventions that save time and money without sacrificing effectiveness. Third, the current review identified contingent rewards as the most effective behavior change technique for promoting smoking cessation during pregnancy across multiple methods of analysis (i.e., subgroup analyses, mediator analyses, and effectiveness ratios).

This review should be considered a first step towards understanding not just *if* an intervention works, but how and why it works.

References

Abel, E. L. (1980). Smoking during pregnancy: a review of effects on growth and development of offspring. *Human Biology*, 593-625.

Abraham, C., Kelly, M.P., West, R., & Michie, S. (2009) The UK national institute for health and clinical excellence public health guidance on behavior change: a brief introduction. *Psychology, Health, & Medicine, 14,* 1-8.

Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology*, *27*(3), 379-387.

Abrahamsson, A., Springett, J., Karlsson, L., & Ottosson, T. (2005). Making sense of the challenge of smoking cessation during pregnancy: a phenomenographic approach. *Health Education Research*, *20*(3), 367-378.

Abroms, L. C., Johnson, P. R., Leavitt, L. E., Cleary, S. D., Bushar, J., Brandon, T. H., & Chiang, S. C. (2017). A Randomized Trial of Text Messaging for Smoking Cessation in Pregnant Women. *American Journal of Preventive Medicine*.

Adams, E.K. & Melvin, C.L. (1998). Costs of Maternal Conditions Attributable to Smoking During Pregnancy. *American Journal of Preventive Medicine, 15*(3), 212-219.

Adams, E.K., Miller, V.P., Ernst, C., Nishimura, B.K., Melvin, C., Merritt, R. (2002). Neonatal health care costs related to smoking during pregnancy. *Health Economics, 11,* 193–206.

Adams, E.K., Melvin, C.L., Raskind-Hood, C., Joski, P.J., Galactionova, E. (2011). Infant delivery costs related to maternal smoking: an update. *Nicotine and Tobacco Research, 13*(8), 627-637.

Adegboye, A. R., Rossner, S., Neovius, M., Lourenço, P. M. C., & Linné, Y. (2010). Relationships between prenatal smoking cessation, gestational weight gain and maternal lifestyle characteristics. *Women and Birth*, *23*(1), 29-35.

Ajzen, I. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211.

Albarracín, D., Durantini, M. R., & Earl, A. (2006). Empirical and theoretical conclusions of an analysis of outcomes of HIV-prevention interventions. *Current Directions in Psychological Science*, *15*(2), 73-78.

Albarracín, D., Gillette, J. C., Earl, A. N., Glasman, L. R., Durantini, M. R., & Ho, M. H. (2005). A test of major assumptions about behavior change: a comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychological Bulletin*, *131*(6), 856-897.

Aligne, C.A, & Stoddard, J.J. (1997). Tobacco and Children: an economic evaluation of the medical effects of parental smoking. *Archives of Pediatrics and Adolescent Medicine, 151,* 648–53.

American College of Obstetricians and Gynecologists (ACOG). (2010). Committee opinion no. 471: smoking cessation during pregnancy. *Obstetrics and Gynecology, 116,* 1241–4.

ACOG (2011, reaffirmed 2013). Committee Opinion No. 503: Tobacco use and women's health. *Obstetrics and Gynecology, 118*(3), 746-750.

Amir, L.H. (2001). Maternal smoking and reduced duration of breastfeeding: a review of possible mechanisms. *Early Human Development*, *64*(1), 45-67.

Andres, R.L., & Day, M.C. (2000). Perinatal complications associated with maternal tobacco use. *Seminars in Neonatology, 5,* 231–241.

Appleton, P. L., & Pharoah, P. O. D. (1998). Partner smoking behavior change is associated with women's smoking reduction and cessation during pregnancy. *British Journal of Health Psychology*, *3*(4), 361-374.

Association of American Medical Colleges. (2007). *Physician behavior and practice patterns related to smoking cessation*. Retrieved January 20, 2015 from: https://www.aamc.org/download/55438/data/.

Association of State and Territorial Health Officials. (2013). *Smoking cessation strategies for women before, during, and after pregnancy: Recommendations for state and territorial health agencies*. Retrieved February 2, 2015 from: http://www.astho.org/prevention/tobacco/smoking-cessation-pregnancy/.

Ayadi, M. F., Adams, E. K., Melvin, C. L., Rivera, C. C., Gaffney, C. A., Pike, J., ... & Ferguson, J. N. (2006). Costs of a smoking cessation counseling intervention for pregnant women: comparison of three settings. *Public Health Reports*, *121*(2), 120-126.

Bacchus, L., Mezey, G., & Bewley, S. (2004). Domestic violence: prevalence in pregnant women and associations with physical and psychological health. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *113*(1), 6-11.

Bandura, A. 1998. Health promotion from the perspective of social cognitive theory. *Psychology and Health*, *13*, 623–649.

Bane, C. M., Ruggiero, L., Dryfoos, J. M., & Rossi, J. S. (1999). Development of a pregnancy-tailored decisional balance measure for smoking cessation. *Addictive Behaviors*, *24*(6), 795-799.

Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Baron, R., Manniën, J., de Jonge, A., Heymans, M. W., Klomp, T., Hutton, E. K., & Brug, J. (2013). Socio-demographic and lifestyle-related characteristics associated with self-reported any, daily and occasional smoking during pregnancy. *PloS one*, *8*(9), e74197.

Barker, D. C., Orleans, C. T., Halpin, H. A., & Barry, M. B. (2004). So near, yet so far: Tobacco dependence treatment for pregnant women. *Nicotine & Tobacco Research*, *6*(Suppl 2), S259-S267.

Bartholomew, L. K., Parcel, G. S., Kok, G., Gottlieb, N. H., & Fernández, M. E. (2011). *Planning health promotion programs: An intervention mapping approach* (3rd ed.). San Francisco, CA: Jossey-Bass.

Bartlett, J.C., Miller, L.S., Rice, D.P., & Max, W.B. (1994). Medical-care expenditures attributable to cigarette smoking--United States, 1993. *Morbidity and Mortality Weekly Report, 43*, 469–472.

Bauman, A.E., Sallis J.F., Dzewaltowski, D.A., et al. (2002). Toward a better understanding of the influences on physical activity: the role of determinants, correlates, causal variables, mediators, moderators, and confounders. *American Journal of Preventive Medicine, 23*, 5-14.

Berman, B.A., & Gritz, E.R. (1991). Women and smoking: current trends and issues for the 1990s. *Journal of Substance Abuse, 13*, 221-238.

Boben, P. J. (2000). Medicaid reform in the 1990s. *Health Care Financing Review*, *22*(2), 1–5.

Borenstein, M., Hedges, L.V., Higgins, J.P., & Rothstein, H.R. (2009). Converting among effect sizes. Hoboken, NJ: *Introduction to Meta-Analysis*. Wiley & Sons. Retrieved August 15, 2016 from: https://www.meta-analysis.com/downloads/Meta-analysis%20Converting%20among%20effect%20sizes.pdf.

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*(2), 431.

Bottorff, J. L., Haines-Saah, R., Oliffe, J. L., & Sarbit, G. (2012). Gender influences in tobacco use and cessation interventions. *Nursing Clinics of North America*, *47*(1), 55-70.

Bottorff, J. L., Kalaw, C., Johnson, J. L., Stewart, M., Greaves, L., & Carey, J. (2006). Couple dynamics during women's tobacco reduction in pregnancy and postpartum. *Nicotine & Tobacco Research*, *8*(4), 499-509.

Bottorff, J.L., Haines-Saah, R., Kelly, M.T., Oliffe, J.L., Torchalla, I., Poole, N., Greaves, L., Robinson, C.A., Ensom, M.H., Chizmuzo, T.O., & Phillips, J.C. (2014). Gender, smoking and tobacco reduction and cessation: A scoping review. *International Journal of Equity in Health, 13*, 114.

Brockwell, S. E., & Gordon, I. R. (2001). A comparison of statistical methods for meta-analysis. *Statistics in Medicine*, *20*(6), 825-840.

Burgess, D. J., Fu, S. S., & van Ryn, M. (2009). Potential unintended consequences of tobacco-control policies on mothers who smoke: a review of the literature. *American Journal of Preventive Medicine*, *37*(2), S151-S158.

Cahill, K., Hartmann-Boyce, J., & Perera, R. (2015). Incentives for smoking cessation. *The Cochrane Library*.

Cane, J., O'Connor, D., & Michie, S. (2012). Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation Science 7*(1), 37.

Carlson, L. E., Goodey, E., Bennett, M. H., Taenzer, P., & Koopmans, J. (2002). The addition of social support to a community-based large-group behavioral smoking cessation intervention: improved cessation rates and gender differences. *Addictive behaviors*, *27*(4), 547-559.

Centers for Disease Control and Prevention (CDC). (1997). Medical Care Expenditures Attributable to Cigarette Smoking During Pregnancy—United States, 1995. *Morbidity and Mortality Weekly Report (MMWR) 46(*44), 1048-1050.

CDC. (2000). *Behavioral Risk Factor Surveillance System: 2000 survey data.* Retrieved February 10, 2015, from https://www.cdc.gov/brfss/annual_data/annual_2000.htm.

CDC. (2002). Annual Smoking-Attributable Mortality, Years of Potential Life Lost, and Economic Costs — United States, 1995–1999. *Morbidity and Mortality Weekly Report, 51*(14), 300-303.

CDC. (2004). State Estimates of Neonatal Health-Care Costs Associated with Maternal Smoking — United States, 1996. *Morbidity and Mortality Weekly Report. 53*(39), 915-917.

CDC. (2008). Smoking prevalence among women of reproductive age — United States, 2006. *Morbidity and Mortality Weekly Report*, *57*, 849-852.

CDC. (2007). National Center for Health Statistics. Births: Final Data for 2005. *National Vital Statistics Reports,* (56)5. Accessed from: https://www.cdc.gov/nchs/data/nvsr/nvsr56/nvsr56_06.pdf.

CDC. (2001). *Women and Smoking: A Report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, CDC, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

CDC. (2013a).“Preventing smoking and exposure to secondhand smoke before, during, and after pregnancy.” Retrieved from: http://www.cdc.gov/nccdphp/publications/factsheets/prevention/pdf/smoking.pdf.

CDC. (2013b). “PRAMS and Smoking.” Retrieved from: http://www.cdc.gov/prams/TobaccoandPrams.htm.

CDC. (2012). “Tobacco Use and Pregnancy.” Retrieved from: http://www.cdc.gov/reproductivehealth/tobaccousepregnancy/.

Chalmers, K., Gupton, A., Katz, A., Hack, T., Hildes-Ripstein, E., Brown, J., … Cann, B. (2004). The description and evaluation of a longitudinal pilot study of a smoking relapse/reduction intervention for perinatal women. *Journal of Advanced Nursing 45*(2), 162-171.

Chamberlain, C., O'Mara-Eves, A., Oliver, S., Caird, J. R., Perlen, S. M., Eades, S. J., & Thomas, J. (2013). Psychosocial interventions for supporting women to stop smoking in pregnancy. *Cochrane Database of Systematic Reviews, 2*: CD001055.

Chapman, S., & Wakefield, M. (2012). Smoking cessation strategies: Time to be more realistic in our expectations of interventions to help quitters. *BMJ, 344*e, 1732.

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, *7*(1), 5-20.

Cohen J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.

Cohen, S., Underwood, L. G., & Gottlieb, B. H. (Eds.). (2000). *Social support measurement and intervention: A guide for health and social scientists*. Oxford University Press.

Coker, A. L., Sanderson, M., & Dong, B. (2004). Partner violence during pregnancy and risk of adverse pregnancy outcomes. *Paediatric and Perinatal Epidemiology*, *18*(4), 260-269.

Coleman-Cowger, V.H. (2012). Smoking cessation intervention for pregnant women: A call for extension into the postnatal period. *Maternal and Child Health Journal, 16*(5), 937-940.

Colman, G. J., & Joyce, T. (2003). Trends in smoking before, during, and after pregnancy in ten states. *American Journal of Preventive Medicine*, *24*(1), 29-35.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation.

Coonrod, D.V., Bruce, N.C., Malcolm, T.D., Drachman, D., Frey, K.A. (2009). Knowledge and attitudes regarding preconception care in a predominantly low-income Mexican American population. *American Journal of Obstetrics and Gynecology,* 200(6), 686, e1-e7.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2013). Developing and evaluating complex interventions: the new Medical Research Council guidance. *International Journal of Nursing Studies*, *50*(5), 587-592.

Crittenden, K. S., Manfredi, C., Cho, Y. I., & Dolecek, T. A. (2007). Smoking cessation processes in low-SES women: the impact of time-varying pregnancy status, health care messages, stress, and health concerns. *Addictive Behaviors*, *32*(7), 1347-1366.

Cummins, S. E., Tedeschi, G. J., Anderson, C. M., & Zhu, S. H. (2016). Telephone intervention for pregnant smokers: a randomized controlled trial. *American Journal of Preventive Medicine*, *51*(3), 318-326.

Curry, S. J., McBride, C., Grothaus, L., Lando, H., & Pirie, P. (2001). Motivation for smoking cessation among pregnant women. *Psychology of Addictive Behaviors*, *15*(2), 126.

Curry, S. J., Grothaus, L., & McBride, C. (1997). Reasons for quitting: intrinsic and extrinsic motivation for smoking cessation in a population-based sample of smokers. *Addictive Behaviors*, *22*(6), 727-739.

Davidson, K. W., Goldstein, M., Kaplan, R. M., Kaufmann, P. G., Knatterud, G. L., Orleans, C. T., ... & Whitlock, E. P. (2003). Evidence-based behavioral medicine: what is it and how do we achieve it?. *Annals of Behavioral Medicine*, *26*(3), 161-171.

Davies, P., Walker, A., & Grimshaw, J. (2003). Theories of behaviour change in studies of guideline implementation. *Proceedings of the British Psychological Society*, *11*(1), 120.

Davies, P., Walker, A. E., & Grimshaw, J. M. (2010). A systematic review of the use of theory in the design of guideline dissemination and implementation strategies and interpretation of the results of rigorous evaluations. *Implementation Science, 5*(14), 5908-5915.

Davis, R. E., Campbell, R., Hildon, Z., Hobbs, L., & Michie, S. (2015). Theories of behaviour and behaviour change across the social and behavioral sciences: A scoping review. *Health Psychology Review, 9,* 323–344.

De Vries, H., Mudde, A. N., Dijkstra, A., & Willemsen, M. C. (1998). Differential beliefs, perceived social influences, and self-efficacy expectations among smokers in various motivational phases. *Preventive Medicine*, *27*(5), 681-689.

Denzin, N.K. (1970). *The research act: A theoretical introduction to sociological methods*. Chicago: Adline Publishing Company.

DerSimonian, R.& Kacker, R. (2007). Random-effects model for meta-analysis of clinical trials: An update. *Contemporary Clinical Trials, 28*, 105–114.

DiClemente, R.J., Crosby, R.A., & Kegler, M.C. (Eds.) (2002). *Emerging theories in health promotion practice and research: Strategies for improving public health.* San Francisco: Jossey-Bass.

DiClemente, C. C., Dolan-Mullen, P., & Windsor, R. A. (2000). The process of pregnancy smoking cessation: implications for interventions. *Tobacco Control*, *9*(suppl 3), iii16-iii21.

DiFranza, J.R. & Lew, R.A. (1995). Effect of Maternal Cigarette Smoking on Pregnancy Complications and Sudden Infant Death Syndrome. *Journal of Family Practice, 40*(4), 385-394.

Dolan-Mullen, P., Ramirez, G., & Groff, J. Y. (1994). A meta-analysis of randomized trials of prenatal smoking cessation interventions. *American Journal of Obstetrics and Gynecology*, *171*(5), 1328-1334.

Donatelle, R. J., Prows, S. L., Champeau, D., & Hudson, D. (2000). Randomised controlled trial using social support and financial incentives for high risk pregnant smokers: Significant Other Supporter (SOS) program. *Tobacco Control*, *9*(suppl 3), iii67-iii69.

Dunn, A. L., Resnicow, K., & Klesges, L. M. (2006). Improving measurement methods for behavior change interventions: opportunities for innovation. *Health Education Research*, *21*(suppl 1), i121-i124.

Durantini, M. R., Albarracin, D., Mitchell, A. L., Earl, A. N., & Gillette, J. C. (2006). Conceptualizing the influence of social agents of behavior change: A meta-analysis of the effectiveness of HIV-prevention interventionists for different groups. *Psychological Bulletin*, *132*(2), 212.

Ebert, L., Fahy, K. (2007). Why do women continue to smoke in pregnancy? *Women and Birth, 20*, 161-168.

Ebrahim, S. H., Floyd, R. L., Merritt II, R. K., Decoufle, P., & Holtzman, D. (2000). Trends in pregnancy-related smoking rates in the United States, 1987-1996. *JAMA*, *283*(3), 361-366.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109-132.

Eccles, M., Grimshaw, J., Walker, A., Johnston, M., & Pitts, N. (2005). Changing the behavior of healthcare professionals: the use of theory in promoting the uptake of research findings. *Journal of Clinical Epidemiology*, *58*(2), 107-112.

Ekblad, M., Gissler, M., Lehtonen, L., & Korkeila, J. (2010). Prenatal smoking exposure and the risk of psychiatric morbidity into young adulthood. *Archives of General Psychiatry, 67*(8), 841-849.

Elsenbruch, S., Benson, S., Rücke, M., Rose, M., Dudenhausen, J., Pincus-Knackstedt, M. K., ... & Arck, P. C. (2007). Social support during pregnancy: effects on maternal depressive symptoms, smoking and pregnancy outcome. *Human Reproduction*, *22*(3), 869-877.

Ershoff, D. H., Quinn, V. P., Boyd, N. R., Stern, J., Gregory, M., & Wirtschafter, D. (1999). The Kaiser Permanente prenatal smoking-cessation trial: When more isn't better, what is enough?. *American Journal of Preventive Medicine*, *17*(3), 161-168.

Everett, K. D., Gage, J., Bullock, L., Longo, D. R., Geden, E., & Madsen, R. W. (2005). A pilot study of smoking and associated behaviors of low-income expectant fathers. *Nicotine & tobacco research*, *7*(2), 269-276.

Fang, W. L., Goldstein, A. O., Butzen, A. Y., Hartsock, S. A., Hartmann, K. E., Helton, M., & Lohr, J. A. (2004). Smoking cessation in pregnancy: a review of postpartum relapse prevention strategies. *The Journal of the American Board of Family Practice*, *17*(4), 264-275.

Fiore, M.C., Jaen, C.R., Baker, T.B. (2000). *Clinical Practice Guideline: Treating Tobacco Use and Dependence*. Rockville, Md, USA: U.S. Department of Health and Human Services, Tobacco Use and Dependence Guideline Panel. http://www.ahrq.gov/path/tobacco.htm#Clinic.

Fisher, J.D., & Fisher, W.A. (1992). Changing AIDS-risk behavior. *Psych Bulletin, 111,* 455–474.

Fisher, J.D., Fisher, W.A., & Shuper, P.A. (2009). The information-motivation-behavioral skills model of HIV prevention behavior. In: R.J. DiClemente, R.A. Crosby, & M.C. Kegler (Eds.), Emerging theories in health promotion practice and research (pp. 21-63). Jossey-Bass, San Francisco, CA.

Flemming, K., Graham, H., Heirs, M., Fox, D., & Sowden, A. (2013). Smoking in pregnancy: a systematic review of qualitative research of women who commence pregnancy as smokers. *Journal of Advanced Nursing*, *69*(5), 1023-1036.

Francis, J. J., Stockton, C., Eccles, M. P., Johnston, M., Cuthbertson, B. H., Grimshaw, J. M., ... & Stanworth, S. J. (2009). Evidence-based selection of theories for designing behaviour change interventions: Using methods based on theoretical construct domains

to understand clinicians' blood transfusion behaviour. *British Journal of Health Psychology*, *14*(4), 625-646.

French, S. D., Green, S. E., O'Connor, D. A., McKenzie, J. E., Francis, J. J., Michie, S., ... & Grimshaw, J. M. (2012). Developing theory-informed behaviour change interventions to implement evidence into practice: a systematic approach using the Theoretical Domains Framework. *Implementation Science*, *7*(1), 38.

Gage, J.D., Everett, K.D., & Bullock, L. (2007). A review of research literature addressing male partners and smoking during pregnancy. *Journal of Obstetrics, Gynecology, & Neonatal Nursing, 36*(6), 574-580.

Gardner, B., Whittington, C., McAteer, J., Eccles, M. P., & Michie, S. (2010). Using theory to synthesise evidence from behaviour change interventions: the example of audit and feedback. *Social Science & Medicine*, *70*(10), 1618-1625.

Gilliland, F. D., Li, Y. F., & Peters, J. M. (2001). Effects of maternal smoking during pregnancy and environmental tobacco smoke on asthma and wheezing in children. *American Journal of Respiratory and Critical Care Medicine*, *163*(2), 429-436.

Glanz, K., & Bishop, D. B. (2010). The role of behavioral science theory in development and implementation of public health interventions. *Annual Review of Public Health, 31*, 399-418.

Glanz, K., Rimer, B. K., & Viswanath, K. (2008). *Health behavior and health education: Theory, research, and practice*. New York, NY: Wiley & Sons.

Goodwin, R. D., Keyes, K., & Simuro, N. (2007). Mental disorders and nicotine dependence among pregnant women in the United States. *Obstetrics & Gynecology*, *109*(4), 875-883.

Greaves, L., Kalaw, C., Bottorff, J.L. (2007). Case studies of power and control related to tobacco use during pregnancy. *Women's Health Issues, 17*, 325-332.

Greaves L, Tungohan E. (2007). Engendering tobacco control: Using an international public health treaty to reduce smoking and empower women. *Tobacco Control, 16*, 148-150.

Greene, J., Sacks, R.M., McMenamin, S.B. (2014). The impact of tobacco dependence treatment coverage and copayments in Medicaid. *American Journal of Preventive Medicine, 46*(4), 331-336.

Grimshaw, J. M., Thomas, R. E., MacLennan, G., Fraser, C., Ramsay, C. R., Vale, L., ... & Wensing, M. (2005). Effectiveness and efficiency of guideline dissemination and implementation strategies. *International Journal of Technology Assessment in Health Care*, *21*(01), 149-149.

Hardeman, W., Sutton, S., Griffin, S., Johnston, M., White, A., Wareham, N. J., & Kinmonth, A. L. (2005). A causal modelling approach to the development of theory-based behaviour change programmes for trial evaluation. *Health Education Research*, *20*(6), 676-687.

Hardeman, W., Johnston, M., Johnston, D.W., Bonetti,, D., Wareham, N., & Kinmonth, A.L. (2002). Application of the theory of planned behaviour in behaviour change interventions: a systematic review.. *Psychology and Health*, 17, 123–158.

Haslam, C., & Draper, E. (2000). Stage of change is associated with assessment of the health risks of maternal smoking among pregnant women. *Social Science & Medicine*, *51*(8), 1189-1196.

Haslam, C., Draper, E. (2001). A qualitative study of smoking during pregnancy. (2001). *Psychology Health & Medicine*, 6(1), 95-99.

Haslam C, Draper E, Goyder E. The pregnant smoker: a preliminary investigation of the social and psychological influences. Journal of Public Health Medicine 1997;19(2):187-192.

Hawkins, S.S., & Baum, C.F. (2014). Impact of state cigarette taxes on disparities in maternal smoking during pregnancy. *American Journal of Public Health, 104*(8), 1464-1470.

Hawkins, S.S., Baum, C.F., Oken, E., & Gillman, M.W. (2014). Associations of tobacco control policies with birth outcomes. *JAMA Pediatrics, 168*(11), e142365.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486.

Heimlich, J. E., & Ardoin, N. M. (2008). Understanding behavior to understand behavior change: A literature review. *Environmental Education Research*, *14*(3), 215-237.

Henningfield, J.E., Fant, R.V., Buchhalter, A.R., Stitzer, M.L. (2005). Pharmacotherapy for nicotine dependence. *CA: A Cancer Journal for Clinicians*, *55*(5), 281–299.

Higgins, S. T., & Petry, N. M. (1999). Contingency management: Incentives for sobriety. *Alcohol Research and Health, 23*, 122–127.

Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta‐analysis. *Statistics in Medicine*, *21*(11), 1539-1558.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557–560.

Higgins, S.T., Washio, Y., Heil, S.H. (2012). Financial incentives for smoking cessation among pregnant and newly postpartum women. *Preventive Medicine, 55*(suppl), S33-S40.

Hoffman, S., & Hatch, M. C. (1996). Stress, social support and pregnancy outcome: a reassessment based on recent research. *Paediatric and perinatal epidemiology*, *10*(4), 380-405.

Holtrop, J.S., Meghea, C., Raffo, J.E., Biery, L., Chartkoff, S.B., & Roman, L. (2010). Smoking among pregnant women with Medicaid insurance: Are mental health factors related? *Maternal and Child Health Journal, 14*(6), 971-977.

Homish, G.G., Eiden, R.D., Leonard, K.E., Kozlowski, L.T. (2012). Social-environmental factors related to prenatal smoking. *Addictive Behaviors, 37*(1), 73-77.

Hotham, E., Atkinson, E., Gilbert, A. (2002). Focus groups with pregnant smokers: barriers to cessation, attitudes to nicotine patch use and perceptions of cessation counselling by care providers. *Drug and Alcohol Review, 21*, 163-168.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index?. *Psychological Methods*, *11*(2), 193-206.

Hughes, E.G., Lamont, D.., Beecroft, M.L., Wilson, D.M., Brennan, B.G., & Rice, S.C. (2000). Randomized trial of a “stage of change” oriented smoking cessation intervention in infertile and pregnant women. *Fertility and Sterility, 74*(3), 498-503.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta‐analysis models: implications for cumulative research knowledge. *International Journal of Selection and Assessment*, *8*(4), 275-292.

Ibrahim, J. K., Schauffler, H. H., Barker, D. C., & Orleans, C. T. (2002). Coverage of tobacco dependence treatments for pregnant women and for children and their parents. *American Journal of Public Health*, 92, 1940–1942.

Improved Clinical Effectiveness through Behavioural Research Group. (2006). Designing theoretically-informed implementation interventions. *Implementation Science*, *1*(1), 4.

Jamtvedt, G., Young, J. M., Kristoffersen, D. T., O’Brien, M. A., & Oxman, A. D. (2006). Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews*, *2*(2).

Jeffery, R.W. (2004). How can health behavior theory be made more useful for intervention research? *International Journal of Behavioral Nutrition and Physical Activity, 1*, 10.

Johnson, B. T., Scott-Sheldon, L. A. J., Huedo-Medina, T. B., & Carey, M. P. (2011). Interventions to reduce sexual risk for HIV in adolescents: A meta-analysis of trials, 1985–2008. *Archives of Pediatrics and Adolescent Medicine, 165*, 77-84.

Kahn, R. S., Certain, L., & Whitaker, R. C. (2002). A reexamination of smoking before, during, and after pregnancy. *American Journal of Public Health*, *92*(11), 1801-1808.

Kane, R.L., Johnson, P.E., Town, R.J., & Butler, M. (2004). *Economic Incentives for Preventive Care: Summary*. Rockville, MD: Agency for Healthcare Research and Quality.

Kelley K., Bond, R., Abraham, C. Effective approaches to persuading pregnant women to quit smoking: a meta-analysis of intervention evaluation studies. *British Journal of Health Psychology, 6,* 207-209.

Kelley, G. A., & Kelley, K. S. (2012). Statistical models for meta-analysis: A brief tutorial. *World Journal of Methodology*, *2*(4), 27-32.

Kleinman, J. C., & Kopstein, A. (1987). Smoking during pregnancy, 1967-80. *American Journal of Public Health*, *77*(7), 823-825.

Kleinman, J. C., Pierre, M. B., Madans, J. H., Land, G. H., & Shramm, W. F. (1988). The effects of maternal smoking on fetal and infant mortality. *American Journal of Epidemiology*, *127*(2), 274-282.

Kogan, M.D., Kotelchuck, M., Alexander, G.R., & Johnson, W.E. (1994). Racial disparities in reported prenatal care advice from health care providers. *American Journal of Public Health, 84*, 82-88.

Kok, G. (2014). A practical guide to effective behavior change: How to apply theory – and evidence-based behavior change methods in an intervention. *European Health Psychologist, 16*(5), 156–170.

Kok, G., Gottlieb, N., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., ... Bartholomew, L. K. (2016). A taxonomy of behavior change methods: An intervention mapping approach. *Health Psychology Review.*

Koshy, P., Mackenzie, M., Tappin, D., & Bauld, L. (2010). Smoking cessation during pregnancy: the influence of partners, family, and friends on quitters and non-quitters. *Health Soc Care Community, 18*(5), 500-510.

Kyrklund-Blomberg, N.B., Gennser, G., & Cnattingius, S. (2001). Placental abruption and perinatal death. *Paediatric and Perinatal Epidemiology, 15,* 290–297.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.

Larsen, K. R., & Bong, C. H. (2016). A Tool for Addressing Construct Identity in Literature Reviews and Meta-Analyses. *MIS Quarterly*, *40*(3), 529-551.

Lightwood JM, Phibbs CS, Glantz SA. (1999). Short-term health and economic benefits of smoking cessation: low birth weight. *Pediatrics*, *104,* 1312-1320.

Likis, F.E., Andrews, J.C., Fonnesbeck, C.J., Hartmann, K.E., Jerome, R.N., Potter, S.A., Surawixa, T.S., McPheeters, M.L. (2014). *Smoking cessation interventions in pregnancy and postpartum care*. Evidence Report/Technology Assessment No. 214. AHRQ Publication No. 14-E001-EF. Accessed February 10, 2015 at: http://www.effectivehealthcare.ahrq.gov/ehc/products/517/1871/smoking-pregnancy-infants-report-140226.pdf.

Lippke, S., & Ziegelmann, J. P. (2008). Theory-Based Health Behavior Change: Developing, Testing, and Applying Theories for Evidence-Based Interventions. *Applied Psychology*, *57*(4), 698-716.

Lipsey, M.W. & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: Sage publications.

Liu, J., Rosenberg, K. D., & Sandoval, A. P. (2006). Breastfeeding duration and perinatal cigarette smoking in a population-based cohort. *American Journal of Public Health*, *96*(2), 309-314.

Lu, Y., Tong, S., & Oldenburg, B. (2001). Determinants of smoking and cessation during and after pregnancy. *Health Promotion International*, *16*(4), 355-365.

Lumley, J., Chamberlain, C., Dowswell, T., Oliver, S., Oakley, L., Watson, L. (2009). Interventions for promoting smoking cessation during pregnancy. *Cochrane Database of Systematic Reviews, 3*, CD001055.

Lundquist, R. S., Seward, G., Byatt, N., Tonelli, M. E., & Kolodziej, M. E. (2012). Using a multidisciplinary approach for pregnant women with nicotine dependence and co-occurring disorders. *Journal of Dual Diagnosis*, *8*(2), 158-167.

Maritz, G. S., Morley, C. J., & Harding, R. (2005). Early developmental origins of impaired lung structure and function. *Early Human Development*, *81*(9), 763-771.

Markus, A.R., Andres, E., West, K.D., Garro, N., & Pellegrini, C. (2013). Medicaid Covered Births, 2008 Through 2010, in the Context of the Implementation of Health Reform. *Women's Health Issues*, *23* (5), e273-e280.

Martin, J. A., Hamilton, B. E., Ventura, S. J., Menacker, F., Park, M. M., & Sutton, P. D. (2002). Births: Final data for 2001. *National Vital Statistics Report, 51*(2), 1–103.

Martin J.A., Hamilton, B.E., Sutton, P.D. (2007). Births: final data for 2005. *National Vital Statistics Report, 56*(6), 1e103.

Maxson, P. J., Edwards, S. E., Ingram, A., & Miranda, M. L. (2012). Psychosocial differences between smokers and non-smokers during pregnancy. *Addictive behaviors*, *37*(2), 153-159.

May, S., & West, R. (2000). Do social support interventions ("buddy systems") aid smoking cessation? A review. *Tobacco Control*, *9*(4), 415-422.

McBride, C. M., Curry, S. J., Grothaus, L. C., Nelson, J. C., Lando, H., & Pirie, P. L. (1998). Partner smoking status and pregnant smoker's perceptions of support for and likelihood of smoking cessation. *Health Psychology*, *17*(1), 63.

McBride, C. M., Curry, S. J., Lando, H. A., Pirie, P. L., Grothaus, L. C., & Nelson, J. C. (1999). Prevention of relapse in women who quit smoking during pregnancy. *American Journal of Public Health, 89*(5), 706-711.

McDonald, S.D., Perkins, S.L., Jodouin, C.A., & Walker, M.C. (2002). Folate levels in pregnant women who smoke: an important gene/environment interaction. *American Journal of Obstetrics and Gynecology, 187,* 620–625.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276-282.

McLeroy, K.R., Bibeau, D. Steckler, A. & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly, 15,* 351–377.

McMenamin, S.B., Halpin, H.A., & Ganiats, T.G. (2012). Medicaid coverage of tobacco-dependence treatment for pregnant women: Impact of the Affordable Care Act. *American Journal of Preventive Medicine, 43*(4), e27-e29.

Melvin, C. L., Dolan-Mullen, P., Windsor, R. A., Whiteside, H. P., & Goldenberg, R. L. (2000). Recommended cessation counselling for pregnant women who smoke: a review of the evidence. *Tobacco Control*, *9*(suppl 3), iii80-iii84.

Melvin, C. L., & Gaffney, C. A. (2004). Treating nicotine use and dependence of pregnant and parenting smokers: an update. *Nicotine & Tobacco Research*, *6*(suppl. 2), s107-S124.

Michie, S., & Abraham, C. (2004). Interventions to change health behaviours: evidence-based or evidence-inspired?. *Psychology & Health*, *19*(1), 29-49.

Michie, S., Abraham, C., Eccles, M. P., Francis, J. J., Hardeman, W., & Johnston, M. (2011). Strengthening evaluation and implementation by specifying components of behavior change interventions: A study protocol. *Implementation Science, 6*, 10.

Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D., & Walker, A. (2005). Making psychological theory useful for implementing evidence based practice: A consensus approach. *Quality and Safety in Health Care*, *14*(1), 26-33.

Michie, S. (2008). Designing and implementing behaviour change interventions to improve population health. *Journal of Health Services Research & Policy*, *13*(suppl 3), 64-69.

Michie, S., Johnston, M., Francis, J., Hardeman, W., & Eccles, M. (2008). From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. *Applied Psychology*, *57*(4), 660-680.

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009a). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychology*, *28*(6), 690.

Michie, S., Fixsen, D., Grimshaw, J. M., & Eccles, M. P. (2009b). Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implement Sci*, *4*(40), 1-6.

Michie, S., & Prestwich, A. (2010). Are interventions theory-based? Development of a theory coding scheme. *Health Psychology*, *29*(1), 1.

Michie, S., Abraham, C., Eccles, M. P., Francis, J. J., Hardeman, W., & Johnston, M. (2011a). Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. *Implementation Science*, *6*(10), 10.

Michie, S., Churchill, S., & West, R. (2011b). Identifying evidence-based competences required to deliver behavioural support for smoking cessation. *Annals of Behavioral Medicine*, *41*(1), 59-70.

Michie, S., Hyder, N., Walia, A., & West, R. (2011). Development of a taxonomy of behavior change techniques used in individual behavioral support for smoking cessation. *Addictive Behaviors, 36*, 315-319.

Michie, S., van Stralen, M. M., & West, R. (2011c). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1), 42.

Michie, S., & Johnston, M. (2012). Theories and techniques of behaviour change: Developing a cumulative science of behaviour change. *Health Psychology Review*, *6*(1), 1-6.

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., ... & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, *46*(1), 81-95.

Miller, D.P., Villa, K.F., Hogue, S.L., & Sivapathasundaram, D. (2001). Birth and first-year costs for mothers and infants attributable to maternal smoking. *Nicotine and Tobacco Research, 3*, 25–35.

Moiduddin, E., & Massey, D. S. (2008). How Neighborhood Disadvantage Reduces Birth Weight. *International Journal of Conflict and Violence*, *2*(1), 113-129.

Moore, L., Campbell, R., Whelan, A., Mills, N., Lupton, P., Misselbrook, E., & Frohlich, J. (2002). Self help smoking cessation in pregnancy: cluster randomised controlled trial. *BMJ*, *325*(7377), 1383.

Nanovskaya, T. N., Oncken, C., Fokina, V. M., Feinn, R. S., Clark, S. M., West, H., ... & Hankins, G. D. (2017). Bupropion sustained release for pregnant smokers: a randomized, placebo-controlled trial. *American Journal of Obstetrics and Gynecology*, *216*(4), 420-e1.

Noar, S. M., Chabot, M., & Zimmerman, R. S. (2008). Applying health behavior theory to multiple behavior change: considerations and approaches. *Preventive Medicine*, *46*(3), 275-280.

Ockene, J. K., Ma, Y., Zapka, J. G., Pbert, L. A., Goins, K. V., & Stoddard, A. M. (2002). Spontaneous cessation of smoking and alcohol use among low-income pregnant women. *American Journal of Preventive Medicine*, *23*(3), 150-159.

Oken, E., Levitan, E. B., & Gillman, M. W. (2008). Maternal smoking during pregnancy and child overweight: systematic review and meta-analysis. *International Journal of Obesity*, *32*(2), 201-210.

Ogden, J. (2016). Celebrating variability and a call to limit systematisation: The example of the behavior change technique taxonomy and the behavior change wheel. *Health Psychology Review, 10*(3), 245–250.

Orleans, C. T., Barker, D. C., Kaufman, N. J., & Marx, J. F. (2000). Helping pregnant smokers quit: meeting the challenge in the next decade. *Tobacco Control*, *9*(suppl 3), iii6-iii11.

Park, E. W., Tudiver, F., Schultz, J. K., & Campbell, T. (2004). Does enhancing partner support and interaction improve smoking cessation? A meta-analysis. *The Annals of Family Medicine*, *2*(2), 170-174.

Park, E.W., Tudiver, F.G., & Campbell, T. (2012). Enhancing partner support to improve smoking cessation. *Cochrane Database of Systematic Reviews, 11*(7), CD002928.

Paterson, T. A., Harms, P. D., Steel, P., & Credé, M. (2016). An assessment of the magnitude of effect sizes: Evidence from 30 years of meta-analysis in management. *Journal of Leadership & Organizational Studies*, *23*(1), 66-81.

Patterson, F., Seravalli, L., Hanlon, A., & Nelson, D. B. (2012). Neighborhood safety as a correlate of tobacco use in a sample of urban, pregnant women. *Addictive Behaviors*, *37*(10), 1132-1137.

Perski, O., Blandford, A., West, R., & Michie, S. (2016). Conceptualizing engagement with digital behavior change interventions: a systematic review using principles from critical interpretive synthesis. *Translational Behavioral Medicine*, 7(2), 254–267.

Perski, O., Michie, S., Blandford, A., & West, R. (2016). Engagement with digital behavior change interventions: a structured review using techniques from critical interpretive synthesis. *European Health Psychologist*, *18*(S), 863.

Peters, G.-J. Y. (2014). A practical guide to effective behavior change: How to identify what to change in the first place. *European Health Psychologist, 16*(4), 142–155.

Peters, G.-J. Y., de Bruin, M., & Crutzen, R. (2015). Everything should be as simple as possible, but no simpler: Towards a protocol for accumulating evidence regarding the active content of health behaviour change interventions. *Health Psychology Review, 9*(1), 1–14.

Peters, G.-J. Y., Ruiter, R. A. C., & Kok, G. (2014). Threatening communication: A qualitative study of fear appeal effectiveness beliefs among intervention developers, policymakers, politicians, scientists and advertising professionals. *International Journal of Psychology, 49*(2), 71–79.

Petersen, R., Garrett, J.M., Melvin, C.L., & Hartmann, K.E. (2006). Medicaid reimbursement for prenatal smoking intervention influences quitting and cessation. *Tobacco Control, 15*(1), 30-34.

Pollak, K. I., Baucom, D. H., Peterson, B. L., Stanton, S., & McBride, C. M. (2006). Rated helpfulness and partner-reported smoking cessation support across the pregnancy-postpartum continuum. *Health Psychology*, *25*(6), 762.

Pollak, K. I., & Mullen, P. D. (1997). An exploration of the effects of partner smoking, type of social support, and stress on postpartum smoking in married women who stopped smoking during pregnancy. *Psychology of Addictive Behaviors*, *11*(3), 182.

Pollak, K. I., McBride, C. M., Baucom, D. H., Curry, S. J., Lando, H., Pirie, P. L., & Grothaus, L. C. (2001). Women's perceived and partners' reported support for smoking cessation during pregnancy. *Annals of Behavioral Medicine*, *23*(3), 208-214.

Prochaska, J.O. & DiClemente, C.C. (1983). Stages and processes of self-change in smoking: towards an integrative model of change. *Journal of Consulting and Clinical Psychology*, 51, 390–395.

Redding, C. A., Maddock, J. E., & Rossi, J. S. (2006). The sequential approach to measurement of health behavior constructs: Issues in selecting and developing measures. *Californian Journal of Health Promotion*, *4*(1), 83-101.

Riley, B. L., MacDonald, J., Mansi, O., Kothari, A., Kurtz, D., von Tettenborn, L. I., & Edwards, N. C. (2008). Is reporting on interventions a weak link in understanding how and why they work? A preliminary exploration using community heart health exemplars. *Implementation Science*, *3*(27), 1-12.

Robbins, C.L., Zapata L.B. Farr, S.L., Kroelinger, C.D., Morrow, B., … Barfield, W.D. (2014). Core State Preconception Health Indicators — Pregnancy Risk Assessment Monitoring System and Behavioral Risk Factor Surveillance System, 2009. *Morbidity and Mortality Weekly Report, 63*(ss03), 1-62.

Rosenthal R. (1987). *Judgment Studies: Design, Analysis, and Meta-analysis.* Cambridge University Press: Paris, France.

Ruggiero, L., Tsoh, J.Y., Everett, K., Fava, J.L., & Guise, B.J. (2000). The transtheoretical model of smoking: Comparison of pregnant and non-pregnant smokers. *Addictive Behaviors, 25*(2), 239-251.

Salihu, H. M., Aliyu, M. H., Pierre-Louis, B. J., & Alexander, G. R. (2003). Levels of excess infant deaths attributable to maternal smoking during pregnancy in the United States. *Maternal and Child Health Journal*, *7*(4), 219-227.

Schauffler, H. H., Mordavsky, J., Barker, D. C., & Orleans, C. T. (2001). State Medicaid coverage for tobacco-dependence treatments—United States, 1998 and 2000. *Mortality and Morbidity Weekly Report, 50*, 979–982.

Schneider, S., Huy, C., Schuetz, J., & Diehl, K. (2010). Smoking cessation during pregnancy: a systematic literature review. *Drug and alcohol review*, *29*(1), 81-90.

Scheibmeir, M. S., O'Connell, K. A., Aaronson, L. S., & Gajewski, B. (2005). Smoking cessation strategy use among pregnant ex-smokers. *Western journal of nursing research*, *27*(4), 411-427.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, *47*(10), 1173.

Scott, J. A., Binns, C. W., Oddy, W. H., & Graham, K. I. (2006). Predictors of breastfeeding duration: evidence from a cohort study. *Pediatrics*, *117*(4), e646-e655.

Severson, H. H., Andrews, J. A., Lichtenstein, E., Wall, M., & Zoref, L. (1995). Predictors of smoking during and after pregnancy: a survey of mothers of newborns. *Preventive Medicine*, *24*(1), 23-28.

Shah, N.R. & Bracken, M.B. (2000). A systematic review and meta-analysis of prospective studies on the association between maternal cigarette smoking and preterm delivery. *American Journal of Obstetrics and Gynecology, 182*, 465-472.

Siu A.L. (2015). Behavioral and pharmacotherapy interventions for tobacco smoking cessation in adults, including pregnant women: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine*, *163*, 622-634.

Slade, P., Laxton-Kane, M., & Spiby, H. (2006). Smoking in pregnancy: The role of the transtheoretical model and the mother's attachment to the fetus. *Addictive behaviors*, *31*(5), 743-757.

Smoak, N. D., Scott-Sheldon, L. A. J., Johnson, B. T., & Carey, M. P. (2006). Sexual risk reduction interventions do not inadvertently increase the overall frequency of sexual behavior: A meta- analysis of 174 studies with 116,735 participants. *Journal of Acquired Immune Deficiency Syndromes, 41*, 374–384.

Solomon, L. J., & Quinn, V. P. (2004). Spontaneous quitting: self-initiated smoking cessation in early pregnancy. *Nicotine & Tobacco Research*, *6*(Suppl 2), S203-S216.

Srivastava, J., & Shu, L.H. (2014). An ontology for unifying behavior-change literature. *CIRP Annals: Manufacturing Technology, 63*(1), 173-176.

Staunton, L., Gellert, P., Knittle, K., & Sniehotta, F. F. (2014). Perceived control and intrinsic vs. extrinsic motivation for oral self-care: A full factorial experimental test of theory-based persuasive messages. *Annals of Behavioral Medicine, 49*, 258–268.

Stotts, A. L., Diclemente, C. C., Carbonari, J. P., & Mullen, P. D. (1996). Pregnancy smoking cessation: a case of mistaken identity. *Addictive Behaviors*, *21*(4), 459-471.

Sutton, S. (2000). A critical review of the transtheoretical model applied to smoking cessation In: Norman, P. Abraham, C. and Conner, M. (Eds.) *Understanding and Changing Health Behaviour: from Health Beliefs to Self-Regulation* pp. 207–225. Harwood Academic. Amsterdam.

Taylor, M. (2009). York Health Economics Consortium for National Institute for Health and Clinical Excellence. 2009. Economic analysis of interventions for smoking cessation aimed at pregnant women: supplementary report. Available from: https://www.nice.org.uk/guidance/ph26/evidence/economic-analysis-review-376281901.

Thorndike, A. N., Rigotti, N. A., Stafford, R. S., & Singer, D. E. (1998). National patterns in the treatment of smokers by physicians. *The Journal of American Medical Association, 279*, 604–608.

Tod, A. (2003). Barriers to smoking cessation in pregnancy: a qualitative study. *British Journal of Community Nursing, 8*(2), 56-62.

Tong, V.T., Jones, J.R., Dietz, P.M., et al. (2009). Trends in smoking before, during, and after pregnancy – Pregnancy Risk Assessment Monitoring System (PRAMS), United States, 31 sites, 2000e2005. *MMWR Surveillance Summary, 58*(4), 1e29.

Tong, V.T., Dietz, P.M., Morrow, B., D'Angelo, D.V., Farr, S.L., Rockhill, K.M., & England, L.J. (2013). Trends in smoking before, during, and after pregnancy: Pregnancy Risk Assessment Monitoring System, United States, 40 sites, 2000-2010. *MMWR Surveillance Summary, 62*(6), 1-19.

Tong, V. T., England, L. J., Dietz, P. M., & Asare, L. A. (2008). Smoking patterns and use of cessation interventions during pregnancy. *American Journal of Preventive Medicine*, *35*(4), 327-333.

Tran, S. T. T., Rosenberg, K. D., & Carlson, N. E. (2010). Racial/ethnic disparities in the receipt of smoking cessation interventions during prenatal care. *Maternal and Child Health Journal*, *14*(6), 901-909.

Upton, M.N., Watt, O.C., Davey Smith, G., McConnachie, A., & Hart, C.L. (1998). Permanent effects of maternal smoking on off-springs' lung function. *Lancet, 352*(9126), 453.

Ussher, M., Etter, J. F., & West, R. (2006). Perceived barriers to and benefits of attending a stop smoking course during pregnancy. *Patient Education and Counseling, 61*(3), 467-472.

U.S. Department of Health and Human Services. (2001). *Smoking and women's health. A report of the Surgeon General.* U.S. Department of Health and Human Services. Atlanta, Georgia.

U.S. Department of Health and Human Services. (2004). *The health consequences of smoking: A report of the Surgeon General*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. Accessed from: http://www.cdc.gov/tobacco/data_statistics/sgr/2004/complete_report/index.htm

U.S. Department of Health and Human Services. (2010a) *How Tobacco Smoke Causes Disease: What It Means to You: A Report of the Surgeon General.* U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. Accessed from: https://www.cdc.gov/tobacco/data_statistics/sgr/2010/consumer_booklet/index.htm.

U.S. Department of Health and Human Services. (2010b). *Overview of Findings Regarding Reproductive Health: A Report of the Surgeon General.* U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. Accessed from: https://www.cdc.gov/tobacco/data_statistics/sgr/2010/highlight_sheets/pdfs/overview_reproductive.pdf.

U.S. Department of Health and Human Services. (2006). *The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General.* U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. Accessed from: http://www.cdc.gov/tobacco/data_statistics/sgr/2006/consumer_summary/index.htm.

U.S. Department of Health and Human Services (n.d). Office of Disease Prevention and Health Promotion. *Healthy People 2020: TU-6, Increase smoking cessation during pregnancy*. Data description. Washington, DC. Accessed from: https://www.healthypeople.gov/node/5364/data_details

VanLandeghem, K., & Schor, E.L. (2012). *New opportunities for integrating and improving health care for women, children, and their families.* The Commonwealth Fund and the Association of Maternal and Child Health Programs. Accessed from: http://www.amchp.org/Policy-Advocacy/health-reform/resources/Documents/New%20Opportunities%20for%20Integrating%20and%20Improving%20Health%20Care%20for%20Women,%20Children,%20and%20Their%20Families.pdf.

Vries, H. D., & Mudde, A. N. (1998). Predicting stage transitions for smoking cessation applying the attitude-social influence-efficacy model. *Psychology and Health*, *13*(2), 369-385.

Wakschlag, L. S., Pickett, K. E., Middlecamp, M. K., Walton, L. L., Tenzer, P., & Leventhal, B. L. (2003). Pregnant smokers who quit, pregnant smokers who don't: Does history of problem behavior make a difference?. *Social science & medicine*, *56*(12), 2449-2460.

Waldron I. (1991). Patterns and causes of gender differences in smoking. *Social Science and Medicine, 32,* 989-1005

Webb, T. L., Sniehotta, F. F., & Michie, S. (2010). Using theories of behaviour change to inform interventions for addictive behaviours. *Addiction*, *105*(11), 1879-1892.

Weinstein, N.D., Rothman, A.J., & Sutton, S.R. (1998). Stage theories of health behavior: conceptual and methodological issues. *Health Psychology*, *17,* 290-299.

West, R., & Michie, S. (2016). A guide to development and evaluation of digital behaviour change interventions in healthcare. Journal of Medical Internet Research. *Silverback Publishing.*

Westmaas, J. L., Bontemps-Jones, J., & Bauer, J. E. (2010). Social support in smoking cessation: reconciling theory and evidence. *Nicotine & Tobacco Research*, 695-707.

Wilson, M., Allen, D.D., & Li, J.C. (2006a). Improving measurement in health education and health behavior research using item response modeling: comparison with the classical test theory approach. *Health Education, 21*(Suppl 1), i19-i32.

Wilson, M., Allen, D.D., & Li, J.C. (2006b). Improving measurement in health education and health behavior research using item response modeling: introducing item response modeling. *Health Education, 21*(Suppl 1), i4-i18.

Windsor, R. A., Boyd, N. R., & Orleans, C. T. (1998). A meta-evaluation of smoking cessation intervention research among pregnant women: improving the science and art. *Health Education Research*, *13*(3), 419-438.

Windsor, R. A., Lowe, J. B., Perkins, L. L., Smith-Yoder, D., Artz, L., Crawford, M., ... & Boyd Jr, N. R. (1993). Health education for pregnant smokers: its behavioral impact and cost benefit. *American Journal of Public Health*, *83*(2), 201-206.

Windsor, R.A., Li, C.Q., Boyd, N.R., & Hartmann, K.E. (1999). The use of significant reduction rates to evaluate health education methods for pregnant smokers: a new harm-reduction behavioral indicator? *Health Education and Behavior, 26*(5), 648-662..

World Health Organization. (2008). *WHO Report on The Global Tobacco Epidemic, The MPOWER package.* Geneva: World Health Organization.

Woodby, L. L., Windsor, R. A., Snyder, S. W., Kohler, C. L., & Diclemente, C. C. (1999). Predictors of smoking cessation during pregnancy. *Addiction*, *94*(2), 283-292.

Wong, W., Liu, W., & Bennamoun, M. (2012). Ontology learning from text: A look back and into the future. *ACM Computing Surveys (CSUR), 44(4)*, 1-36.

Young, S., Sherrill, D.L., Arnott, J., Diepeveen, D., LeSouëf, P.N., & Landau, L.I. (2000). Parental factors affecting respiratory function during the first year of life. *Pediatric Pulmonology, 29,* 331-340.

Yu, S.M., Park, C. H., & Schwalberg, R. H. (2002). Factors associated with smoking cessation among US pregnant women. *Maternal and Child Health Journal*, *6*(2), 89-97.

Yunzai-Butler, C., Joyce, T., & Racine, A.D. (2010). Maternal smoking and the timing of WIC enrollment. *Maternal and Child Health Journal, 14*(3), 318-331.

Zapka, J. G., Pbert, L., Stoddard, A. M., Ockene, J. K., Goins, K. V., & Bonollo, D. (2000). Smoking cessation counseling with pregnant and postpartum women: a survey of community health center providers. *American Journal of Public Health*, *90*(1), 78.

APPENDIX

APPENDIX: CASE STUDY, TABLES, AND FIGURES

APPENDIX: CASE STUDY, TABLES, AND FIGURES

Case Study: Applying the Taxonomy

The results of the meta-analysis could be applied in several different ways to design a smoking cessation intervention for pregnant women. The following paragraphs describe one such application, starting with the behavior change techniques that were identified as most effective, then selecting an appropriate theoretical foundation upon which to build the intervention, and finally, designing an effective measurement strategy to assess key determinants of behavior change.

Of the 8 behavior change techniques that were identified as “active ingredients” in the meta-analysis (e.g., techniques that had effect sizes that were significantly larger than the respective control groups), two were related to the provision of information, one was related to the provision of instructions, and four were related to goal setting, achievement, and/or rewards for achieving goals. These results align closely with the determinants of behavior change described by the Information-Motivation-Behavioral Skills (IMB) Model, which postulates that three main constructs influence health behavior change: information and knowledge about the behavior; motivation to perform the behavior; and behavioral skills to perform the behavior. Applied to smoking cessation for pregnant women, the key determinants can be grouped into the following domains and targeted with the associated BCTs:

Information

-Provide information about the general and health-related effects of smoking and quitting (for the woman)

-Provide information about the general and health-related effects of smoking and quitting (for the baby)

-Provide information about accessing smoking cessation-related resources

-Provide information about personal susceptibility

Example: Provide personalized information about the health and economic benefits of staying smoke-free at the two-day mark, the one week mark, the two week mark, the one month mark, etc., as participants reach those milestones.

Motivation

-Restructure social and physical environment to support smoking cessation

-Facilitate goal-setting activities

-Ask participant to sign behavioral contract to reinforce commitment to quitting smoking

-Provide rewards for achieving goals

Example: Women will be assigned to a trained cessation counselor ("quit buddy"), who will help participants set clear and achievable daily goals. If the goal is met, the participant will earn points to "spend" on a mobile app featuring music, podcasts, e-books, and other online prizes.

Behavioral Skills

-Help participants identify and utilize emotional and practical support from family and social circle

-Relapse prevention

Example: A trained cessation counselor will work with participants to identify likely challenges to remaining smoke-free during the postpartum period, then practice problem-solving activities aimed at minimizing those challenges.

Key behavioral determinants (e.g., knowledge about consequences of smoking, perceived social support, perceived risk/susceptibility, motivation to quit smoking, etc.) should be measured at baseline and again at the midpoint and end of the intervention, in order to look for changes and potential moderators. Key study outcomes would include self-reported smoking abstinence, biochemically validated smoking abstinence, reduction in cigarettes per day, and reduction in postpartum relapse.

APPENDIX: TABLES AND FIGURES

Aim One Tables and Figures

Table 1.0.1: Study Characteristics

Study Characteristics	No. (K)	%
Design		
RCT	36	95%
Cluster R	2	5%
Number of arms in trial		
Two	30	79%
More than two	8	21%
Risk of Bias		
Low	13	34%
High	12	32%
Unclear	13	34%
Sample size		
0-50 participants	3	8%
51-100 participants	7	18%
101-500 participants	23	61%
> 501 participants	5	13%

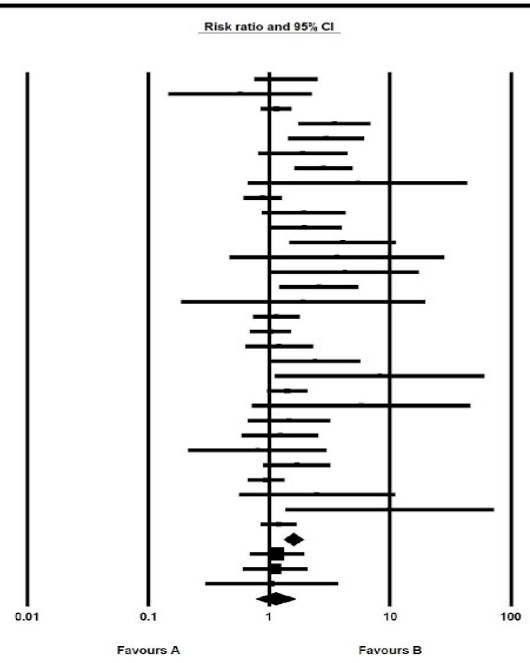
Table 1.0.2: Intervention Characteristics

Intervention Characteristic	No. (K)	%
Type of Intervention		
Single	**16**	**42%**
Multiple	**22**	**58%**
Main Intervention Strategy		
Counseling	**19**	**50%**
Vouchers/Incentives	**9**	**24%**
Social Support	**6**	**16%**
NRT (+supplement)	**3**	**8%**
Educational	**1**	**2%**
Deliverer		
Study personnel	**13**	**34%**
Mental health counselors	**8**	**21%**
Medical providers	**7**	**18%**
Peer educators	**4**	**11%**
Other	**3**	**8%**
Primary mode of Delivery		
Face-to-face	**19**	**50%**
Telephone, video, or computer	**16**	**42%**
Equal mix of both	**3**	**8%**
Setting (of trial)		
Community clinics	**12**	**32%**
Hospital-based clinics	**10**	**26%**
Medicaid/WIC clinics	**9**	**24%**
Other	**7**	**18%**
Setting (of delivery)		

Primarily within clinic	**24**	**64%**
Primarily outside of clinic	**14**	**36%**
Tailored		
For culture or ethnicity	**24**	**64%**
For smoking habits or beliefs	**4**	**11%**
Low SES sample		
Yes	**28**	**74%**
No	**10**	**26%**
Majority Minority sample		
Yes	**10**	**26%**
No	**28**	**74%**
High Psychosocial Risk		
Yes	**22**	**58%**
No	**16**	**42%**

Table 1.1.0. Primary Outcome: Late Pregnancy Smoking Abstinence

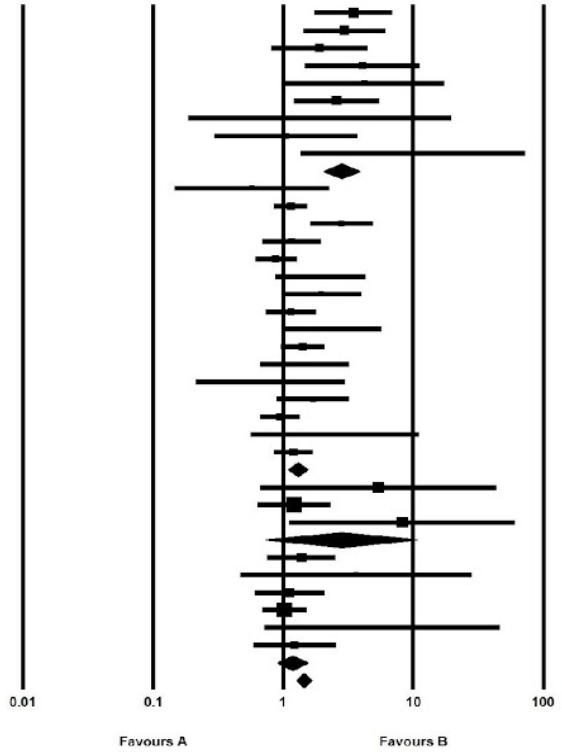
Group by Cultural Tailoring	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
						Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
0.00	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
0.00	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
0.00	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
0.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
0.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
0.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
0.00	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
0.00	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
0.00	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
0.00	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
0.00	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
0.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
0.00	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
0.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
0.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
0.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
0.00	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
0.00	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
0.00	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
0.00	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
0.00	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
0.00	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
0.00	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
0.00	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
0.00	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
0.00	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
0.00	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.692	0.895	3.201	1.618	0.106
0.00	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
0.00	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.564	10.892	1.202	0.230
0.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
0.00	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
0.00						1.591	1.332	1.900	5.125	0.000
1.00	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
1.00	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
1.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
1.00						1.133	0.781	1.643	0.658	0.510



Meta Analysis

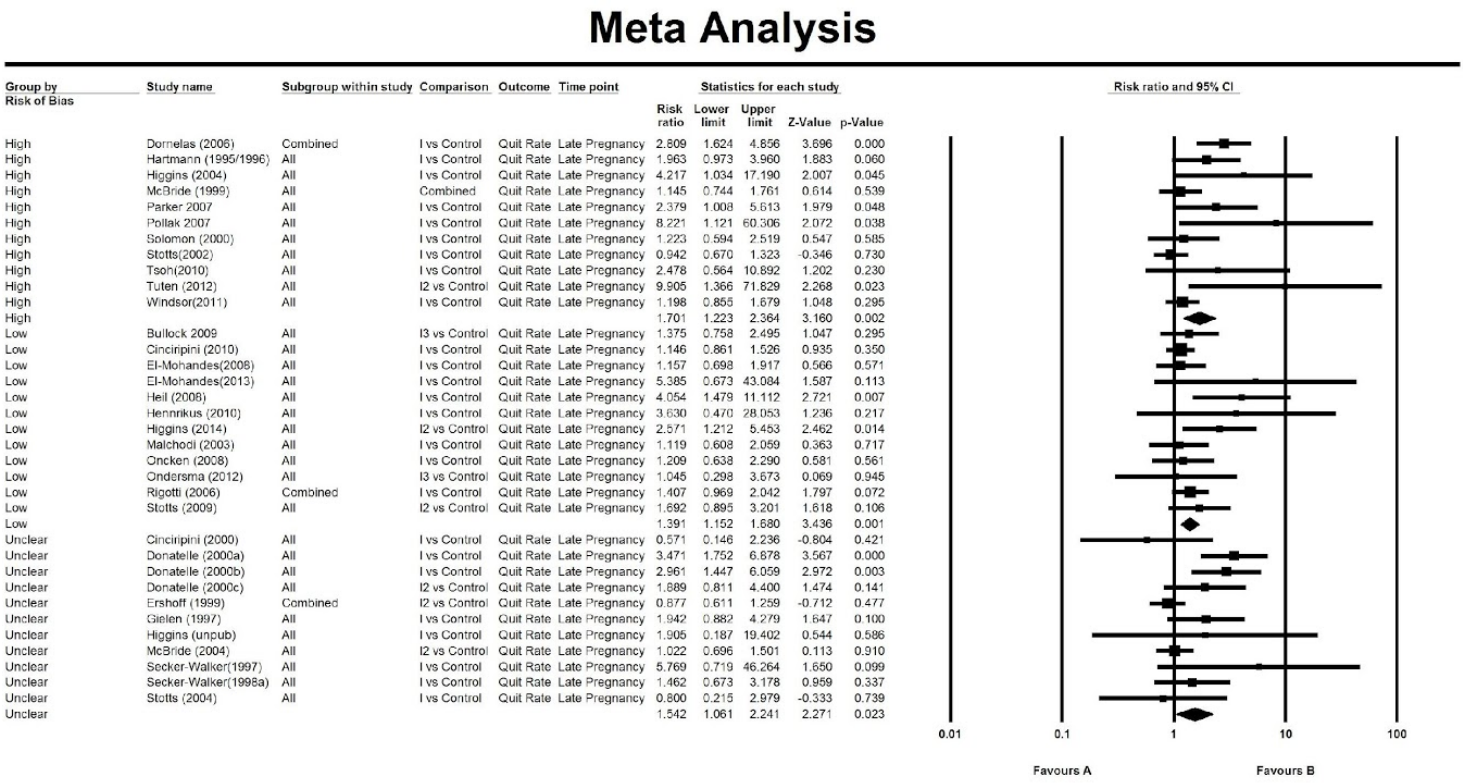
Table 1.1.1: Subgroup Analysis: Late pregnancy smoking cessation by intervention type

Group by Intervention Type	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
						Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
Contingent Vouchers	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
Contingent Vouchers	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
Contingent Vouchers	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
Contingent Vouchers	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
Contingent Vouchers	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
Contingent Vouchers	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
Contingent Vouchers	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
Contingent Vouchers	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
Contingent Vouchers	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
Contingent Vouchers						2.819	2.049	3.878	6.365	0.000
Counseling	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
Counseling	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
Counseling	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
Counseling	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
Counseling	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
Counseling	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
Counseling	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
Counseling	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
Counseling	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
Counseling	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
Counseling	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
Counseling	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
Counseling	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.692	0.895	3.201	1.618	0.106
Counseling	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
Counseling	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.564	10.892	1.202	0.230
Counseling	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
Counseling						1.298	1.096	1.538	3.018	0.003
NRT + Counseling	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
NRT + Counseling	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
NRT + Counseling	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
NRT + Counseling						2.808	0.737	10.697	1.513	0.130
Social Support	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
Social Support	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
Social Support	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
Social Support	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
Social Support	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
Social Support	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
Social Support						1.182	0.910	1.534	1.253	0.210
Overall						1.451	1.275	1.651	5.646	0.000



Meta Analysis

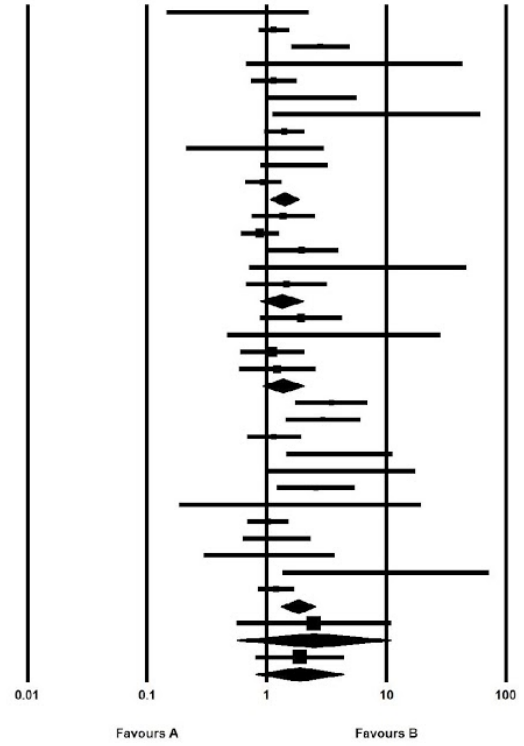
Table 1.1.2: Subgroup Analysis: Late pregnancy smoking abstinence by risk of bias



Meta Analysis

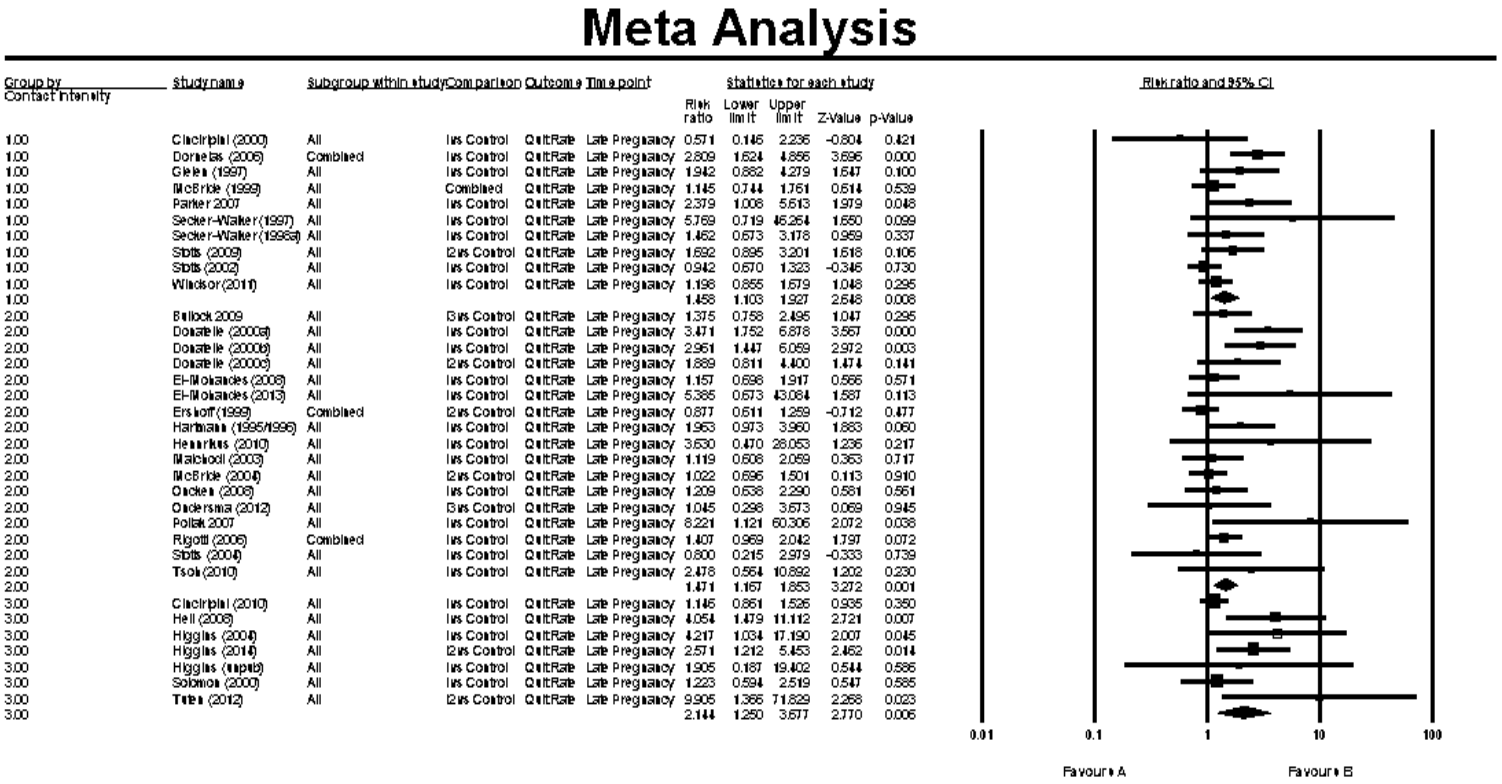
Table 1.1.3: Subgroup Analysis: Late Pregnancy smoking abstinence by intervention deliverer

Group by Deliverer	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
						Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
Counselor	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
Counselor	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
Counselor	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
Counselor	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
Counselor	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
Counselor	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
Counselor	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
Counselor	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
Counselor	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
Counselor	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.692	0.895	3.201	1.618	0.106
Counselor	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
Counselor						1.415	1.080	1.854	2.516	0.012
Medical Provider	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
Medical Provider	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
Medical Provider	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
Medical Provider	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
Medical Provider	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
Medical Provider						1.346	0.889	2.037	1.405	0.160
Peer	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
Peer	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
Peer	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
Peer	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
Peer						1.377	0.929	2.041	1.591	0.112
Study Personnel	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
Study Personnel	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
Study Personnel	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
Study Personnel	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
Study Personnel	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
Study Personnel	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
Study Personnel	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
Study Personnel	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
Study Personnel	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
Study Personnel	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
Study Personnel	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
Study Personnel	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
Study Personnel						1.835	1.311	2.569	3.537	0.000
Tech	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.564	10.892	1.202	0.230
Tech						2.478	0.564	10.892	1.202	0.230
Volunteers	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
Volunteers						1.889	0.811	4.400	1.474	0.141



Meta Analysis

Table 1.1.4: Subgroup Analysis: Late pregnancy smoking abstinence by contact intensity



Meta Analysis

Table 1.1.5: Subgroup Analysis: Late Pregnancy smoking abstinence by context of delivery (within prenatal care [y/n])

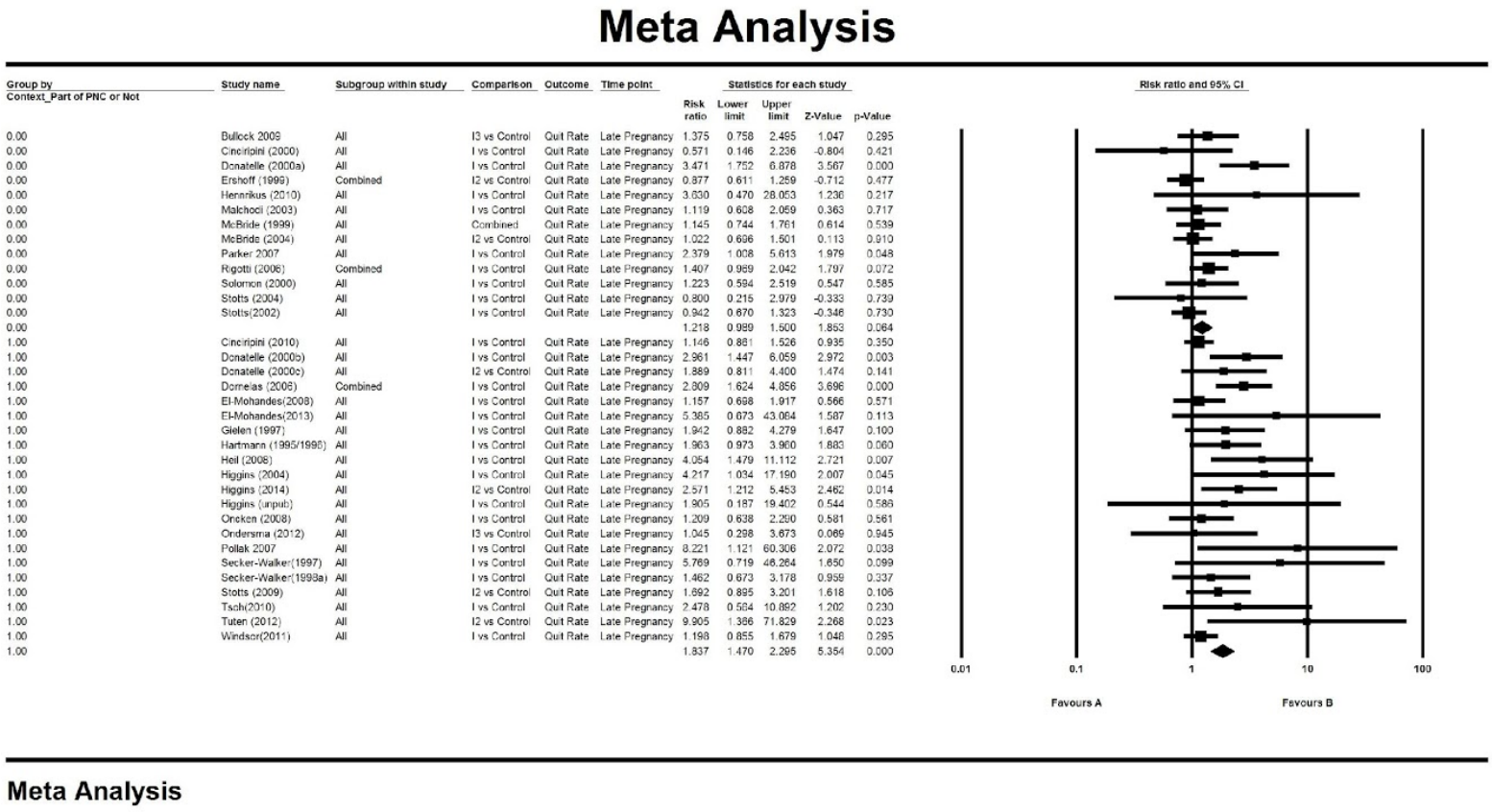
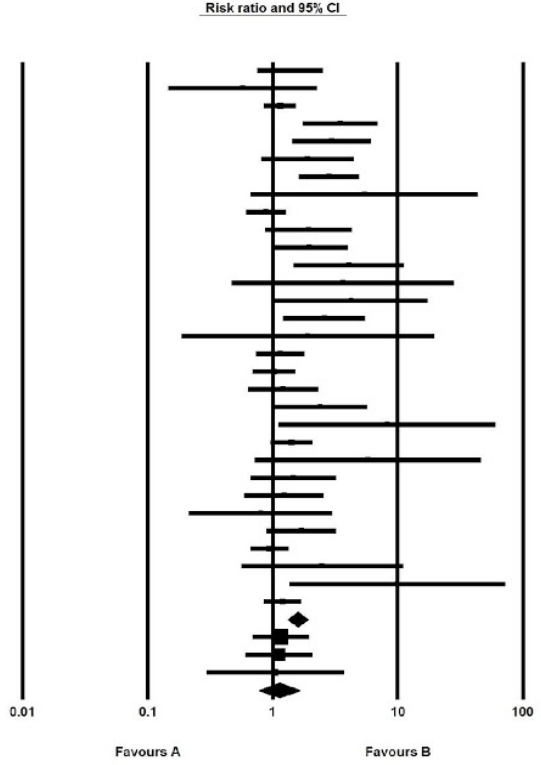


Table 1.1.6: Subgroup Analysis: Late Pregnancy smoking abstinence by cultural tailoring

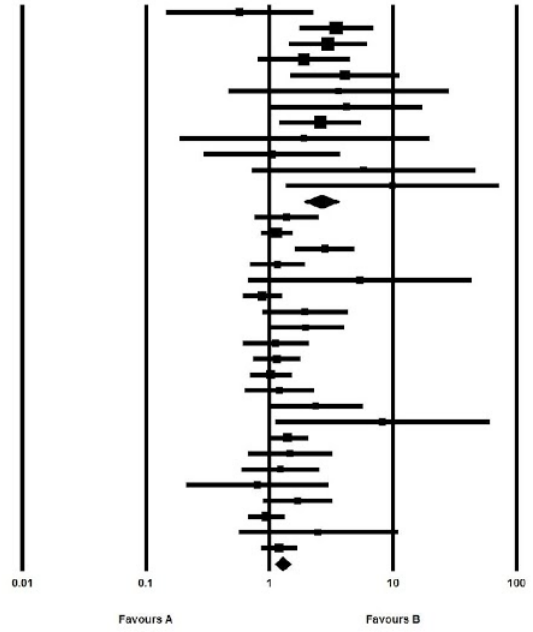
Group by Cultural Tailoring	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
						Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
0.00	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
0.00	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
0.00	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
0.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
0.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
0.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
0.00	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
0.00	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
0.00	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
0.00	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
0.00	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
0.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
0.00	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
0.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
0.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
0.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
0.00	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
0.00	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
0.00	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
0.00	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
0.00	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
0.00	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
0.00	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
0.00	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
0.00	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
0.00	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
0.00	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.692	0.895	3.201	1.618	0.106
0.00	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
0.00	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.564	10.892	1.202	0.230
0.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
0.00	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
0.00						1.591	1.332	1.900	5.125	0.000
1.00	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
1.00	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
1.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
1.00						1.133	0.781	1.643	0.658	0.510



Meta Analysis

Table 1.1.7: Subgroup Analysis: Late Pregnancy smoking abstinence by organizational/provider-level strategies

Group by Organizational-Provider Level Strategies	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study: Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
0.00	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
0.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
0.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
0.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.689	0.811	4.400	1.474	0.141
0.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
0.00	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
0.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
0.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
0.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
0.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
0.00	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
0.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.386	71.829	2.268	0.023
0.00						2.647	1.919	3.650	5.938	0.000
1.00	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
1.00	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.148	0.881	1.526	0.935	0.350
1.00	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.896	0.000
1.00	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
1.00	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
1.00	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
1.00	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.892	4.279	1.647	0.100
1.00	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
1.00	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
1.00	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
1.00	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
1.00	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
1.00	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
1.00	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
1.00	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
1.00	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
1.00	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
1.00	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
1.00	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.692	0.895	3.201	1.618	0.106
1.00	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
1.00	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.584	10.892	1.202	0.230
1.00	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
1.00						1.285	1.118	1.478	3.527	0.000



Meta Analysis

Table 1.1.8: Subgroup Analysis: Late Pregnancy smoking abstinence by SES

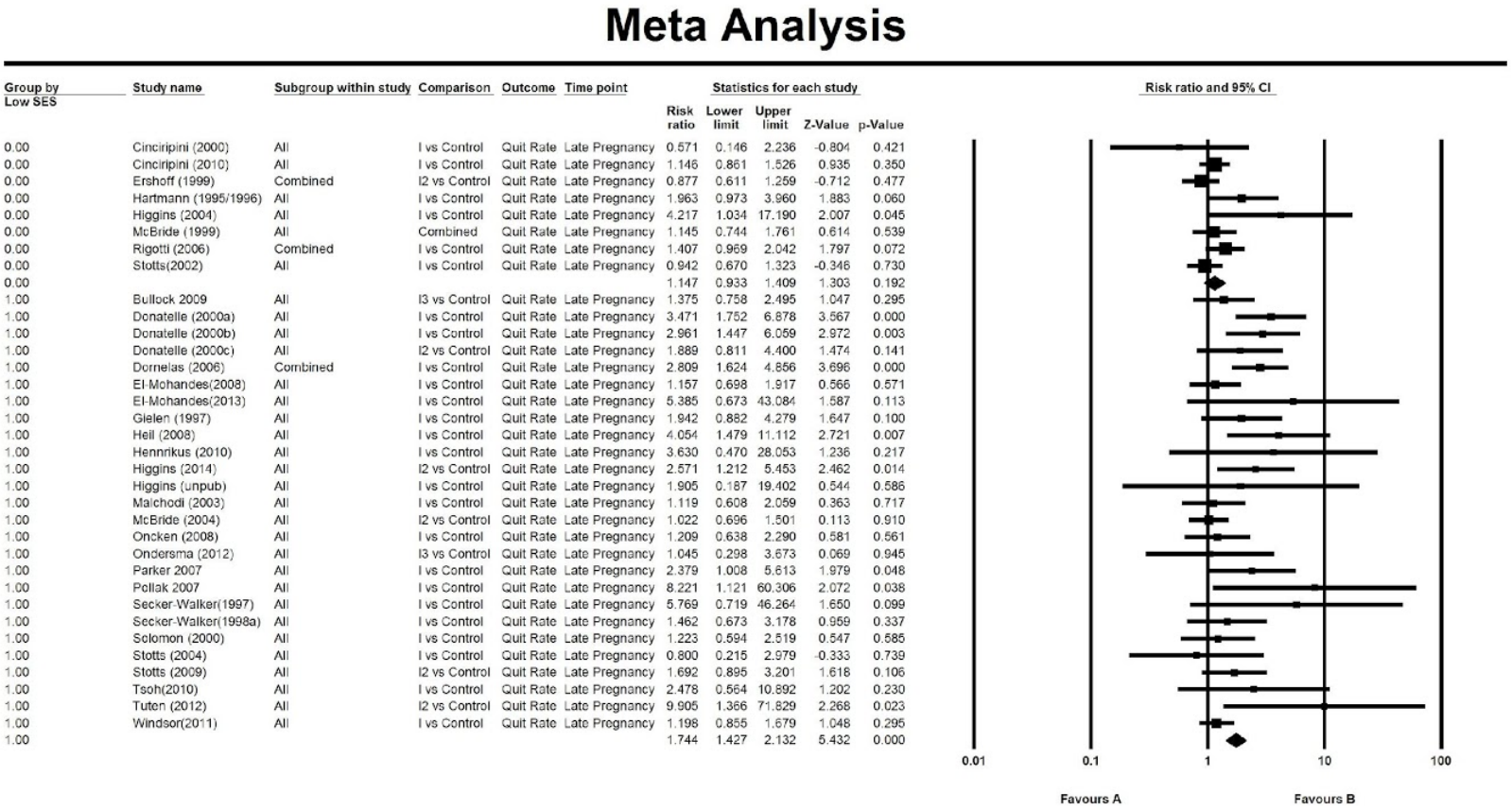
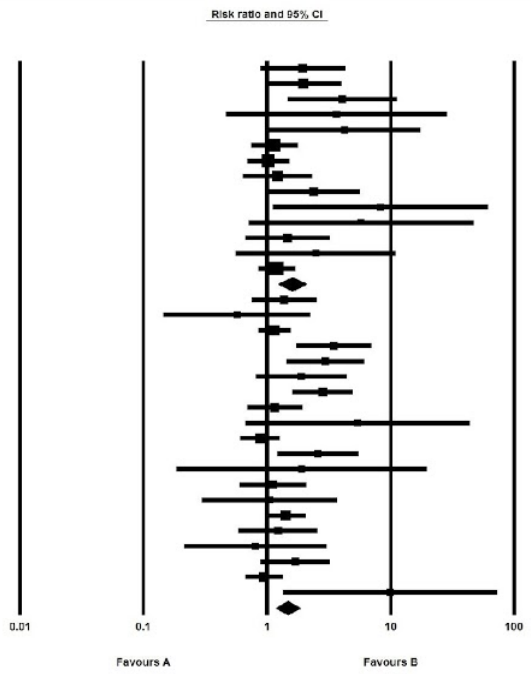


Table 1.1.9: Subgroup Analysis: Late Pregnancy smoking abstinence by psychosocial risk

Meta Analysis

Group by High Psychosocial Risk	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
						Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
0.00	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
0.00	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
0.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
0.00	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
0.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
0.00	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
0.00	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
0.00	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
0.00	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
0.00	Polisk 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
0.00	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.284	1.650	0.099
0.00	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
0.00	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.564	10.892	1.202	0.230
0.00	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.188	0.855	1.679	1.048	0.295
0.00						1.596	1.240	2.055	3.625	0.000
1.00	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.756	2.495	1.047	0.295
1.00	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
1.00	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
1.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
1.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
1.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
1.00	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
1.00	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
1.00	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
1.00	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
1.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
1.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
1.00	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
1.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
1.00	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
1.00	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
1.00	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
1.00	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.682	0.895	3.201	1.618	0.106
1.00	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.348	0.730
1.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
1.00						1.480	1.192	1.837	3.546	0.000

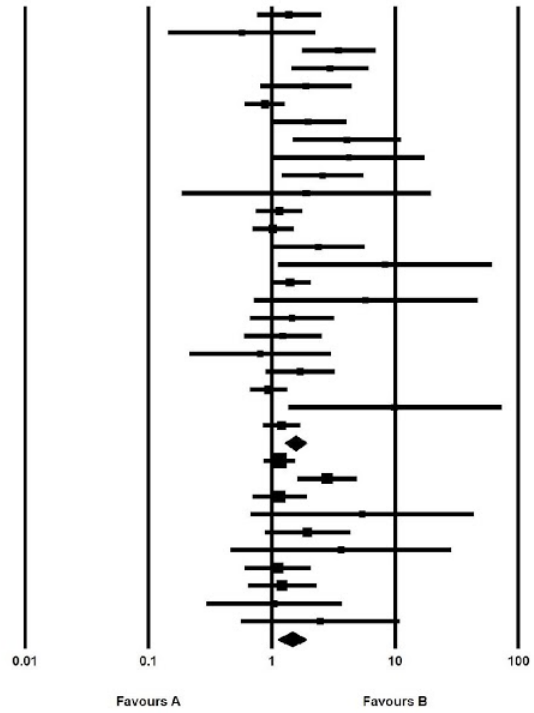


Meta Analysis

Table 1.1.10: Subgroup Analysis: Late Pregnancy smoking abstinence by majority minority sample

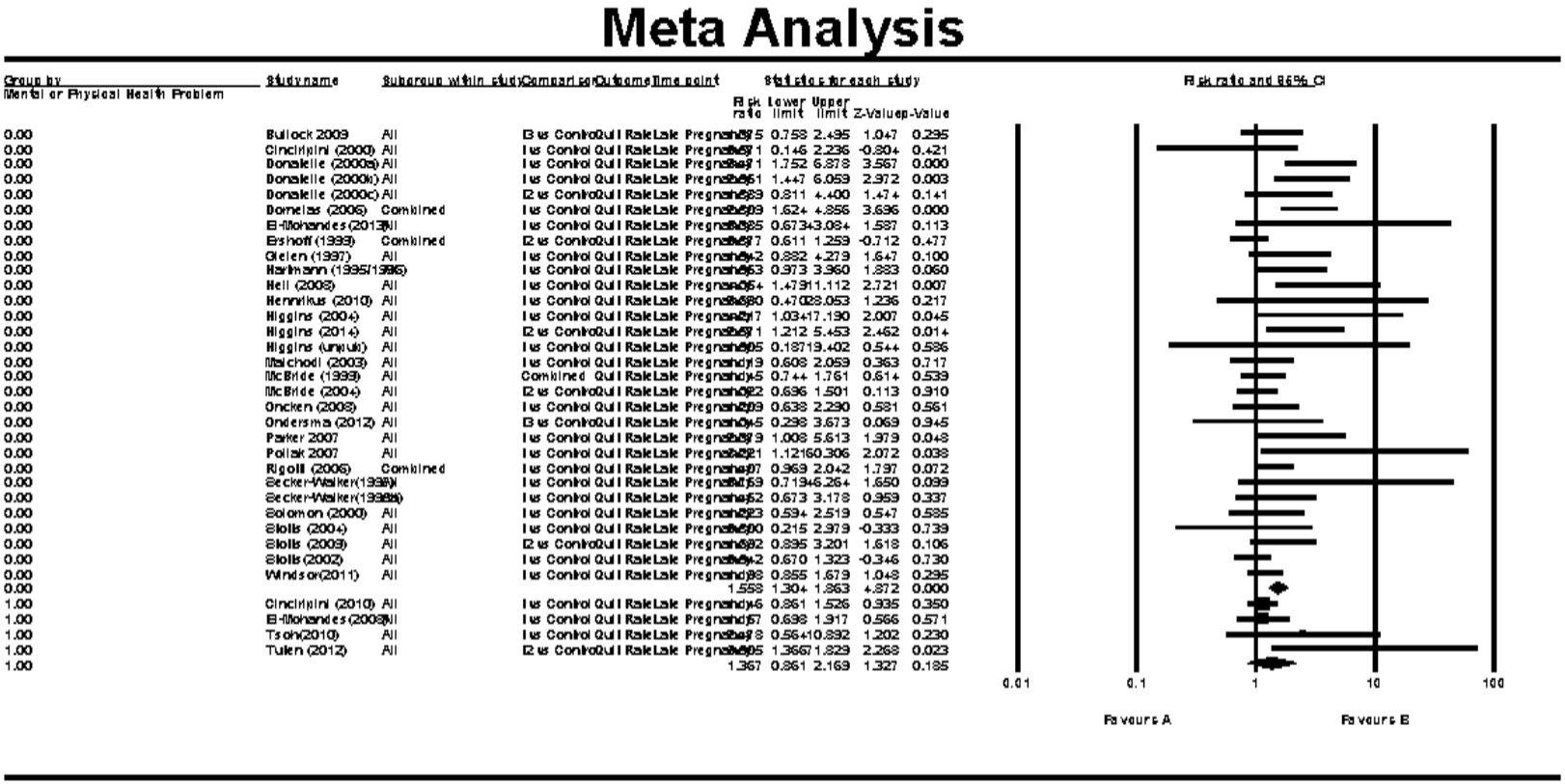
Meta Analysis

Group by Majority Minority	Study name	Subgroup within study	Comparison	Outcome	Time point	Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
0.00	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
0.00	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
0.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
0.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
0.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
0.00	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
0.00	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
0.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
0.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
0.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
0.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
0.00	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
0.00	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
0.00	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
0.00	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
0.00	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
0.00	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
0.00	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
0.00	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
0.00	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
0.00	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.692	0.895	3.201	1.618	0.106
0.00	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
0.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
0.00	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
0.00						1.564	1.274	1.920	4.277	0.000
1.00	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
1.00	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
1.00	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
1.00	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
1.00	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
1.00	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
1.00	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
1.00	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
1.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
1.00	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.564	10.892	1.202	0.230
1.00						1.464	1.118	1.917	2.774	0.006



Meta Analysis

Table 1.1.11: Subgroup Analysis: Late Pregnancy smoking abstinence by health status



Meta Analysis

Table 1.1.12: Meta-regression: Late Pregnancy smoking abstinence by baseline smoking (cig./day)

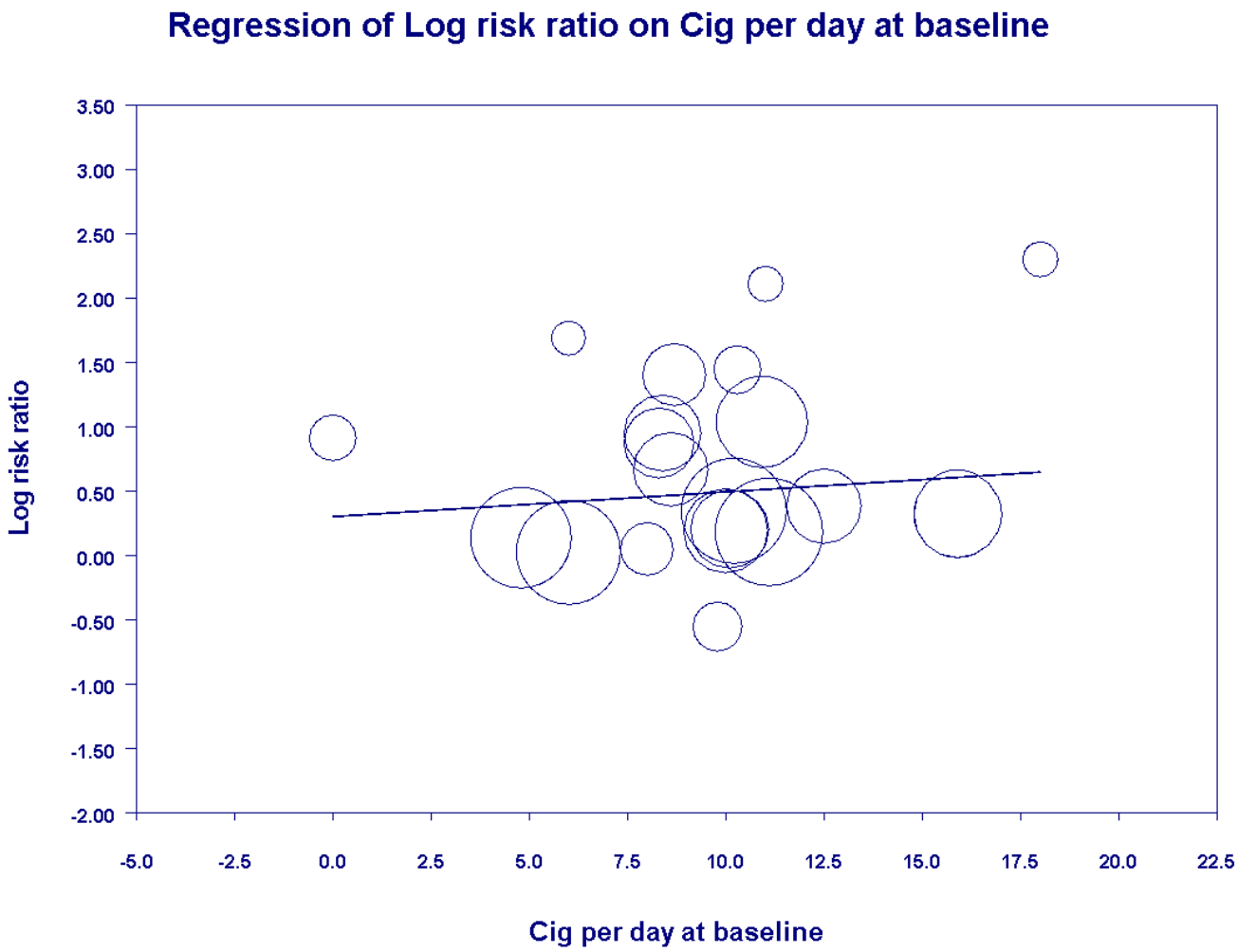


Table 1.1.13: Meta-regression: Late pregnancy smoking cessation by gestational age

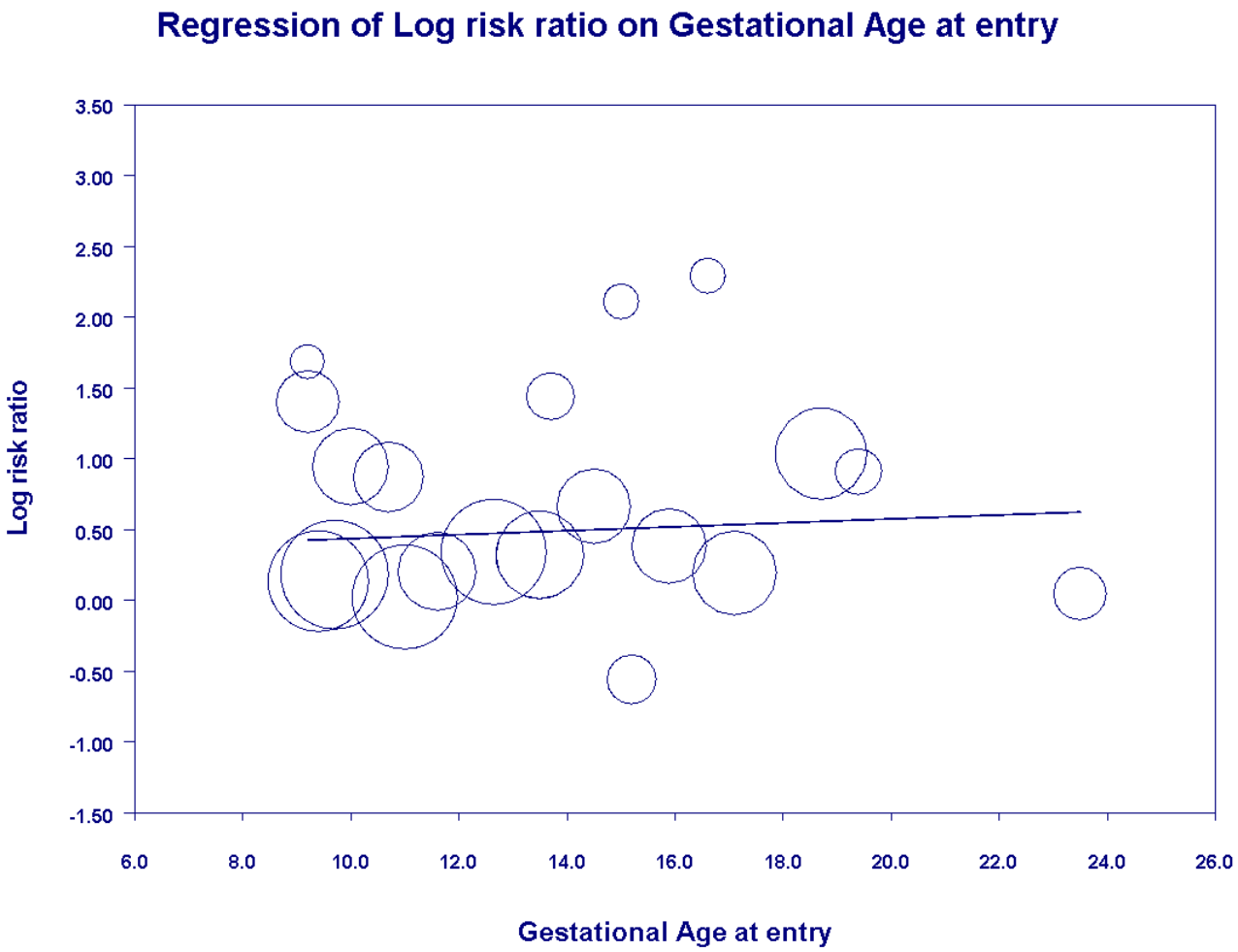
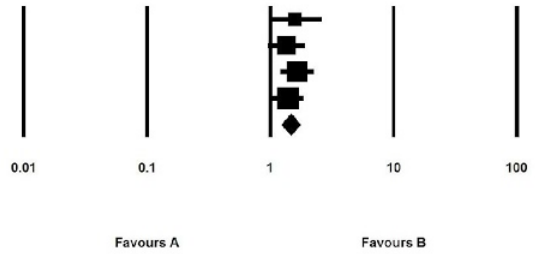


Table 1.2.0. Secondary Outcome: Significant Reduction in Smoking

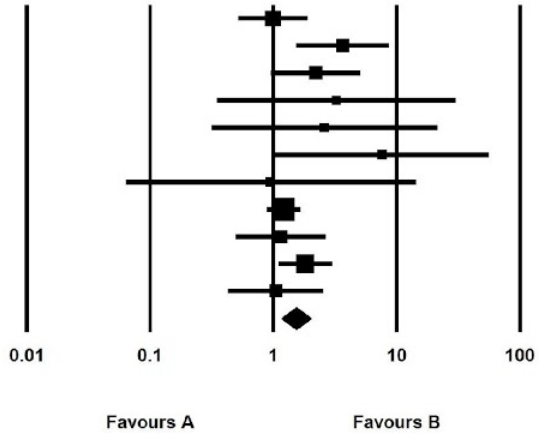
Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
					Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
Hartmann (1995/1998)	All	I vs Control	Reduction in smoking by at least half	Late Pregnancy	1.589	0.983	2.569	1.889	0.059
Rigotti (2006)	All	I vs Control	Reduction in smoking by at least half	Late Pregnancy	1.345	0.965	1.874	1.752	0.080
Secker-Walker(1998a)	All	I vs Control	All (Indep)	Combined	1.643	1.219	2.215	3.264	0.001
Windsor(2011)	All	I vs Control	Sig Reduction	Late Pregnancy	1.386	1.049	1.832	2.297	0.022
					1.471	1.250	1.732	4.635	0.000



Meta Analysis

Table 1.2.1. Secondary Outcome: Early (<6 mos) Postpartum Smoking Cessation

Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
					Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
Bullock 2009	All	I3 vs Control	Quit Rate	0-5 Months PP	0.991	0.523	1.879	-0.026	0.979
Donatelle (2000a)	All	I vs Control	Quit Rate	0-5 Months PP	3.631	1.536	8.582	2.938	0.003
El-Mohandes(2008)	All	I vs Control	Quit Rate	0-5 Months PP	2.204	0.965	5.036	1.874	0.061
Heil (2008)	All	I vs Control	Quit Rate	0-5 Months PP	3.243	0.353	29.819	1.039	0.299
Hennrikus (2010)	All	I vs Control	Quit Rate	0-5 Months PP	2.593	0.318	21.128	0.890	0.373
Higgins (2004)	All	I vs Control	Quit Rate	0-5 Months PP	7.667	1.056	55.670	2.014	0.044
Higgins (unpub)	All	I vs Control	Quit Rate	0-5 Months PP	0.952	0.064	14.218	-0.035	0.972
McBride (1999)	All	Combined	Quit Rate	0-5 Months PP	1.209	0.891	1.641	1.221	0.222
Oncken (2008)	All	I vs Control	Quit Rate	0-5 Months PP	1.149	0.499	2.647	0.326	0.745
Stotts(2002)	All	I vs Control	Quit Rate	0-5 Months PP	1.813	1.109	2.965	2.373	0.018
Strecher(2000)	All	I vs Control	Quit Rate	0-5 Months PP	1.046	0.432	2.534	0.099	0.921
					1.537	1.161	2.034	3.005	0.003



Meta Analysis

Table 1.2.2. Secondary Outcome: Late (6 mos+) Postpartum Smoking Cessation

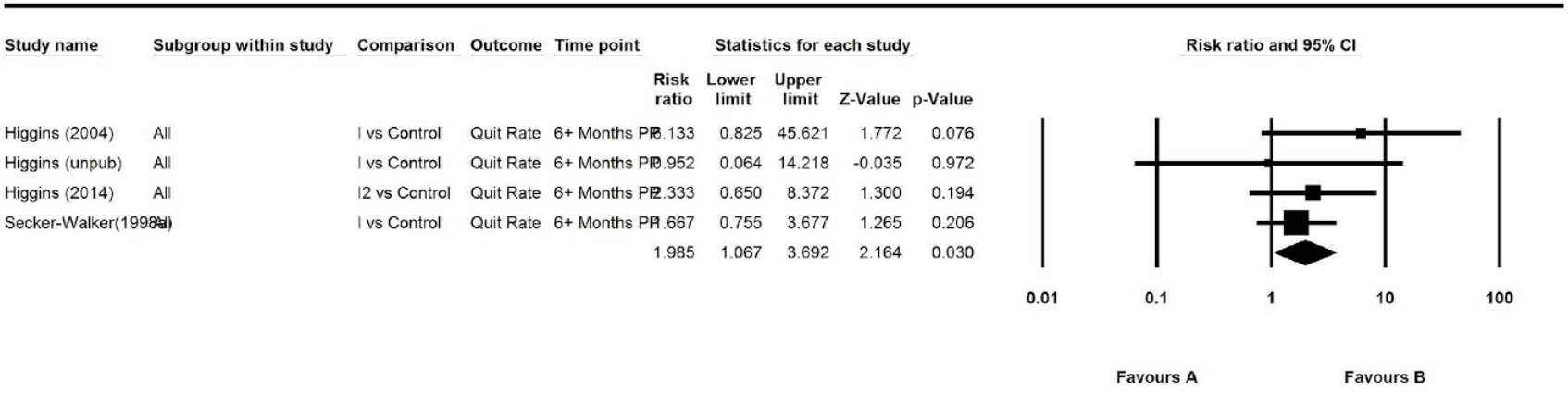
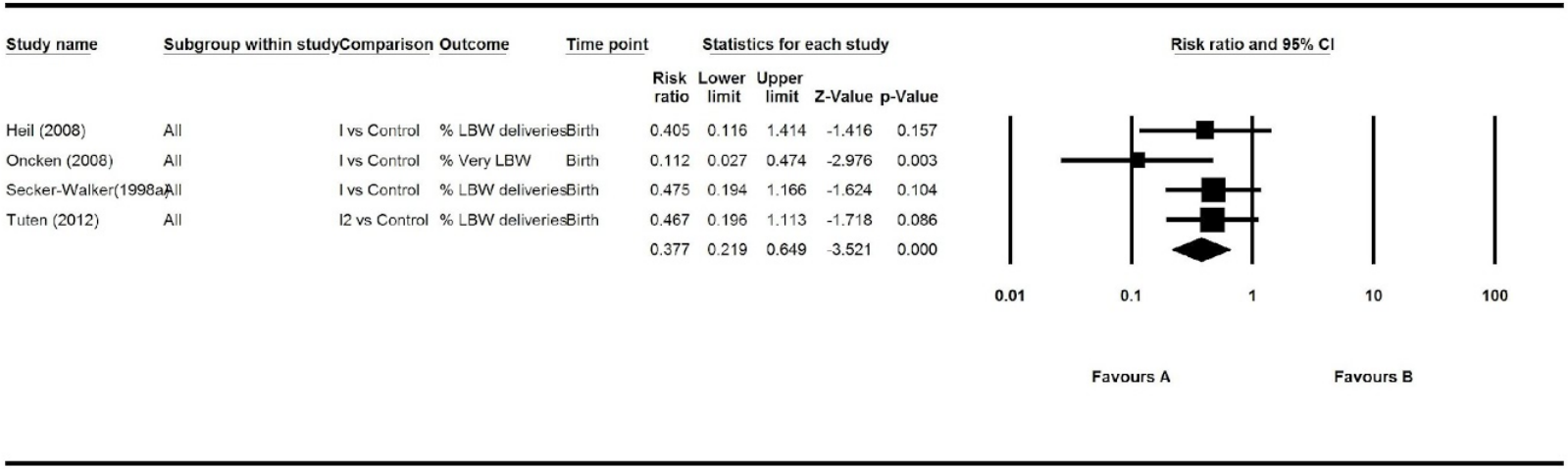


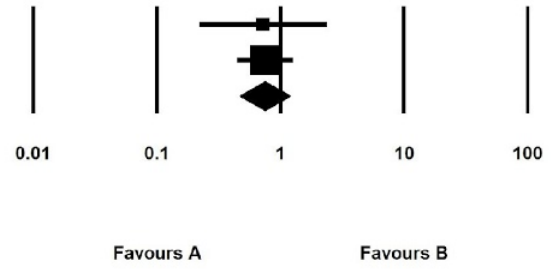
Table 1.2.3. Secondary Outcome: Low Birthweight or Very Low Birthweight



Meta Analysis

Table 1.2.4. Secondary Outcome: NICU Admissions

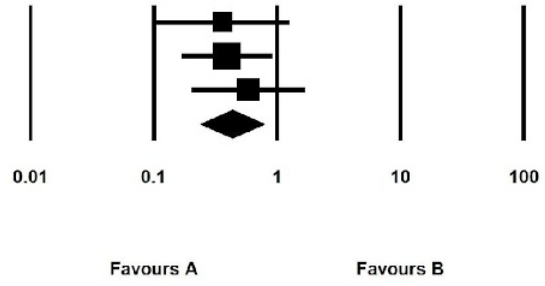
Study name	Subgroup within study	Comparison	Outcome	Time point	Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
Heil (2008)	All	I vs Control	% NICU admission	Birth	0.721	0.221	2.354	-0.542	0.588
Tuten (2012)	All	I2 vs Control	% NICU Admission	Birth	0.754	0.453	1.254	-1.088	0.276
					0.749	0.469	1.195	-1.214	0.225



Meta Analysis

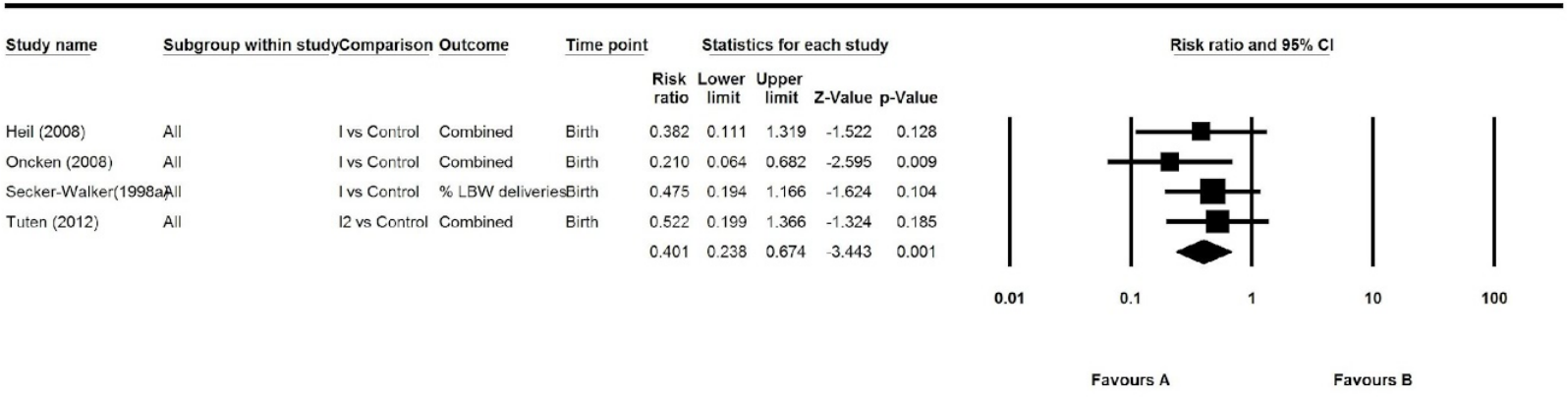
Table 1.2.5. Secondary Outcome: Preterm Birth

Subgroup within study	Comparison	Outcome	Time point	Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
Heil (2008) All	I vs Control	% Preterm birth	Birth	0.360	0.106	1.230	-1.629	0.103
Oncken (2008) All	I vs Control	% Preterm birth	Birth	0.392	0.169	0.909	-2.184	0.029
Tuten (2012) All	I2 vs Control	% Preterm birth	Birth	0.583	0.205	1.663	-1.008	0.313
				0.434	0.244	0.774	-2.827	0.005



Meta Analysis

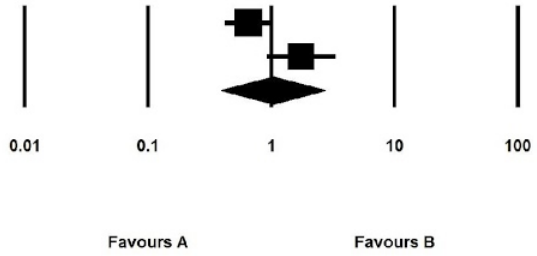
Table 1.2.6. Secondary Outcome: Preterm Birth or Low Birthweight/Very Low Birthweight



Meta Analysis

Table 1.2.7. Secondary Outcome: Any serious adverse event

Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study				
					Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
Oncken (2008)	All	I vs Control	Any serious adverse event	Birth	0.652	0.421	1.012	-1.908	0.056
Pollak 2007	All	I vs Control	Serious adverse events	Birth	1.745	0.929	3.277	1.732	0.083
					1.039	0.396	2.720	0.077	0.939



Meta Analysis

APPENDIX: TABLES AND FIGURES

Aim 2 Tables & Figures

Table 2.1.0. Theory Coding Scheme (kappa and % yes for each item)

Item no. (& Kappa)	Item	Description	Yes (N)	%
1 (0.94)	Explicit mention of the use of health behavior theory	The study **explicitly mentioned** using a health behavior theory (or model), defined as "a set of interrelated concepts, definitions, and propositions that presents a *systematic* view of events or situations by specifying relations among variables in order to *explain* and *predict* events or situations" (Glanz et al., 1997, p. 21). *Note: this is an independent assessment from their **actual use** of theory.	26	68%
2 (0.67)	Targeted constructs mentioned as predictors of behavior.	1) The study explicitly mentioned how targeted constructs are theorized to predict behavior, where "targeted constructs" refer to theoretical constructs that the intervention is hypothesized to change, **AND** 2) The study provided evidence that the construct targeted relates to behavior in the introduction or methods section (not discussion section).	26	68%
3 (0.69)	Intervention based on a single theory	The intervention is based on a single theory, rather than a combination of theories or theory and predictors.	9	24%
4 (0.78)	Theory used to select participants	Participants were screened/selected based on achieving a particular score/level on a theory-relevant construct/predictor.	1	3%
5 (0.71)	Theory used to select and/or develop intervention techniques	The intervention techniques are explicitly based on a theory or predictor or combination of theories or predictors.	25	66%

6 (0.84)	Theory used to tailor intervention techniques to participants	The intervention differs for different sub-groups that vary on a psychological construct or predictor at baseline.	8	21%
7 (0.73)	**ALL** intervention techniques are explicitly linked to at least one theory-relevant construct/predictor	Each intervention technique is explicitly linked to at least one theory-relevant construct/predictor.	1	3%
8 (0.77)	At least one, but not all, of the intervention techniques are explicitly linked to at least one theory-relevant construct/ predictor.	At least one, but not all, of the intervention techniques are explicitly linked to at least one theory-relevant construct/predictor.	19	50%
9 (0.74)	Group of techniques are linked to a group of constructs/ predictors	A cluster of techniques is linked to a cluster of constructs/predictors	5	13%
10 (0.70)	All theory-relevant constructs/predictors are explicitly linked to at least	Every theoretical construct within a stated theory, or every stated predictor (see item 5), is linked to at least one intervention technique.	2	5%

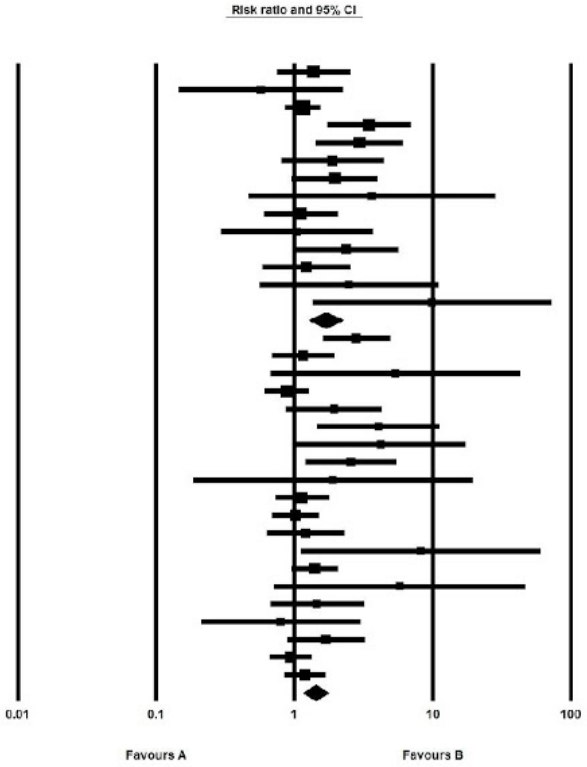
	one intervention technique			
11 (0.67)	**At least one**, but not all, of the theory relevant constructs/predictors are explicitly linked to at least one intervention technique.	At least one, but not all, of the theoretical constructs within a stated theory **or** at least one, but not all, of the stated predictors (see item 5) are linked to at least one intervention technique.	21	55%
12 (0.94)	Theory-relevant constructs/ predictors are measured	a) At least one construct of theory (or predictor) mentioned in relation to the intervention is measured POST-INTERVENTION. **OR**	5	13%
		b) At least one construct of theory (or predictor) mentioned in relation to the intervention is measured PRE AND POST-INTERVENTION.		
13 (Mean *k* = 0.81)	Quality of Measures	a) All of the measures of theory relevant constructs/predictors had some evidence for their reliability.	2	5%
		b) At least one, but not all, of the measures of theory relevant constructs/predictors had some evidence for their reliability	3	8%
		c) All of the measures of theory relevant constructs/predictors have been previously validated	2	5%
		d) At least one, but not all, of the measures of theory relevant constructs/predictors have been previously validated	3	8%
		e) The behavior measure had some evidence for its reliability	38	100%

		f) The behavior measure has been previously validated	38	100%
14 (0.95)	Randomization of participants to condition	a) Do the authors claim randomization?	38	100%
		b) Is a method of random allocation to condition described (e.g., random number generator; coin toss)	25	66%
		c) Was the success of randomization tested?	38	100%
		d) Was the randomization successful (or baseline differences between intervention and control group statistically controlled)?	38	100%
15 (0.98)	Changes in measured theory-relevant constructs/predictors	The intervention leads to sig. change in at least one theory-relevant construct/predictor (vs.control group) in favor of the intervention group.	3	8%
16 (1)	Mediational analysis of construct/s / predictors	In addition to 14, do the following effects emerge?:	0	0%
		a) Mediator predicts DV? (or change in mediator leads to change in DV)	0	0%
		b) Mediator predicts DV (when controlling for IV)?	0	0%
		c) Intervention does not predict DV (when controlling for mediator)?	0	0%
		d) Mediated effect statistically significant?	0	0%
17 (0.71)	Results discussed in relation to theory	Results are discussed in terms of the theoretical basis of the intervention	20	53%

18 (1)	Appropriate support for theory	Support for the theory is based on appropriate mediation OR refutation of the theory is based on obtaining appropriate null effects (i.e. changing behavior without changing the theory-relevant constructs).	0	0%
19 (1)	Results used to refine theory	The authors attempt to refine the theory upon which the intervention was based by either: a) adding or removing constructs to the theory, or b) specifying that the interrelationships between the theoretical constructs should be changed and spelling out which relationships should be changed.	0	0%

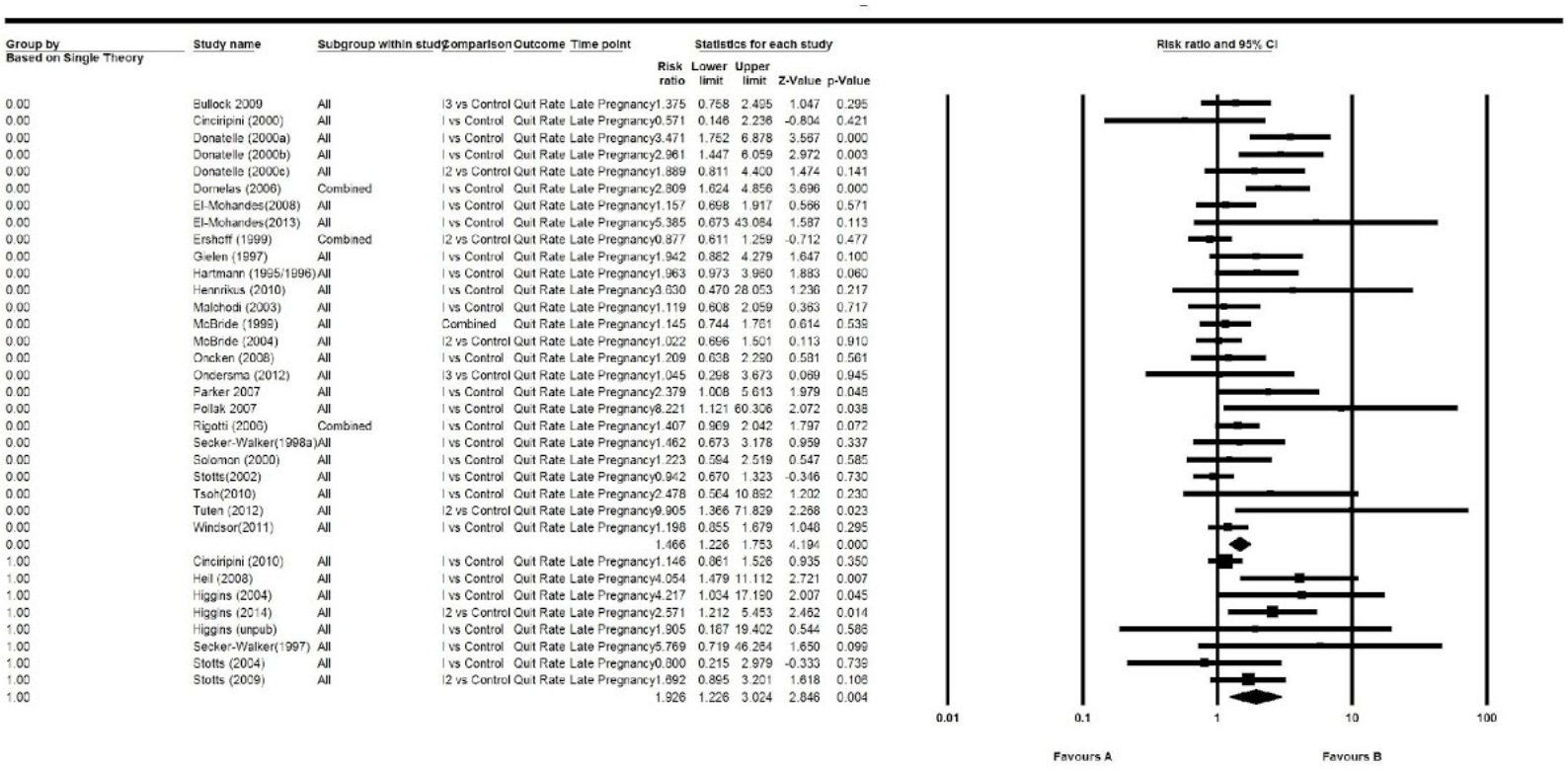
Table 2.1.1. Late pregnancy smoking abstinence by explicit mention of theory (y/n)

Group by Explicit Mention of Theory	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study: Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
0.00	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
0.00	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
0.00	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
0.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
0.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
0.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
0.00	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
0.00	Henrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
0.00	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
0.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
0.00	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
0.00	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
0.00	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.584	10.892	1.202	0.230
0.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.386	71.829	2.268	0.023
0.00						1.696	1.282	2.244	3.698	0.000
1.00	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
1.00	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
1.00	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
1.00	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
1.00	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
1.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
1.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
1.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
1.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
1.00	McBride (1999)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
1.00	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
1.00	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
1.00	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
1.00	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.989	2.042	1.797	0.072
1.00	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
1.00	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
1.00	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
1.00	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.692	0.895	3.201	1.616	0.106
1.00	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
1.00	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
1.00						1.438	1.176	1.758	3.536	0.000



Meta Analysis

Table 2.1.2. Late pregnancy smoking abstinence by *single* theory-based (y/n)



Group by Based on Single Theory	Study name	Subgroup within study	Comparison	Outcome	Time point	Statistics for each study: Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
0.00	Bullock 2009	All	I3 vs Control	Quit Rate	Late Pregnancy	1.375	0.758	2.495	1.047	0.295
0.00	Cinciripini (2000)	All	I vs Control	Quit Rate	Late Pregnancy	0.571	0.146	2.236	-0.804	0.421
0.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
0.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
0.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
0.00	Dornelas (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	2.809	1.624	4.856	3.696	0.000
0.00	El-Mohandes(2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.157	0.698	1.917	0.566	0.571
0.00	El-Mohandes(2013)	All	I vs Control	Quit Rate	Late Pregnancy	5.385	0.673	43.084	1.587	0.113
0.00	Ershoff (1999)	Combined	I2 vs Control	Quit Rate	Late Pregnancy	0.877	0.611	1.259	-0.712	0.477
0.00	Gielen (1997)	All	I vs Control	Quit Rate	Late Pregnancy	1.942	0.882	4.279	1.647	0.100
0.00	Hartmann (1995/1996)	All	I vs Control	Quit Rate	Late Pregnancy	1.963	0.973	3.960	1.883	0.060
0.00	Hennrikus (2010)	All	I vs Control	Quit Rate	Late Pregnancy	3.630	0.470	28.053	1.236	0.217
0.00	Malchodi (2003)	All	I vs Control	Quit Rate	Late Pregnancy	1.119	0.608	2.059	0.363	0.717
0.00	McBride (1998)	All	Combined	Quit Rate	Late Pregnancy	1.145	0.744	1.761	0.614	0.539
0.00	McBride (2004)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.022	0.696	1.501	0.113	0.910
0.00	Oncken (2008)	All	I vs Control	Quit Rate	Late Pregnancy	1.209	0.638	2.290	0.581	0.561
0.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
0.00	Parker 2007	All	I vs Control	Quit Rate	Late Pregnancy	2.379	1.008	5.613	1.979	0.048
0.00	Pollak 2007	All	I vs Control	Quit Rate	Late Pregnancy	8.221	1.121	60.306	2.072	0.038
0.00	Rigotti (2006)	Combined	I vs Control	Quit Rate	Late Pregnancy	1.407	0.969	2.042	1.797	0.072
0.00	Secker-Walker(1998a)	All	I vs Control	Quit Rate	Late Pregnancy	1.462	0.673	3.178	0.959	0.337
0.00	Solomon (2000)	All	I vs Control	Quit Rate	Late Pregnancy	1.223	0.594	2.519	0.547	0.585
0.00	Stotts(2002)	All	I vs Control	Quit Rate	Late Pregnancy	0.942	0.670	1.323	-0.346	0.730
0.00	Tsoh(2010)	All	I vs Control	Quit Rate	Late Pregnancy	2.478	0.564	10.892	1.202	0.230
0.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
0.00	Windsor(2011)	All	I vs Control	Quit Rate	Late Pregnancy	1.198	0.855	1.679	1.048	0.295
0.00						1.466	1.226	1.753	4.194	0.000
1.00	Cinciripini (2010)	All	I vs Control	Quit Rate	Late Pregnancy	1.146	0.861	1.526	0.935	0.350
1.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
1.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
1.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
1.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.167	19.402	0.544	0.586
1.00	Secker-Walker(1997)	All	I vs Control	Quit Rate	Late Pregnancy	5.769	0.719	46.264	1.650	0.099
1.00	Stotts (2004)	All	I vs Control	Quit Rate	Late Pregnancy	0.800	0.215	2.979	-0.333	0.739
1.00	Stotts (2009)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.892	0.895	3.201	1.618	0.106
1.00						1.926	1.226	3.024	2.846	0.004

Meta Analysis

Table 2.2.1. Meta-Regression: Late pregnancy smoking abstinence regressed on TCS Category 1 Score

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.6793	0.2031	0.2812	1.0775	3.34	0.0008
TCS Cat1 Total	-0.1508	0.107	-0.3604	0.0589	-1.41	0.1587

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 1.99, df = 1, p = 0.1587
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0804, Tau = 0.2836, I^2 = 44.83%, Q = 58.00, df = 32, p = 0.0033

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.08

Table 2.2.2. Meta-Regression: Late pregnancy smoking abstinence regressed on TCS Category 2 Score

Main results for Model 2, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.5441	0.146	0.2579	0.8303	3.73	0.0002
TCS Cat2 Total_Constructs Targeted	-0.0449	0.0451	-0.1332	0.0434	-1	0.3193

Statistics for Model 2

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 0.99, df = 1, p = 0.3193
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0926, Tau = 0.3043, I^2 = 47.96%, Q = 61.49, df = 32, p = 0.0013

Comparison of Model 2 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 2
R^2 analog = 0.00 (computed value is -0.06)

Table 2.2.3. Meta-Regression: Late pregnancy smoking abstinence regressed on TCS Category 3 Score

Main results for Model 3, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.5176	0.0929	0.3355	0.6997	5.57	0
TCS Cat3 Total_Select/Tailor	-0.3601	0.1659	-0.6851	-0.035	-2.17	0.0299

Statistics for Model 3

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 4.71, df = 1, p = 0.0299
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0654, Tau = 0.2557, I^2 = 39.29%, Q = 52.71, df = 32, p = 0.0120

Comparison of Model 3 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 3
R^2 analog = 0.25

Table 2.2.4. Meta-Regression: Late pregnancy smoking abstinence regressed on TCS Category 4 Score

Main results for Model 4, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.4707	0.0888	0.2966	0.6448	5.3	0
TCS Cat4 Total_Constructs Measured	-0.3717	0.2454	-0.8527	0.1093	-1.51	0.1299

Statistics for Model 4

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 2.29, df = 1, p = 0.1299
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0880, Tau = 0.2966, I^2 = 47.02%, Q = 60.40, df = 32, p = 0.0018

Comparison of Model 4 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 4
R^2 analog = 0.00 (computed value is 0.00)

Table 2.2.5. Meta-Regression: Late pregnancy smoking abstinence regressed on TCS Category 5 Score

Main results for Model 5, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.5835	0.0988	0.3899	0.777	5.91	0
TCS Cat5 Total_Theory Tested/Refined				-0.3794	0.1311	-0.6363

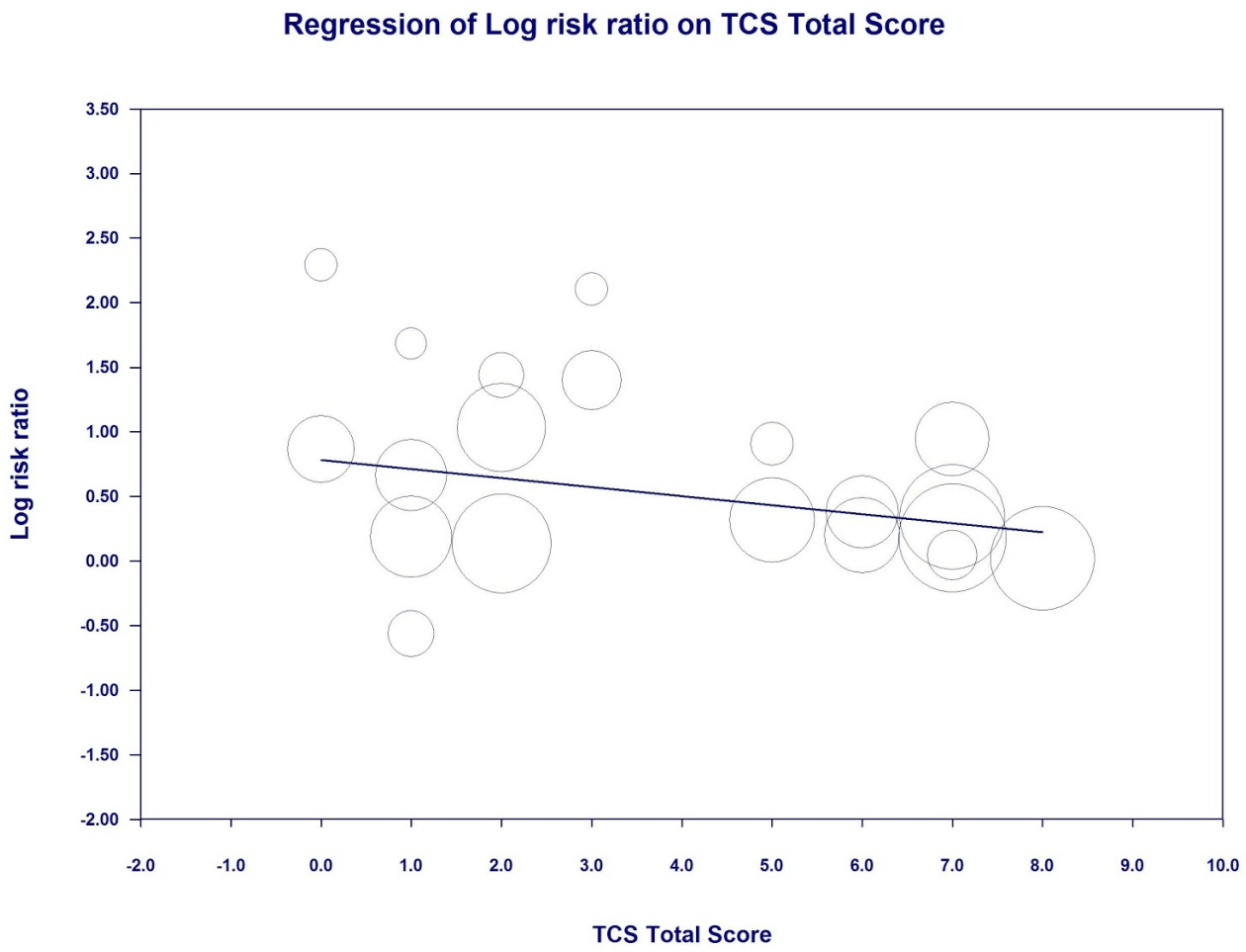
Statistics for Model 5

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 8.38, df = 1, p = 0.0038
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0485, Tau = 0.2202, I^2 = 33.09%, Q = 47.83, df = 32, p = 0.0357

Comparison of Model 5 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 5
R^2 analog = 0.45

Table 2.2.6. Meta-Regression: Late pregnancy smoking abstinence regressed on TCS Total Score



APPENDIX: TABLES AND FIGURES

Aim 3 Tables & Figures

Table 3.1.0. Behavior Change Techniques: Intercoder Reliability and Frequencies

Behavior Change Technique	Associated theory(ies)	Intercoder Reliability (*k*)	Number of studies: Total *K* (out of 38)	Number of studies: Active *K*
1: Provide info on health-behavior link	IMB	0.62	19	12
2: Provide info on consequences (negative)	TRA, TPB, SCogT, & IMB	0.65	16	10
3: Provide info on consequences (positive)	TRA, TPB, SCogT, & IMB	0.71	12	7
4: Provide information about others' approval	TRA, TPB, IMB, SCogT	0.82	1	1
5: Prompt intention formation	TRA, TPB, SCogT, & IMB	0.66	13	7
6: Prompt barrier identification	SCogT	0.71	10	7
7: Provide general encouragement	SCogT	0.69	17	12
8: Set graded tasks	SCogT	0.83	3	2
9: Provide instructions	SCogT	0.63	29	8
10: Model/demonstrate the behavior	SCogT	0.85	6	5
11: Prompt specific goal setting	CT	0.71	25	10
12: Prompt review of behavioral goals	CT	0.7	12	5
13: Prompt self-monitoring of behavior	CT	0.68	7	1
14: Provide feedback on performance	CT	0.73	11	8
15: Provide contingent rewards	OC	0.94	9	9
16: Teach to use prompts/cues	OC	0.73	4	3
17: Agree to behavioral contract	OC	0.91	4	3
18: Prompt practice	OC	0.76	1	1
19: Use follow-up prompts	OC	0.71	10	8
20: Provide opportunity for social comparison	SCogT	0.75	4	3

21: Plan social support/social change	Social support theories	0.71	10	7
22: Prompt identification as role model	Stress & coping theories	0.90	0	0
23: Prompt self-talk	IMB	0.67	7	3
24: Relapse prevention	Relapse prev. therapy	0.73	10	5
25: Stress management	Stress & coping theories	0.71	6	2
26: Motivational interviewing	SCogT, IMB	0.84	11	8
27: Time management	IMB	1.0	0	0
Other			26	

Table 3.1.1. Total Behavior Change Techniques and *Active* Ingredients by Study

Study	Intervention Arm	Total BCT's (out of 27)	Total Active BCT's
Bullock (2009)	I1	3	2
	I2	10	7
	I3	10	7
	Control	3	3
Cinciripini (2000)	I	11	4
	Control	7	7
Cinciripini (2010)	I	17	4
	Control	13	9
Donatelle (2000a)	I	4	3
	Control	1	1
Donatelle (2000b)	I	3	2
	Control	1	1
Donatelle (2000c)	I	3	2
	I2	5	4
	Control	1	1
Dornelas (2006)	I	6	5
	Control	1	1
El-Mohandes, et al (2008)	I	9	9
	Control	0	0
El-Mohandes (2013)	I	3	2
	Control	3	1
Ershoff (1999)	I1	7	2

	I2	13	9
	Control	4	4
Gielen (1997)	I	12	10
	Control	2	2
Hartmann (1995/1996)	I	10	9
	Control	1	1
Heil (2008)	I	4	1
	Control	3	3
Hennrikus (2010)	I	3	3
	Control	1	0
Higgins (2004)	I	3	1
	Control	2	2
Higgins, unpublished	I	4	2
	Control	3	1
Higgins (2014)	I1	4	1
	I2	5	2
	Control	3	3
Kendrick (1995)	I	7	7
	Control	0	0
Malchodi (2003)	I	3	3
	Control	3	0
McBride (1999)	I1	8	4
	I2	8	4
	Control	5	1
McBride (2004)	I1	3	2
	I2	5	4
	Control	1	1

Oncken (2008)	I	11	5
	Control	11	6
Ondersma (2012)	I1	6	6
	I2	2	2
	I3	8	8
	Control	0	0
Parker (2007)	I	6	6
	Control	0	0
Patten (2012)	I	6	4
	Control	2	2
Pbert (2004)	I	4	4
	Control	0	0
Pollak (2007)	I	6	5
	Control	5	1
Rigotti (2006)	I	7	7
	Control	0	0
Secker-Walker (1997)	I	4	4
	Control	1	0
Secker-Walker (1998a)	I	7	7
	Control	1	0
Solomon (2000)	I	4	4
	Control	2	0
Stotts (2002)	I	12	12
	Control	1	0
Stotts (2004)	I	6	6
	Control	0	0
Stotts (2009)	I	3	3

	I2	8	8
	Control	2	0
Strecher (2000)	I	6	6
	Control	1	0
Tsoh (2010)	I	6	6
	Control	0	0
Tuten (2012)	I1	3	3
	I2	4	4
	Control	2	0
Windsor (2011)	I	8	7
	Control	1	1

Table 3.1.2. *Active* Behavior Change Techniques: Effect sizes (compared to control group) and heterogeneity statistics for late pregnancy smoking cessation grouped by inclusion of each intervention technique

Behavior Change Technique	K	Risk Ratio	95% CI	Homog. (Q-statistic)	I2
BCT1: Provide general information	12	1.68*	1.26-2.12	25.73	57.25
BCT2: Provide information on consequences (negative)	10	1.38*	1.08-1.77	15.37	41.45
BCT3: Provide information on consequences (positive)	7	1.03*	0.86-1.24	3.5	0.000
BCT4: Provide information about others' approval	1	NA	NA	NA	NA
BCT5: Prompt intention formation	7	1.24*	1.00-1.53	5.68	0.000
BCT6: Prompt barrier identification	7	1.40	0.97-2.01	15.44	61.14
BCT7: Provide general encouragement	12	1.19	0.99-1.42	14.46	23.93
BCT8: Set graded tasks	2	NA	NA	NA	NA
BCT9: Provide instructions	8	1.51*	1.21-1.89	4.9	0.00
BCT10: Model/demonstrate the behavior	5	1.16	0.94-1.44	3.8	0.000

BCT11: Prompt specific goal setting	10	1.48*	1.17-1.88	15.15	40.58
BCT12: Prompt review of behavioral goals	5	1.20	0.90-1.60	8.24	51.44
BCT13: Prompt self-monitoring of behavior	1	NA	NA	NA	NA
BCT14: Provide feedback on performance	8	1.23	0.97-1.57	6.82	0.000
BCT15: Provide contingent rewards	9	2.82*	2.05-3.88	6.16	0.000
BCT16: Teach to use prompts/cues	3	1.63*	1.03-2.59	1.09	0.000
BCT17: Agree to behavioral contract	3	2.14*	1.29-3.56	1.87	0.000
BCT18: Prompt practice	1	NA	NA	NA	NA
BCT19: Use follow-up prompts	8	1.32	0.97-1.79	17.24	39.77
BCT20: Provide opportunities for social comparison	3	1.22	0.54-2.76	3.32	39.77
BCT21: Plan social support/social change	7	1.14	0.93-1.40	6.90	13.00
BCT22: Prompt identification as role model	0	NA	NA	NA	NA
BCT23: Prompt self-talk	3	1.12	0.85-1.47	2.87	30.29
BCT24: Relapse prevention	5	1.14	0.91-1.43	5.29	24.45

BCT25: Stress management	2	NA	NA	NA	NA
BCT26: Motivational interviewing	8	1.09	0.93-1.29	7.29	3.92
BCT27: Time management	0	NA	NA	NA	NA

Table 3.2.1. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 1

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.37	0.1085	0.1573	0.5826	3.41	0.0007
BCT1: 1	0.1343	0.1704	-0.1996	0.4683	0.79	0.4305

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 0.62, df = 1, p = 0.4305
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0916, Tau = 0.3027, I^2 = 47.86%, Q = 61.37, df = 32, p = 0.0014

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.00 (computed value is -0.05)

Table 3.2.2. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 2

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.477	0.1049	0.2713	0.6826	4.55	0
BCT 2: 1	-0.1427	0.1776	-0.4908	0.2055	-0.8	0.4219

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 0.65, df = 1, p = 0.4219
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0959, Tau = 0.3096, I^2 = 49.00%, Q = 62.75, df = 32, p = 0.0009

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.00 (computed value is -0.09)

Table 3.2.3. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 5

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.4689	0.096	0.2809	0.657	4.89	0
BCT 5: 1	-0.1875	0.2012	-0.5818	0.2068	-0.93	0.3513

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 0.87, df = 1, p = 0.3513
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0944, Tau = 0.3073, I^2 = 48.53%, Q = 62.18, df = 32, p = 0.0011

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.00 (computed value is -0.08)

Table 3.2.4. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 9

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.4127	0.0966	0.2234	0.602	4.27	0
BCT9	0.0464	0.193	-0.3319	0.4247	0.24	0.81

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 0.06, df = 1, p = 0.8100
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0914, Tau = 0.3023, I^2 = 48.13%, Q = 61.70, df = 32, p = 0.0012

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.00 (computed value is -0.04)

Table 3.2.5. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 11

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.4358	0.1056	0.2289	0.6427	4.13	0
BCT11	-0.0226	0.1776	-0.3707	0.3256	-0.13	0.8989

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 0.02, df = 1, p = 0.8989
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0969, Tau = 0.3113, I^2 = 49.17%, Q = 62.95, df = 32, p = 0.0009

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.00 (computed value is -0.11)

Table 3.2.6. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 15

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.2497	0.0686	0.1153	0.3841	3.64	0.0003
BCT15	0.7853	0.187	0.4187	1.1518	4.2	0

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 17.63, df = 1, p =
0.0000
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0244, Tau = 0.1561, I^2 = 20.32%, Q = 40.16, df = 32, p = 0.1524

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.72

Table 3.2.7. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 16

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.4158	0.0867	0.2459	0.5857	4.8	0
BCT16	0.116	0.3243	-0.5196	0.7516	0.36	0.7206

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 0.13, df = 1, p = 0.7206
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0911, Tau = 0.3018, I^2 = 48.70%, Q = 62.37, df = 32, p = 0.0010

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.00 (computed value is -0.04)

Table 3.2.8. Meta-Regression, Late pregnancy smoking cessation regressed on BCT 17

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.3918	0.0842	0.2267	0.5568	4.65	0
BCT17: 1	0.4101	0.3316	-0.2398	1.06	1.24	0.2162

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 1.53, df = 1, p = 0.2162
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0825, Tau = 0.2873, I^2 = 46.39%, Q = 59.69, df = 32, p = 0.0021

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012
Proportion of total between-study variance explained by Model 1
R^2 analog = 0.06

Table 3.3.1. Meta-Regression, Late pregnancy smoking cessation regressed on Total # of BCTs

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	0.6878	0.1831	0.329	1.0467	3.76	0.0002
Active BCTs	-0.0481	0.0297	-0.1062	0.01	-1.62	0.1048

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero

Q = 2.63, df = 1, p = 0.1048

Goodness of fit: Test that unexplained variance is zero

Tau^2 = 0.0899, Tau = 0.2998, I^2 = 47.19%, Q = 60.59, df = 32, p = 0.0017

Comparison of Model 1 with the null model

Total between-study variance (intercept only)

Tau^2 = 0.0876, Tau = 0.2960, I^2 = 47.65%, Q = 63.04, df = 33, p = 0.0012

Proportion of total between-study variance explained by Model 1

R^2 analog = 0.00 (computed value is -0.03)

Table 3.4.1. Late pregnancy smoking cessation within interventions providing contingent rewards (Moderator: BCT 1 [y/n])

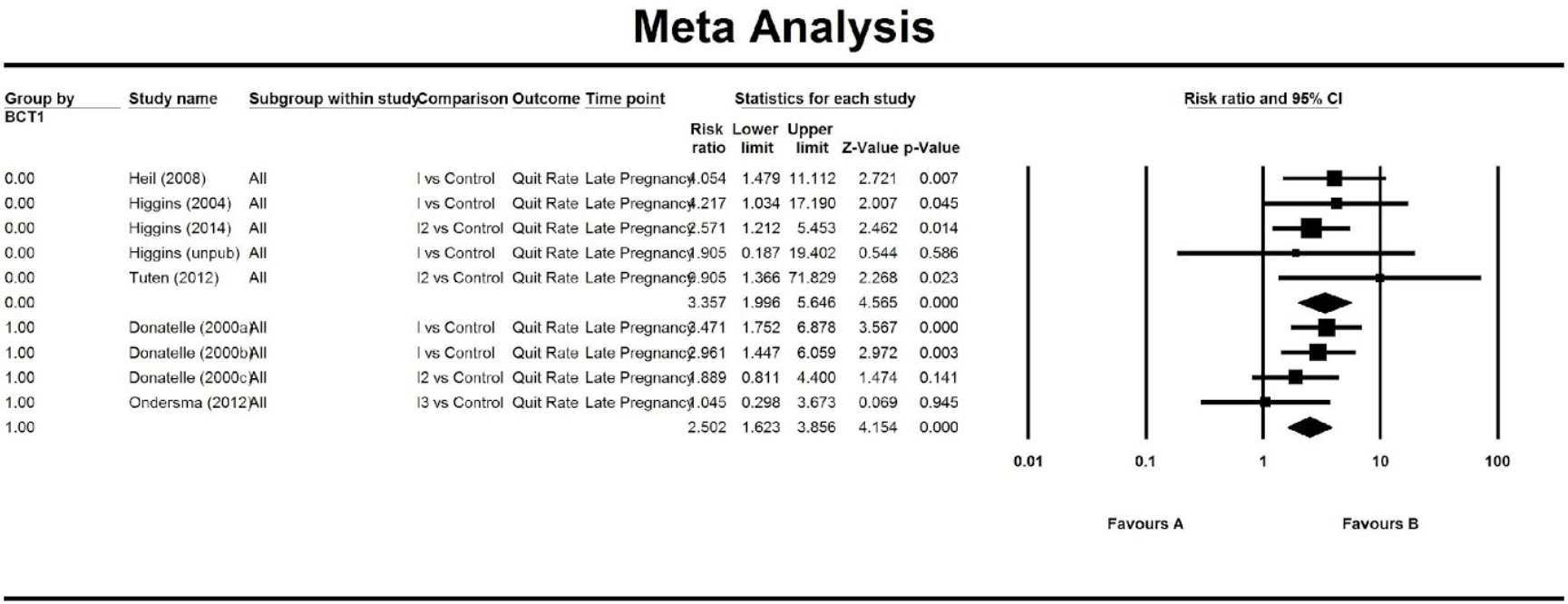


Table 3.4.2. Late pregnancy smoking cessation within interventions providing contingent rewards (Moderator: BCT 14 [y/n])

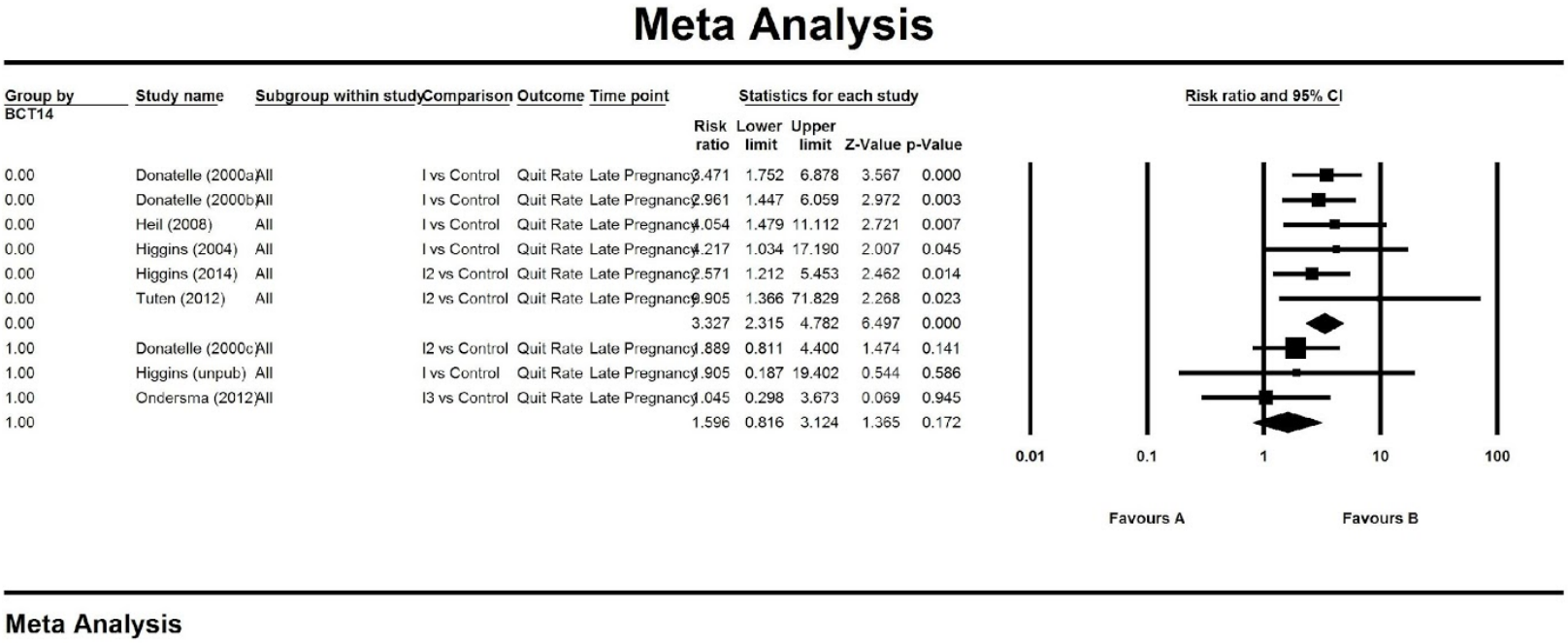


Table 3.4.3. Late pregnancy smoking cessation within interventions providing contingent rewards (Moderator: Assessed smoking in social network [y/n])

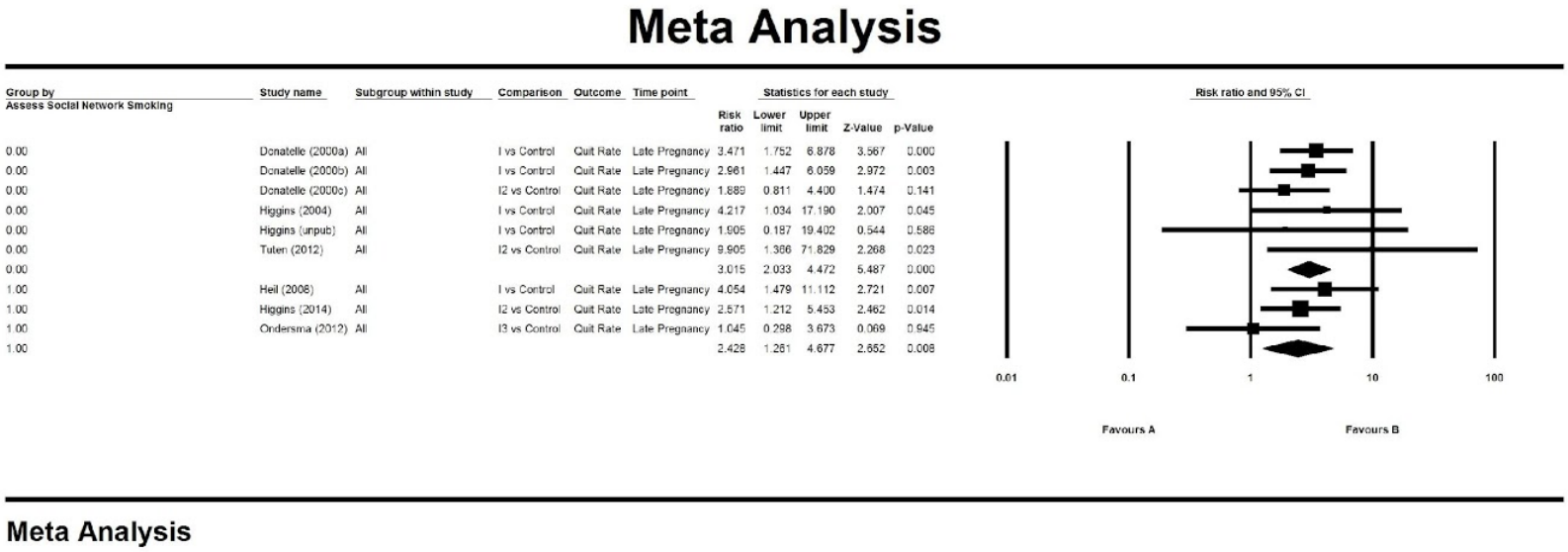


Table 3.4.4. Late pregnancy smoking cessation within interventions providing contingent rewards (Moderator: Referral to community resources [y/n])

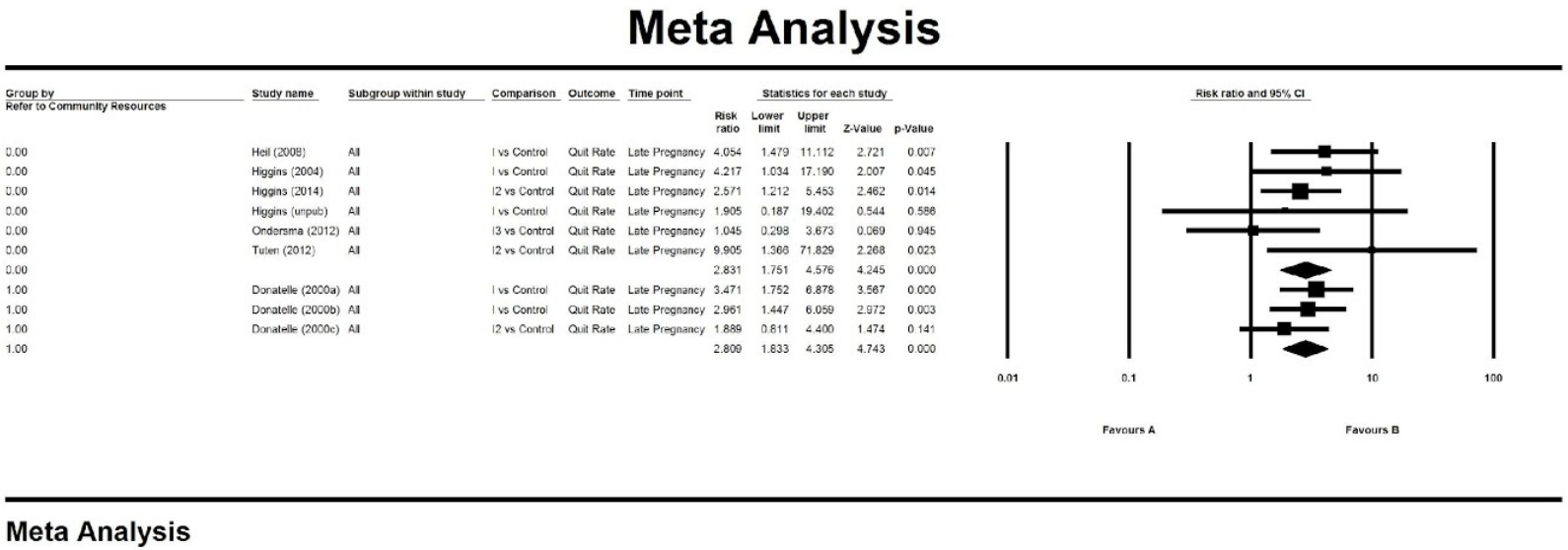
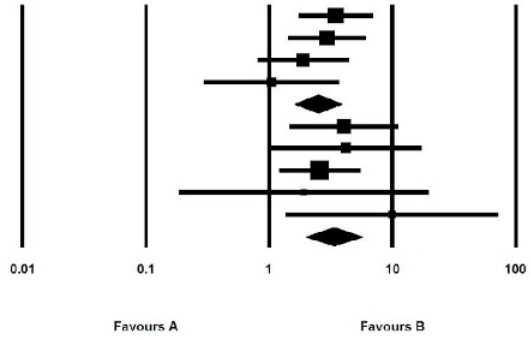


Table 3.4.5. Late pregnancy smoking cessation within interventions providing contingent rewards (Moderator: Assessed smoking in social network [y/n])

Meta Analysis

Group by Contact Intensity	Study name	Subgroup within study	Comparison	Outcome	Time point	Risk ratio	Lower limit	Upper limit	Z-Value	p-Value
2.00	Donatelle (2000a)	All	I vs Control	Quit Rate	Late Pregnancy	3.471	1.752	6.878	3.567	0.000
2.00	Donatelle (2000b)	All	I vs Control	Quit Rate	Late Pregnancy	2.961	1.447	6.059	2.972	0.003
2.00	Donatelle (2000c)	All	I2 vs Control	Quit Rate	Late Pregnancy	1.889	0.811	4.400	1.474	0.141
2.00	Ondersma (2012)	All	I3 vs Control	Quit Rate	Late Pregnancy	1.045	0.298	3.673	0.069	0.945
2.00						2.502	1.623	3.856	4.154	0.000
3.00	Heil (2008)	All	I vs Control	Quit Rate	Late Pregnancy	4.054	1.479	11.112	2.721	0.007
3.00	Higgins (2004)	All	I vs Control	Quit Rate	Late Pregnancy	4.217	1.034	17.190	2.007	0.045
3.00	Higgins (2014)	All	I2 vs Control	Quit Rate	Late Pregnancy	2.571	1.212	5.453	2.462	0.014
3.00	Higgins (unpub)	All	I vs Control	Quit Rate	Late Pregnancy	1.905	0.187	19.402	0.544	0.586
3.00	Tuten (2012)	All	I2 vs Control	Quit Rate	Late Pregnancy	9.905	1.366	71.829	2.268	0.023
3.00						3.357	1.996	5.646	4.565	0.000



Meta Analysis

Table 3.4.6. Late pregnancy smoking cessation regressed on gestational age (among studies that provided contingent rewards)

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	1.7641	0.6771	0.437	3.0912	2.61	0.0092
Gestational Age at entry	-0.0494	0.045	-0.1376	0.0389	-1.1	0.2728

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero

Q = 1.20, df = 1, p = 0.2728

Goodness of fit: Test that unexplained variance is zero

Tau^2 = 0.0000, Tau = 0.0000, I^2 = 0.00%, Q = 4.84, df = 6, p = 0.5642

Comparison of Model 1 with the null model

Total between-study variance (intercept only)

Tau^2 = 0.0000, Tau = 0.0000, I^2 = 0.00%, Q = 6.04, df = 7, p = 0.5346

Proportion of total between-study variance explained by Model 1

R^2 analog = 0.00

Table 3.4.7. Late pregnancy smoking cessation regressed on baseline smoking [cig/day] (among studies that provided contingent rewards)

Main results for Model 1, Random effects (MM), Z-Distribution, Log risk ratio

Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value
Intercept	-0.4692	1.0214	-2.4712	1.5327	-0.46	0.646
Cig per day at baseline	0.1655	0.1073	-0.0448	0.3757	1.54	0.1229

Statistics for Model 1

Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero
Q = 2.38, df = 1, p = 0.1229
Goodness of fit: Test that unexplained variance is zero
Tau^2 = 0.0000, Tau = 0.0000, I^2 = 0.00%, Q = 2.42, df = 3, p = 0.4893

Comparison of Model 1 with the null model

Total between-study variance (intercept only)
Tau^2 = 0.0700, Tau = 0.2646, I^2 = 16.72%, Q = 4.80, df = 4, p = 0.3081
Proportion of total between-study variance explained by Model 1
R^2 analog = 1.00

Table 3.5.1 Ratio of Effective BCT's to Active BCT's

Behavior Change Technique	# of studies (*K*): Active	# of studies: Significant RR	Ratio of Effective BCTs: Active BCTs
1: Provide info on health-behavior link*	12	3	1:4
2: Provide info on consequences (negative)*	10	1	1:10
3: Provide info on consequences (positive)	7	0	0:7
4: Provide information about others' approval	1	0	0:1
5: Prompt intention formation*	7	0	0:7
6: Prompt barrier identification	7	1	1:7
7: Provide general encouragement	12	1	1:12
8: Set graded tasks	2	1	1:2
9: Provide instructions*	8	1	1:8
10: Model/demonstrate the behavior	5	0	0:5
11: Prompt specific goal setting*	10	1	1:10
12: Prompt review of behavioral goals	5	0	0:5
13: Prompt self-monitoring of behavior	1	0	0:1
14: Provide feedback on performance	8	0	0:8
15: Provide contingent rewards*	9	6	2:3
16: Teach to use prompts/cues*	3	0	0:3
17: Agree to behavioral contract*	3	1	1:3
18: Prompt practice	1	0	0:1
19: Use follow-up prompts	8	1	1:8
20: Provide opportunity for social comparison	3	0	0:3
21: Plan social support/social change	7	0	0:7
22: Prompt identification as role model	0	0	-

23: Prompt self-talk	3	0	0:3
24: Relapse prevention	5	0	0:5
25: Stress management	2	0	0:2
26: Motivational interviewing	8	0	0:8
27: Time management	0	0	-

Note: BCTs marked by an asterisk (*) indicate that the risk ratio for that subgroup of interventions was statistically significant in subgroup analyses.

APPENDIX: SUPPLEMENTARY MATERIALS

Coding Documents & Instructions

DATA EXTRACTION FORM

Study ID: **Article ID (if needed):**

Study Brief Citation (Author, year):

Name of review author completing this form:

Date form completed:

Notes *(Unpublished – for own use) Eg. References to be followed up, questions or need for clarity, etc.*

STUDY CHARACTERISTICS

Study Characteristics	Description	Answer/Code
Source of study?	What is the source of the study? Write the number in the box to the right. 1 = Peer-reviewed journal 2 = Non-peer-reviewed journal 3 = Government Report 9 = Other	
Journal Name	If the study was published in a journal what is the name of the journal?	
Year of publication?	What year did the study actually appear in print?	
Study Purpose?	What was the purpose of the study, as stated by the authors? (Write in the box to the right)	
Accuracy of stated purpose?	Did the author's statement of the study's purpose accurately represent the study, as it was actually carried out? (Y/N) If no, briefly describe why.	
Funding?	Who funded the study? 1 = Federal agency 2 = State agency 3 = Local agency	

	4 = Foundation 5 = University supported 9 = Other 0 = No sources listed	
Consumer Involvement?	Was there consumer involvement in the study and/or intervention? 1 = Yes, in design of study and/or intervention 2 = Yes, in delivery of intervention 3 = Yes, in evaluation of intervention 4 = Yes, in interpretation of study findings 5 = Yes, in multiple areas specified above 6 = No	
Conflict of Interest?	Did the authors report any conflicts of Interest? 0 = No COI reported 1 = Yes, COI reported 9 = No mention of COI	
Geographical Setting?	What was the geographical setting of the intervention? 0 = Not reported 1 = Urban 2 = Rural 3 = Suburban 9 = Other	

METHODS

Criteria	Description	Answer/Code
Study Design	1 = Randomized Controlled Trial 2 = Cluster Randomized Trial 3 = Randomized Crossover Trial	
No. of sites	At how many sites did data collection take place?	
Data collection timeline	What year(s) were data collected?	
IRB Approval	Did study mention IRB approval? (Y/N)	
Informed Consent	Was Informed Consent obtained from participants? (Y/N/Unclear)	
Recruitment Methods	How were potential participants approached and invited to participate?	
Inclusion/Exclusion Criteria	What were the inclusion/exclusion criteria for participation in study?	
Statistical Methods	What statistical methods were used to analyze data?	

Appropriateness of Statistical Methods	Were statistical methods appropriate? (Y/N/Unclear) Rationale?	

Quality Assessment

Criteria	Description	Answer/Code

Adequate Sequence Generation?	**Yes**: The investigators describe a random component in the sequence generation, such as: ● Referring to a random number table ● Using a computer random number generator ● Coin tossing ● Shuffling cards or envelopes ● Throwing dice ● Drawing of lots ● Minimization w/ or w/out a random element. **No:** The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example: ● Sequence generated by odd or even date of birth; ● Sequence generated by some rule based on date (or day) of admission; ● Sequence generated by some rule based on hospital or clinic record number. Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgement or some method of non-random categorization of participants, for example: ● Allocation by judgement of the clinician; ● Allocation by preference of the participant; ● Allocation based on the results of a laboratory test or a series of tests; ● Allocation by availability of the intervention **Unclear:** Insufficient information about the sequence generation to permit judgment of **yes** or **no**.	

Allocation Concealment?*	**Yes:** Participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: ● Central allocation (including telephone, web-based, and pharmacy-controlled, randomization); ● Sequentially numbered drug containers of identical appearance; ● Sequentially numbered, opaque, sealed envelopes. **No:** Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: ● Using an open random allocation schedule (e.g. a list of random numbers); ● Assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed); ● Alternation or rotation; ● Date of birth; ● Case record number; ● Any other explicitly unconcealed procedure. **Unclear:** Any one of the following: ● Insufficient information to permit judgement of 'Yes' or 'No'; ● The study did not address this outcome	**Note: it is rarely feasible in psychosocial interventions to blind women or the intervention providers to group allocation.*

BLINDING OF PARTICIPANTS, PERSONNEL AND OUTCOME ASSESSORS. Was knowledge of the allocated interventions adequately prevented during the study?	**Yes**: Any one of the following: ● No blinding, but the review authors judge that the outcome and the outcome measurement are not likely to be influenced by lack of blinding; ● Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken; ● Either participants or some key study personnel were not blinded, but outcome assessment was blinded and the non- blinding of others unlikely to introduce bias. **No:** Any one of the following: ● No blinding or incomplete blinding, and the outcome or outcome measurement is likely to be influenced by lack of blinding. ● Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken; ● Either participants or some key study personnel were not blinded, and the non-blinding of others likely to introduce bias. **Unclear**: Any one of the following: ● Insufficient information to permit judgement of 'Yes' or 'No'; ● The study did not address this outcome.	
Incomplete Outcome Data Addressed?	**Yes:** Any one of the following: ● No missing outcome data;	

	● Reasons for missing outcome data unlikely to be related to true outcome; ● Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; ● For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate; ● For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size; ● Missing data have been imputed using appropriate methods. **No:** Any one of the following: ● Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; ● For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate; ● For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size;	

	● 'As-treated' analysis done with substantial departure of the intervention received from that assigned at randomization; ● Potentially inappropriate application of simple imputation. **Unclear:** Any one of the following: ● Insufficient reporting of attrition/exclusions to permit judgement of 'Yes' or 'No' (e.g. number randomized not stated, no reasons for missing data provided); ● The study did not address this outcome.	
Free of Selective Outcome Reporting?	**Yes:** Any of the following: ● The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way; ● The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified. **No**: Any one of the following: ● Not all of the study's pre-specified primary outcomes have been reported; ● One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; ● One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect);	

	● One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; ● The study report fails to include results for a key outcome that would be expected to have been reported for such a study. **Unclear:** Insufficient information to permit judgement of 'Yes' or 'No'. It is likely that the majority of studies will fall into this category.	
Free of other forms of bias?	**Yes:** The study appears to be free of other sources of bias. **No:** There is at least one important risk of bias. For example, the study: ● Had a potential source of bias related to the specific study design used; or ● Stopped early due to some data-dependent process (including a formal-stopping rule); or ● Had extreme baseline imbalance; or ● Has been claimed to have been fraudulent; or ● Had some other problem. **Unclear:** There may be a risk of bias, but there is either: ● Insufficient information to assess whether an important risk of bias exists; or ● Insufficient rationale or evidence that an identified problem will introduce bias.	

Criteria for Summary Assessments for Risk of Bias for Each Important Outcome Across Domains (w/in trials) & Across Trials

Risk of Bias	Interpretation	Within a Trial (Across Domains)	Across Trials
Low risk of bias	Bias, if present, is unlikely to alter the results seriously	Low risk of bias for all key domains	Most information is from trials at low risk of bias
Unclear risk of bias	A risk of bias that raises some doubt about the results	Low or unclear risk of bias for all key domains	Most information is from trials at low or unclear risk of bias
High risk of bias	Bias may alter the results seriously	High risk of bias for one or more key domains	The proportion of information from trials at high risk of bias is sufficient to affect the interpretation of results

PARTICIPANTS

Criteria	Description	Answer/Code
Number eligible	Total number of eligible participants	
Number randomized to intervention	Total number of participants randomized to intervention	
Number randomized to control group	Total number of participants randomized to control group	
Number included in analysis	Total number of participants included in analysis	
Age	What was the range of participants' ages? What was the mean & SD of participants' ages?	Range: Mean(SD):
Race/Ethnicity	What was the race/ethnicity of participants?	% Asian_____ % Black _____ %Hispanic/Latina_____ % Native American/Alaskan Native_____ % White_____ % Other _____
Medicaid eligible?	Were participants a Medicaid eligible population (Yes/No)?	If yes, what % of sample was Medicaid eligible?
Education Level	What was the education level of study participants?	% No High School Diploma: % High School Diploma only:

		% Undergrad degree or higher:
Geographical Location	Did the intervention target a specific geographical region (Ye s/No) If so, what type of geographical region did the study target? 1 = Urban 2 = Rural 3 = Other (Specify)	
Income Level	What was the range of income levels? What was the mean income level?	Range = Mean (SD) =
Poverty	What % of study participants fell below poverty level?	
Marital Status	What % of study participants were: Married? Divorced? Single?	% Married_____ % Divorced _____ % Single_____
Pregnancy Timing	How far along (in weeks) were study participants when intervention began?	Mean (SD) in Weeks = OR % in 1st trimester: % in 2nd trimester: % in 3rd trimester:
Health Status	What was the health of participants?	

	1 = Intervention targeted generally healthy pregnant women 2 = Intervention targeted women with specific health and/or mental health condition(s).	
Parity		
Unintended Pregnancy	1= Yes 0 = No	
Nicotine Dependence	Did study provide data on participants' smoking habits? (Y/N) If so: How long has participant smoked? How many cigarettes/day?	

INTERVENTION CHARACTERISTICS

Variable	Description	Answer/Code
Intervention Type	Was the intervention: 1 = Single-component intervention (only one main strategy) 2 = Multiple-component intervention (several strategies offered to all women) 3 = Tailored intervention (additional strategies available for some women)	
Pharmacological assistance?	Did the intervention include a pharmacological cessation aid (e.g., nicotine patch, nicotine gum medication assistance, etc) (Y/N) If yes, specify:	
Intervention Target	Did the intervention target any outcomes other than smoking? (Y/N)	If Yes, what other outcomes were targeted?

Comparison Group Type	What type of comparison groups were used? 1 = Usual care or no additional intervention 2 = Less intensive version of intervention 3 = Alternative intervention of similar intensity	
Delivery of intervention	**Frequency**: Total # of sessions/appointments/meetings Length of contact: How long (in minutes) was each **Duration**: Total length (in weeks) of intervention:	
Deliverer of intervention	Who delivered the intervention? 0 = Not specified 1 = Doctor 2 = Nurse(s) 3 = Mental health professional 4 = Community health worker 5 = Health educator 6 = Peer-led 9 = Other	

Setting of Intervention	Where did the intervention take place? 0 = Not reported 1 = Hospital 2 = Public/community clinic 3 = Private doctor's office 4 = Home 5 = Community 9 = Other	
Group or Individual?	Was the intervention delivered *primarily* in a one-on-one (individual) setting or in a group setting? 1 = Individual 2 = Group	
Part of prenatal care?	**Yes:** the intervention was included/embedded as part of routine prenatal care **No:** the intervention was separate from routine prenatal care	
Intervention Component(s)	Intervention Component(s) Included in *each arm of the study* (choose all that apply for the control group and again for the intervention group; use separate coding document provided):	

	1 = Provide information on behavior– health link 2 = provide information on consequences 3 = provide information about others' approval 4 = prompt intention formation 5 = prompt barrier identification 6 = provide general encouragement 7 = set graded tasks 8 = provide instruction 9 = model/ demonstrate the behavior 10 = prompt specific goal setting 11 = prompt review of behavioral goals 12 = prompt self-monitoring of behavior 13 = provide feedback on performance 14 = provide contingent rewards 15 = teach to use prompts/cues 16 = agree a behavioral contract 17 = prompt practice 18 = use of follow-up prompts 19 = provide opportunities for social comparison 20 = plan social support/social change 21 = prompt identification as role model/ position advocate 22 = prompt self talk	

	23 = relapse prevention 24 = stress management 25 = motivational interviewing 26 = time management. 27 = other (specify)	
Number of Intervention Components	How many of the above techniques did the control arm employ? How many of the above techniques did the intervention arm employ? How many active techniques did the study employ?	_____ in control arm _____ in intervention arm _____ active techniques
Process evaluation?	**Yes:** the intervention included process evaluation measures **No:** the intervention did not include process evaluation measures	
Fidelity/integrity?	Was the intervention delivered as described? (Y/N/Unclear)	

OUTCOMES

Principal and secondary outcome measures of interest (*operationalize).*
Smoking point prevalence:

For each outcome:

Methods of assessing outcome measures *(e.g, phone survey, questionnaire, physical measurements)*

Validity and reliability of outcome measures

Methods of follow-up for non-respondents

Timing of outcome assessment *(including frequency, length of follow up (for each outcome))*

Adverse events (*e.g complaints, levels of dissatisfaction, adverse incidents, side effects))*

RESULTS

Dichotomous outcomes

Outcome	Timing of outcome assessment (days/months)	Intervention group*		Control group		Notes
		Observed (n)	Total (N)	Observed (n)	Total (N)	

**Note: add additional columns if there is more than one intervention group, e.g. Intervention Group A, Intervention Group B…*

Study ID	Biochemically validated (Y/N)	Timing of outcome assessment	Total N	Total Tx Group N	%	N	Total Control N	%	N

Continuous outcomes

Outcome	Timing of outcome assessment (days/months)	Intervention group			Control group			Notes
		*Mean / Mean change	Standard deviation	N	*Mean / Mean change	Standard deviation	N	

THEORY CODING SCHEME (TCS)

(Complete Table Below)

Item TCS1	**Item Description** Explicit mention of the use of health behavior theory	**Item Definition** The study **explicitly mentioned** using a health behavior theory (or model), defined as "a set of interrelated concepts, definitions, and propositions that presents a *systematic* view of events or situations by specifying relations among variables in order to *explain* and *predict* events or situations" (Glanz et al., 1997, p. 21). *Note: this is an independent assessment from their **actual use** of theory.	**Examples** Health Belief Model, Theory of Planned Behavior, Social Cognitive Theory, Transtheoretical/ Stages of Change Model	**Yes/No/Don't Know**	**Supporting evidence** Name of theory mentioned?
TCS2	Targeted constructs mentioned as predictors of behavior.	1) The study explicitly mentioned how targeted constructs are theorized to predict behavior, where "targeted constructs" refer to theoretical constructs that the intervention is hypothesized to change. **AND** 2) The study provided evidence that the construct targeted construct relates to behavior in the introduction or methods section (not discussion section).	Self-efficacy, perceived risk/threat, social support, knowledge, intentions		
TCS3	Intervention based on a single theory	The intervention is based on a single theory, rather than a combination of theories or theory and predictors.			
TCS4	Theory used to select participants	Participants were screened/selected based on achieving a particular score/level on a theory-relevant construct/predictor.	Selecting participants with low levels of social support		What construct or predictor? What threshold or level was used?

TCS5	Theory used to select and/or develop intervention techniques	The intervention techniques are explicitly based on a theory or predictor or combination of theories or predictors.			What theory or predictor or combination(s)?
TCS6	Theory used to tailor intervention techniques to participants	The intervention differs for different sub-groups that vary on a psychological construct or predictor at baseline.	Cessation advice tailored to participants' readiness for change on TTM measure.		What construct or predictor? What were the groups?
TCS7	**ALL** intervention techniques are explicitly linked to at least one theory-relevant construct/predictor	Each intervention technique is explicitly linked to at least one theory-relevant construct/predictor.			List all intervention techniques and the constructs/predictors they are linked to.
TCS8	**At least one**, but not all, of the intervention techniques are explicitly linked to at least one theory-relevant construct/ predictor.	At least one, but not all, of the intervention techniques are explicitly linked to at least one theory-relevant construct/ predictor.			List all intervention techniques and the constructs/predictors they are linked to.
TCS9	Group of techniques are linked to a group of constructs/ predictors	A cluster of techniques is linked to a cluster of constructs/predictors			List all clusters of techniques & clusters of constructs/ predictors they are linked to.

TCS10	**All** theory-relevant constructs/predictors are explicitly linked to at least one intervention technique	Every theoretical construct within a stated theory, or every stated predictor (see item 5), is linked to at least one intervention technique.			List all constructs/predictors and the intervention technique they are linked to.
TCS11	**At least one**, but not all, of the theory relevant constructs/predictors are explicitly linked to at least one intervention technique.	At least one, but not all, of the theoretical constructs within a stated theory or at least one, but not all, of the stated predictors (see item 5) are linked to at least one intervention technique.			List all constructs/predictors and the intervention technique they are linked to.
TCS12	Theory-relevant constructs/ predictors are measured	a) At least one construct of theory (or predictor) mentioned in relation to the intervention is measured POST-INTERVENTION. **OR** b) At least one construct of theory (or predictor) mentioned in relation to the intervention is measured PRE AND POST-INTERVENTION.			List constructs/predictors and when they were measured

TCS13	Quality of Measures	a) All of the measures of theory relevant constructs/predictors had some evidence for their reliability b) At least one, but not all, of the measures of theory relevant constructs/predictors had some evidence for their reliability c) All of the measures of theory relevant constructs/predictors have been previously validated d) At least one, but not all, of the measures of theory relevant constructs/predictors have been previously validated e) The behavior measure had some evidence for its reliability f) The behavior measure has been previously validated			
TCS14	Randomization of participants to condition	a) Do the authors claim randomization? b) Is a method of random allocation to condition described (e.g., random number generator; coin toss) c) Was the success of randomization tested? d) Was the randomization successful (or baseline differences between intervention and control group statistically controlled)?			
TCS15	Changes in measured theory-relevant constructs/predictors	The intervention leads to sig. change in at least one theory-relevant construct/predictor (vs.control group) in favor of the intervention group.			What construct(s) and/or predictors?

TCS16	Mediational analysis of construct/s / predictors	In addition to 14, do the following effects emerge?: a) Mediator predicts DV? (or change in mediator leads to change in DV) b) Mediator predicts DV (when controlling for IV)? c) Intervention does not predict DV (when controlling for mediator)? d) Mediated effect statistically significant?			
TCS17	Results discussed in relation to theory	Results are discussed in terms of the theoretical basis of the intervention			
TCS18	Appropriate support for theory	Support for the theory is based on appropriate mediation **OR** refutation of the theory is based on obtaining appropriate null effects (i.e. changing behavior without changing the theory-relevant constructs).			
TCS19	Results used to refine theory	The authors attempt to refine the theory upon which the intervention was based by either: a) adding or removing constructs to the theory, **or** b) specifying that the interrelationships between the theoretical constructs should be changed and spelling out which relationships should be changed.			a) Constructs added or removed from theory: b) Interrelationships between the theoretical constructs to be changed:

BEHAVIOR CHANGE TECHNIQUES: CODING MANUAL

Adapted version of the coding manual from: Abraham, C. & Michie, S (2007). *A taxonomy of behavior change techniques used in interventions: The Coding Manual.*

BCT Coding instructions

Carefully read the taxonomy before coding materials for behavioral change techniques. Discuss the techniques with co-coders to make sure that these are interpreted similarly by all coders. Always practice coding on practice materials comparable to your final materials and discuss these coding results before starting actual coding.

Suggestions for optimal coding (for coders individually):

- Read the published study once before actual coding. Highlight or underline relevant sections.
- Scan the different techniques (and associated definitions) presented in the coding table.
- Start coding the relevant sections using the scoring table (below). In case of any doubt between techniques, always turn to the description of the techniques presented in this document.
- FOR EACH ARM OF THE STUDY: If a technique is identified, code 1 for ‘yes’. If a technique has been ruled out, code 0 for ‘no’. If unsure, make a note and return to the item to make a final judgment. Make sure all items are coded 0 or 1 before assessing intercoder reliability.
- If techniques that are not included in the taxonomy are identified in the published article, make note of them (highlight or underline the relevant text; code ‘yes’ for ‘other’). When all studies have been coded, evaluate the techniques identified as ‘other’ to determine if additional categories should be added to capture these techniques.

Note:

- Most BCT's will be found in the introduction and methods sections. If only mentioned in the discussion, make sure there is evidence that the technique was actually used and not simply discussed in relation to the results.

BCT Coding Form

Behavior Change Technique	Definition	Control arm? Y/N	Treatment arm? Y/N	Notes & text (page #, keywords)
Provide information on health-behavior link	Provide general information about behavior risk (e.g., susceptibility to poor health outcomes for mother or fetus)			
Provide information on consequences (negative)	Provide information about the costs of action or inaction, focusing on what negative outcomes could happen if the person does or does not perform the behavior.			
Provide information on consequences (positive)	Provide information about the benefits of action or inaction, focusing on what positive outcomes could happen if the person does or does not perform the behavior.			
Provide information about others' approval	Provide information about what others think about the person's behavior and whether others will approve or disapprove of proposed behavior change(s)			

Prompt intention formation	Encouraging the person to decide to act or set a general goal (e.g., to quit or cut back on smoking)
Prompt barrier identification	Identify barriers to performing the barriers and plan ways of overcoming them
Provide general encouragement	Provide praise or reward for effort or performance without this being contingent on specified behaviors or standards of performance
Set graded tasks	Set easy tasks, and increase difficulty until target behavior is achieved
Provide instructions	Advise or agree on how to perform the behavior
Model/demonstrate the behavior	An expert shows the person how to perform a behavior (may be in person or on a video/computer)
Prompt specific goal setting	Set or agree on a goal defined in terms of the behavior to be achieved
Prompt review of behavioral goals	Review behavior goal(s) jointly with the person and consider modifying goal(s) or behavior change strategy depending on achievement. This may result in setting a new goal instead of (or in addition to) the initial goal, or to no change in goals.

Prompt self-monitoring of behavior	Establish a method for the person to monitor and record their behavior(s) as part of a behavior change strategy
Provide feedback on performance	Monitor and provide informative or evaluative feedback on performance of behavior (e.g., form, frequency, duration, intensity, etc)
Provide contingent rewards	Provide praise, encouragement, or material rewards that are explicitly linked to the achievement of specified behaviors
Teach to use prompts/cues	Teach the person to identify environmental cues that can be used to remind them to perform a behavior, including times of day or elements of context
Agree to behavioral contract	Create a written or verbal specification of the behavior to be performed, agreed on by the person, and witnessed by another person (may be the intervention deliverer)
Prompt practice	Prompt the person to rehearse the behavior and/or preparatory behaviors

Use follow-up prompts	Briefly contacting the person again after the primary intervention is complete; not reaching the level of social support
Provide opportunities for social comparison	Facilitate observation of non-expert others' performance of the behavior (e.g., in a group or using a video)
Plan social support/social change	Prompting consideration of how others could change their behavior to offer the person help (instrumental support) or encouragement (emotional support), including buddy systems and partner support
Prompt identification as role model	Indicating how the person may be an example to others and influence their behavior or provide an opportunity for the person to set a good example
Prompt self-talk	Encourage use of self-instruction and self-encouragement to support action (aloud or silently)
Relapse prevention	Following initial behavior change, help identify situations likely to result in readopting risk behaviors or failure to maintain new behaviors, and help the person plan to avoid or manage these situations

Stress management	May involve a variety of specific techniques that do not target the behavior but seek to reduce anxiety and stress
Motivational interviewing	Prompting the person to provide self-motivating statements and evaluations of their own behavior to minimize resistance to change
Time management	Helping the person make time for the behavior (e.g., fitting it into daily schedule/routine)
Other	Describe the technique that was identified in the text but not included in the taxonomy

DEFINITIONS: THEORETICAL DOMAINS AND CONSTRUCTS

Domain	Definition	Constructs & related constructs
Behavioral regulation	An awareness of the existence of something.	Knowledge Procedural knowledge Knowledge of task environment
Beliefs about capabilities	Acceptance of the truth, reality, or validity about an ability, talent, or facility that a person can put to constructive use.	Self-confidence Perceived competence Self-efficacy Perceived behavioral control Beliefs Self esteem Empowerment Professional confidence
Beliefs about consequences	Acceptance of the truth, reality, or validity about outcomes of a behavior in a given situation.	Outcome expectancies Beliefs Consequents Anticipated regret
Emotions	A complex reaction pattern, involving experiential, behavioral, and physiological elements, by which the individual attempts to deal with a personally significant matter or event.	Anxiety Fear Affect Stress Depression Positive/negative affect Burnout
Environmental context and resources	Any circumstance of a person's situation or environment that discourages or encourages the development of skills and abilities,	Environmental stressors Resources/material resources Barriers and facilitators Organizational culture/climate Person X environment interaction

	independence, social competence, and adaptive behavior.	Salient events/critical incidents
Goals	Mental representations of outcomes or end states that an individual wants to achieve.	Goals (distal/proximal) Goal priority Goal/target setting Goals (autonomous/controlled) Action planning Implementation intention
Intentions	A conscious decision to perform a behavior or a resolve to act in a certain way.	Stability of intentions Stages of change
Knowledge	An awareness of the existence of something.	Knowledge Procedural knowledge Knowledge of task environment
Memory, attention, and decision processes	The ability to retain information, focus selectively on aspects of the environment, and choose between two or more alternatives.	Memory Attention Attention control Decision making Cognitive overload
Optimism	The attitude that outcomes will be positive and that people's wishes or aims will ultimately be fulfilled.	Optimism Pessimism Unrealistic optimism Identity
Reinforcement	Increasing the probability of a response by arranging a dependent relationship, or contingency, between the response and a given stimulus.	Rewards Incentives Punishment Sanctions Contingencies Reinforcement Consequents
Skills	An ability or proficiency acquired through training and/or practice	Skills Skills development

		Competence Ability Interpersonal Skills Practice Skills Assessment
Social Influences	Those interpersonal processes that can cause individuals to change their thoughts, feelings, or behaviors.	Social pressure Social norms Group conformity Group norms Social support Intergroup conflict Power Group identity
Social role/identity	A coherent set of behaviors and displayed personal qualities of an individual in a social setting	Professional identity Professional role Social identity Identity Group identity Leadership Organizational commitment